



Industria y Comercio
S U P E R I N T E N D E N C I A

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-177590- -00000-0000

Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 563

Act. 411 PRESENTACION Folios: 563

DELEGATURA DE PROPIEDAD INDUSTRIAL

División de Nuevas Creaciones

**SOLICITUD
PATENTE DE INVENCION**

21. EXPEDIENTE No. _____

54. TÍTULO FRACCIONAMIENTO DE LA BIOMASA MICROBIANA QUE
CONTIENE ACEITE

51. CLASIFICACIÓN INTERNACIONAL _____

71. SOLICITANTE SOLAZYME, INC

DOMICILIO SOUTH SAN FRANCISCO, CALIFORNIA, ESTADOS UNIDOS

74. APODERADO ALVARO CORDERA ORDOÑEZ

22. BOGOTÁ, D.C., Diciembre 22, 2011

PET

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-177590-00000-0000

Fecha: 2011-12-23 12:00:23
Tra. 11 PATENTEIYII
Act. 411 PRESENTACIONDep. 2020 DIR.NUEVASCR
Eve: 379 FASENACIONALII
Folios: 563

icación

Industria y Comercio
SUPERINTENDENCIA

Act. 411 PRESENTACION Folios: 562

FORMULARIO ÚNICO DE SOLICITUD DE PATENTE
PCT
ENTRADA EN FASE NACIONAL

①

SOLICITUD DE :

☒

Patente de Invención

☐

Patente de Modelo de Utilidad

☐

Capítulo I.

☒

Capítulo II.

②

SOLICITANTE
(71)

Nombre: SOLAZYME, INC.

Dirección: 225 Gateway Boulevard, South San Francisco,
California 94080, Estados Unidos

Nacionalidad o Domicilio: Estados Unidos de América

Lugar de Constitución: Estados Unidos de América

Teléfono:

Fax:

E-Mail

IDENTIFICACION

C.C. ☐

NIT

☐C.E. ☐

Otro

☐

Cual

Número

③

REPRESENTANTE
O APODERADO

Nombre: ALVARO CORREA ORDOÑEZ

Dirección: Avenida 82 No. 10-62 Piso 6

Teléfono: 6341500

Fax: 3762211

E-Mail: office.bogota@bakermckenzie.com

IDENTIFICACION

C.C. ☒

NIT

☐C.E. ☐

Otro

☐

Cual

Número 79.143.366 Usaquén TP 20281

④

INVENTOR (ES)

(72)

Nombre: DILLON, Harrison F.; ELEFANT, Dan; DAY, Anthony G.; FRANKLIN, Scott; WITTENBERG, Jon

Dirección: 225 Gateway Boulevard South San Francisco, California 94080 (US); 225 Gateway Boulevard South San Francisco, California 94080 (US); 225 Gateway Boulevard South San Francisco, California 94080 (US); 225 Gateway Boulevard South San Francisco, California 94080 (US); 225 Gateway Boulevard South San Francisco, California 94080 (US).

Nacionalidad o Domicilio: Estados Unidos de América (Respectivamente)

⑤

PCT (85)

Solicitud Internacional No. PCT/US2010/036238

Fecha 26-05-2010

Publicación Internacional No. WO2010/138620

Fecha 02-12-2010

⑥

Título (54)

FRACCIONAMIENTO DE LA BIOMASA MICROBIANA QUE CONTIENE ACEITE

⑦

Clasificación Internacional (51)

C12N 13/00; B01D 3/00

⑧

Prioridad

SI ☒NO ☐

(33) País de Origen

US

(32) Fecha

26-05-2009

(31) Número de Solicitud

61/181,252

9 Comprobante de pago No.

11-111720 - 11-111647 - 11-0064238 - 11-
0079905 - 11-0098063 - 11-111594 - 11-
111626 - 11-104001 - 11-0098012 - 11-
111721 -Fecha 16-11-11 - 01-07-11 - 18-08-11 - 05-10-
11 - 24-10-2011

10

ANEXOS

- ☒ Comprobante de pago de la tasa de presentación de la solicitud
- ☒ Documento que acredite la existencia y representación legal cuando el solicitante sea persona jurídica
- ☒ Poderes, si fuere el caso
- ☐ Documento de cesión del inventor al solicitante o a su causante
- ☐ Declaraciones
- ☒ Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos)
- ☒ Traducción de la solicitud internacional al castellano. (Resumen, descripción, reivindicaciones, dibujos)
- ☐ Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen preliminar (IPEA)
- ☐ Traducción del examen preliminar internacional efectuado por la IPEA
- ☐ De ser el caso, copia del contrato de acceso
- ☐ De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas
- ☐ De ser el caso, certificado de depósito del material biológico
- ☒ Arte final 12 x 12
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud
- ☐ De ser el caso, modificaciones efectuadas a la solicitud internacional.
- ☒ Otros: Extractos , Listado de Secuencias y (CD), Modificación efectuada en Fase Nacional Art. 34.

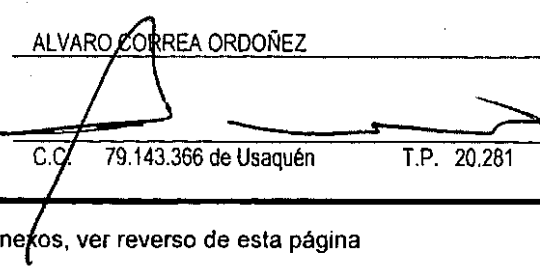
11

FIGURA CARATERISTICA

12

Solicito la concesión de la patente,

NOMBRE: ALVARO CORREA ORDOÑEZ

FIRMA: 
C.C. 79.143.366 de Usaquén T.P. 20.281

Instrucciones para la presentación de los anexos, ver reverso de esta página

2011/018

3

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 111720

Bogotá D.C., Noviembre 16 de 2011 - 17:33:43

RECIBIDO DE : BAKER & MCKENZIE

NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	Vr. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	325105933	16/11/2011	113.695.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-01	SOLICITUDES	114 PTC CAP-II EXAMEN DE PATENTIBILIDAD DE INVENSIÓN	485.000.00
				485.000.00
				\$485.000.00

SON: **CUATROCIENTOS OCHENTA Y CINCO MIL PESOS MONEDA CORRIENTE**

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-177590-00000-0000

Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 563

Act. 411 PRESENTACION Folios: 562

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.slc.gov.co e-mail: info@slc.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Noviembre 16 de 2011 - 17:34:03

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 111647

Bogotá D.C., Noviembre 16 de 2011 - 17:33:43

RECIBIDO DE : BAKER & MCKENZIE NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	Vr. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	325105933	16/11/2011	113.695.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-08	PRORROGA DE TERMINOS	86.000.00	86.000.00
		1608 N.CREACION PRORROGA DE TERMINOS O PLAZO ADICIONAL		\$86.000.00
			=====	=====

SON: **OCHENTA Y SEIS MIL PESOS MONEDA CORRIENTE**

Responsable: _____
Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-177590- -00000-0000

Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 563

ACT. 411 PRESENTACION Folios: 563

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C. - Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Noviembre 16 de 2011 - 17:34:00

2011/0115

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -



RECIBO DE CAJA

No. 11 - 0064238

Bogotá D.C., Julio 01 de 2011 - 14:27:42

RECIBIDO DE : BAKER & MCKENZIE NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	Nº. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	325065742	01/07/2011	95.266.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. LINDITARIO	Vr. CONCEPTO
1	50005-01-09	OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99	OTROS SERVICIOS ADMINISTRATIVOS
			4.000.00	4.000.00
				\$4.000.00

SON: **CUATRO MIL PESOS MONEDA CORRIENTE***

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 11-177590-00000-0000
Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 563
Act. 411 PRESENTACION Folios: 562

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 0079905

Bogotá D.C., Agosto 18 de 2011 - 16:59:22

RECIBIDO DE : BAKER & MCKENZIE

NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	411692035	17/08/2011	82.279.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	5.000.00	5.000.00
				\$5.000.00

SON: **CINCO MIL PESOS MONEDA CORRIENTE***

Responsable

Recibo de Caja Aplicado al Expediente No.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-177590- -00000-0000

Fecha: 2011-12-23 12:00:23

Tra. 11 PATENTE IYII

Act. 411 PRESENTACION

Dep. 2020 DIR. NUEVAS CR

Eve: 379 FASE NACIONAL II

Folios: 563

Tra. 11 PATENTE IYII

Act. 411 PRESENTACION

Folios: 562

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C. - Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Agosto 18 de 2011 - 16:59:47

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -



RECIBO DE CAJA

No. 11 - 0098063

Bogotá D.C., Octubre 05 de 2011 - 17:08:26

RECIBIDO DE : BAKER & MCKENZIE

NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	V.R. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	412134424	04/10/2011	79.821.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1 50005-01-09	OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	10.000.00	10.000.00
				\$10.000.00

SON: **DIEZ MIL PESOS MONEDA CORRIENTE***

Responsable

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-177690-00000-0000

Fecha: 2011-12-23 12:00:23

Tra. 11 PATENTEIYII

Act. 411 PRESENTACION

Dep. 2020 DIR.NUEVASCR

Eve: 379 FASENACIONALII

Folios: 563

Act. 411 PRESENTACION

Folios: 562

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-810165 Call Center: (571) 6513240

Octubre 5 de 2011 - 17:08:47

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 111594

Bogotá D.C., Noviembre 16 de 2011 - 17:33:43

RECIBIDO DE : BAKER & MCKENZIE NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	325105933	16/11/2011	113.895.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNIDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	10.000.00	10.000.00
				\$10.000.00

SON: **DIEZ MIL PESOS MONEDA CORRIENTE***

Responsable: _____
Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 11-177590- -00000-0000
Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 563
Act. 411 PRESENTACION Folios: 562

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -

Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 111626

Bogotá D.C., Noviembre 16 de 2011 - 17:33:43

RECIBIDO DE : BAKER & MCKENZIE

NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	Vr. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	325105933	16/11/2011	113.695.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	54 INFORMACION DETALLADA DE UN TRAMITE	27.000.00	27.000.00
				\$27.000.00

SON: **VEINTISIETE MIL PESOS MONEDA CORRIENTE***

Responsable

Recibo de Caja Aplicado al Expediente No.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 11-177590-00000-0000
Fecha: 2011-12-23 12:00:23
Tra. 11 PATENTE IYII
Act. 411 PRESENTACION
Dep. 2020 DIR. NUEVAS CR
Eve: 379 FASENACIONAL II
Folios: 563
Folios: 562

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C. - Colombia

Web: www.sic.gov.co e-mail: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Noviembre 16 de 2011 - 17:34:00

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

-/-



Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 104001

Bogotá D.C., Octubre 24 de 2011 - 16:26:54

RECIBIDO DE : BAKER & MCKENZIE

NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	Vr. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	410784006	24/10/2011	109.520.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-01 SOLICITUDES	83 INVOCACION DE UNA PRIORIDAD EN SIGNOS DISTINTIVOS	160.000.00	160.000.00
				\$160.000.00

SON: **CIENTO SESENTA MIL PESOS MONEDA CORRIENTE***

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-177590-00000-0000

Fecha: 2011-12-23 12:00:23

Dep. 2020 DIR.NUEVASCR

Tra. 11 PATENTEIYII

Eve: 379 FASENACIONALII

Act. 411 PRESENTACION

Folios: 563

Act. 411 PRESENTACION

Folios: 562

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C.- Colombia

Web: www.sic.gov.co e-mai: info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

, Octubre 24 de 2011 - 16:27:10

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -


Industria y Comercio
SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 0098012

Bogotá D.C., Octubre 05 de 2011 - 17:08:26

RECIBIDO DE : BAKER & MCKENZIE

NI 900.034.906

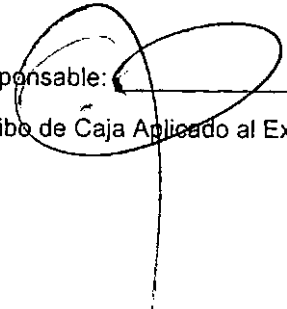
*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	V.R. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	412134424	04/10/2011	79.821.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-09 OTROS SERVICIOS DE PROPIEDAD INDUSTRIAL	99 OTROS SERVICIOS ADMINISTRATIVOS	1.000.00	1.000.00
				\$1.000.00

SON: **MIL PESOS MONEDA CORRIENTE***

Responsable: 

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 11-177690-00000-0000
Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR. NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 563
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 562

Sede Centro: Carrera 13 No. 27 - 00 Pisos 3,4,5 y 10 Bogotá, D.C. - Colombia

Web: www.sic.gov.co e-mail: Info@sic.gov.co Conmutador: (571) 5870000 Fax: (571) 5870284 Línea: 018000-910165 Call Center: (571) 6513240

Octubre 5 de 2011 - 17:08:45

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -



RECIBO DE CAJA

No. 11 - 111721

Bogotá D.C., Noviembre 16 de 2011 - 17:33:43

RECIBIDO DE : BAKER & MCKENZIE NI 900.034.906

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	Vr. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	325105933	16/11/2011	113.695.000.00

*** Conceptos Pagados ***

CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-01 SOLICITUDES	114 PTC CAP-III EXAMEN DE PATENTIBILIDAD DE INVENSIÓN	485.000.00	485.000.00
				\$485.000.00

SON: **CUATROCIENTOS OCHENTA Y CINCO MIL PESOS MONEDA CORRIENTE**

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
No. 11-177590-00000-0000
Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTE I y II Eve: 379 FASE NACIONAL II
Act. 411 PRESENTACION Folios: 563
Act. 411 PRESENTACION Folios: 564

 Industria y Comercio		SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO DIVISION DE NUEVAS CREACIONES TARJETA ARCHIVO PROPIETARIOS	
SUPERINTENDENCIA		PATENTE DE INVENCION ☐ PCT ☒ MODELO DE UTILIDAD ☐ DISEÑO INDUSTRIAL ☐	
(21) N° de solicitud:		(22) Fecha de solicitud:	
(51) Clasificación Internacional: C12N 13/00; B01D 3/00			
(71) Solicitante(s): SOLAZYME, INC.			
(72) Inventor(es): DILLON, Harrison F.; ELEFANT, Dan; DAY, Anthony G.; FRANKLIN, Scott; WITTENBERG, Jon			
(30) Prioridad	(31) No. Prioridad	32) Fecha	(33) País
	61/181,252	26-05-2009	US
(85) Fecha inicio fase nacional: 26-12-2011		(87) Publicación internacional	
(86) Datos relativos a la presentación de la solicitud PCT		Fecha: 02-12-2010	
Fecha de presentación de la solicitud: 26-05-2010		N° publicación: WO2010/138620	
No. de solicitud: PCT/US2010/036238			
(54) Título: FRACCIONAMIENTO DE LA BIOMASA MICROBIANA QUE CONTIENE ACEITE			

 Industria y Comercio		SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO DIVISION DE NUEVAS CREACIONES TARJETA ARCHIVO TEMATICO	
<u>SUPERINTENDENCIA</u> <u>PATENTE DE INVENCION</u> <u> </u> <u>PCT</u> <u>X</u> <u>MODELO DE UTILIDAD</u> <u> </u> <u>DISEÑO INDUSTRIAL</u> <u> </u>			
(21) N° de solicitud: (22) Fecha de solicitud:			
(51) Clasificación Internacional: C12N 13/00; B01D 3/00			
(71) Solicitante(s) SOLAZYME, INC.			
(72) Inventor(es): DILLON, Harrison F.; ELEFANT, Dan; DAY, Anthony G.; FRANKLIN, Scott; WITTENBERG, Jon			
(74) Apoderado: ALVARO CORREA ORDOÑEZ			
(30) Prioridad	(31) No. Prioridad	32) Fecha	(33) Pais
	61/181,252	26-05-2009	US
(85) Fecha límite inicio fase nacional: 26-12-2011		(87) Publicación internacional:	
(86) Datos relativos a la presentación de la solíc. PCT		Fecha: 02-12-2010	
Fecha de presentación de la solicitud: 26-05-2010		No. Publicación: WO2010/138620	
No. de solicitud: PCT/US2010/036238			
(54) Título: FRACCIONAMIENTO DE LA BIOMASA MICROBIANA QUE CONTIENE ACEITE			
(57) Resumen: La invención se refiere generalmente a la producción de composiciones de hidrocarburo, tales como un lípido, en microorganismos. En particular, la invención proporciona métodos para la extracción, recuperación, aislamiento y obtención de un lípido a partir de un microorganismo y composiciones que comprenden el lípido. La invención describe, además, métodos para producir composiciones de hidrocarburo para usar como biodiesel, diesel renovable, combustible de reactor, y otros materiales.			
CONTINUA AL RESPALDO ...			

 Industria y Comercio SUPERINTENDENCIA		SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO DIVISION DE NUEVAS CREACIONES EXTRACTO PARA PUBLICACIÓN SOLICITUD DE PATENTE DE INVENCION PCT	
(21) N° de solicitud:		(51) Int Cl: C12N 13/00; B01D 3/00	
22) Fecha de solicitud:		(71) Solicitante(s) SOLAZYME, INC.	
(30) Prioridad	(31) No. Prioridad	(32) Fecha	(33) País
	61/181,252	26-05-2009	US
(85)	Fecha limite inicio fase nacional: 26-12-2011		
(86)	Datos relativos a la presentación de la solicitud PCT		
	Fecha de presentación de la solicitud: 26-05-2010		
	No. de PCT/US2010/036238		
		(72) Inventor(es) DILLON, Harrison F.; ELEFANT, Dan; DAY, Anthony G.; FRANKLIN, Scott; WITTENBERG, Jon	
		(74) Apoderado: ALVARO CORREA ORDOÑEZ	
		(87) Publicación internacional	
		Fecha: 02-12-2010	
		N° publicación: WO2010/138620	
(54) Título: FRACCIONAMIENTO DE LA BIOMASA MICROBIANA QUE CONTIENE ACEITE			

Resumen: La invención se refiere generalmente a la producción de composiciones de hidrocarburo, tales como un lípido, en microorganismos. En particular, la invención proporciona métodos para la extracción, recuperación, aislamiento y obtención de un lípido a partir de un microorganismo y composiciones que comprenden el lípido. La invención describe, además, métodos para producir composiciones de hidrocarburo para usar como biodiesel, diesel renovable, combustible de reactor, y otros materiales.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
2 December 2010 (02.12.2010)(10) International Publication Number
WO 2010/138620 A1(51) International Patent Classification:
C12N 13/00 (2006.01) B01D 3/00 (2006.01)(21) International Application Number:
PCT/US2010/036238(22) International Filing Date:
26 May 2010 (26.05.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/181,252 26 May 2009 (26.05.2009) US

(71) Applicant (for all designated States except US): SO-LAZYME, INC. [US/US]; 225 Gateway Boulevard, South San Francisco, California 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DILLON, Harrison F. [US/US]; 225 Gateway Boulevard, South San Francisco, California 94080 (US). ELEFANT, Dan [US/US]; 225 Gateway Boulevard, South San Francisco, California 94080 (US). DAY, Anthony G. [US/US]; 225 Gateway Boulevard, South San Francisco, California 94080 (US). FRANKLIN, Scott [US/US]; 225 Gateway Boulevard, South San Francisco, California 94080 (US). WITTENBERG, Jon [US/US]; 225 Gateway Boulevard, South San Francisco, California 94080 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: FRACTIONATION OF OIL-BEARING MICROBIAL BIOMASS

(57) Abstract: The invention generally relates to the production of hydrocarbon compositions, such as a lipid, in microorganisms. In particular, the invention provides methods for extracting, recovering, isolating and obtaining a lipid from a microorganism and compositions comprising the lipid. The invention also discloses methods for producing hydrocarbon compositions for use as biodiesel, renewable diesel, jet fuel, and other materials.



WO 2010/138620 A1

FRACTIONATION OF OIL-BEARING MICROBIAL BIOMASS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit under 35 U.S.C. 119(e) of US Provisional Application No. 61/181,252, filed May 26, 2009, which is incorporated herein by reference in its entirety for all purposes.

REFERENCE TO A SEQUENCE LISTING

[0002] This application includes a Sequence Listing, appended hereto as pages 1-45.

FIELD OF THE INVENTION

[0003] The invention generally relates to the production of oil compositions, such as a lipid, in microorganisms. In particular, the invention provides methods for extracting, recovering, isolating and obtaining a lipid from a microorganism and compositions comprising the lipid. The invention also discloses methods for producing hydrocarbon or lipid compositions for production of biodiesel, renewable diesel, jet fuel, and lipid surfactants having various carbon chain lengths, including C8, C10, C12 and C14.

BACKGROUND OF THE INVENTION

[0004] Fossil fuel is a general term for buried combustible geologic deposits of organic materials, formed from decayed plants and animals that have been converted to crude oil, coal, natural gas, or heavy oils by exposure to heat and pressure in the earth's crust over hundreds of millions of years.

[0005] Fossil fuels are a finite, non-renewable resource. With global modernization in the 20th and 21st centuries, the thirst for energy from fossil fuels, especially gasoline derived from oil, is one of the causes of major regional and global conflicts. Increased demand for energy by the global economy has also placed increasing pressure on the cost of hydrocarbons. Aside from energy, many industries, including plastics and chemical manufacturers, rely heavily on the availability of hydrocarbons as a feedstock for their manufacturing processes. Alternatives to current sources of supply could help mitigate the upward pressure on these raw material costs.

[0006] Lipids for use in biofuels can be produced in microorganisms, such as algae, fungi, and bacteria. Typically, manufacturing a lipid in a microorganism involves growing microorganisms, such as algae, fungi, or bacteria, which are capable of producing a desired lipid in a fermentor or bioreactor, isolating the microbial biomass, drying it, and extracting the intracellular lipids.

[0007] There is a need for a process for extracting lipids from microorganism which solves the above problems of low efficiency and high cost of lipid extraction from microorganism. The present invention provides a solution to these prior art problems.

BRIEF SUMMARY OF THE INVENTION

[0008] In one aspect, the present invention is directed to a method of extracting a lipid from a microorganism. In one embodiment, the method comprises lysing a cultured microorganism that has not been subjected to a drying step between culturing and lysing, and which contains a lipid, to produce a lysate, treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent, and separating the lysate into layers comprising a lipid:organic solvent layer and an aqueous layer, whereby the lipid is extracted from the microorganism.

[0009] In another aspect, the present invention is directed to a composition comprising a lipid isolated from a microorganism, and an oil obtained from a source other than the microorganism. In some cases, the ratio of the lipid to the oil is between 1 and 100. In other cases, the ratio of the lipid to the oil is between 1 and 10. In some embodiments, the lipid comprises at least 50% of a C18:1 lipid, or at least 10% of a C12 and C14 lipid combined.

[0010] In another aspect, the present invention is directed to a method of producing biodiesel. In one embodiment, the method comprises lysing a lipid-containing microorganism to produce a lysate, treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent, separating the lysate into a lipid:organic solvent composition and an aqueous composition and, optionally, an emulsified composition and/or cell pellet composition, removing the lipid:organic solvent composition from the aqueous composition, emulsion composition, or cell pellet composition, and transesterifying the lipid:organic solvent composition to produce biodiesel. In some cases, the biodiesel meets or exceeds the ASTM D6751 biodiesel standard and/or the EN 14214 biodiesel standard. In one embodiment, the microorganism is not subjected to a drying step.

[0011] In some embodiments of the biodiesel production method at least 20% w/v of total lipid from the microorganism is C18:1. In one embodiment, the layers further comprise a lipid:aqueous emulsion layer and/or a cell pellet. Transesterification of the lipid:organic solvent can be performed without substantially separating the lipid from the organic solvent.

[0012] In still another aspect, the present invention is directed to a method of producing renewable diesel. In one embodiment, the method comprises lysing a lipid-containing microorganism to produce a lysate, treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent, separating the lysate into a lipid:organic solvent composition and an aqueous composition and, optionally, an emulsified composition and/or cell pellet composition, removing the lipid:organic solvent composition from the aqueous composition, emulsion composition, or cell pellet composition, and treating the lipid:organic solvent composition to produce a straight chain alkane renewable diesel product. In some cases, the renewable diesel meets or exceeds the ASTM D 975 standard.

[0013] In various embodiments of the renewable diesel production method, the microorganism has not been subjected to a drying step. In some cases, treating the lipid:organic solvent composition is performed without substantially separating the lipid from the organic solvent. In some embodiments, treating the lipid:organic solvent composition comprises hydrotreating, hydroprocessing, or indirect liquefaction.

[0014] In another aspect, the present invention is directed to a method of producing a jet fuel. In one embodiment, the method comprises lysing a lipid-containing microorganism to produce a lysate, treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent, separating the lysate into a lipid:organic solvent composition and an aqueous composition and, optionally, an emulsified composition or cell pellet composition, removing the lipid:organic solvent composition from the aqueous composition, emulsion composition, or cell pellet composition, treating the lipid:organic solvent composition to produce a straight chain alkane, and cracking the straight chain alkane to produce the jet fuel product.

[0015] In various embodiments of the jet fuel production method, the microorganism has not been subjected to a drying step. In some cases, treating the lipid:organic solvent composition is performed without substantially separating the lipid from the organic solvent. In some embodiments, treating the lipid:organic solvent composition is performed by flowing

the lipid:organic solvent composition to a fluid catalytic cracking zone, and can further comprise contacting the lipid:organic solvent composition with a catalyst at cracking conditions. In one embodiment, treating the lipid:organic solvent composition is performed by hydrodeoxygenating the lipid:organic solvent composition. In some cases, the method further comprises subjecting the hydrodeoxygenated lipid:organic solvent composition to isomerization.

[0016] In another aspect, the present invention is directed to a method of extracting lipid from a microorganism by contacting a microorganism containing a lipid with an acid to produce a lysate, separating the lysate into layers comprising an aqueous layer and a lipid:aqueous emulsion layer, and extracting lipid from the emulsion layer.

[0017] In various embodiments of the method of extracting lipid from a microorganism, the microorganism is contacted with the acid to produce an acid concentration of 5-200 mN, and contacting the microorganism can be performed above 25°C. In various embodiments, an organic solvent is added to the microorganism or lysate before, simultaneously with, or after contacting the microorganism with the acid. In some cases, the microorganism is contacted with an acid at a pH of no more than 4, at a pH of no more than 3, or at a pH of no more than 2. In some embodiments, in addition to contacting the microorganism with the acid, one or more additional methods of lysing the microorganism is also utilized. In some cases, contacting the microorganism with the acid is performed at a temperature of 50-160 °C. In other cases, the step of contacting the microorganism with the acid is performed at a temperature of 20-65°C. In one embodiment, the lipid is extracted from the emulsion by contacting the emulsion with an organic solvent, whereby the lipid partitions from the emulsion into the organic solvent. In some cases, separation of the lysate is performed by cooling the emulsion below 25°C, below 10°C, or to a temperature at or below 0°C, whereby a lipid layer separates from the emulsion layer. In one embodiment, the method further comprises centrifuging the emulsion after the cooling to separate the lipid layer. In some embodiments, the method further comprises separating the emulsion from the aqueous layer before separating the lipid from the emulsion. In one embodiment, the lipid is extracted from the lipid:aqueous emulsion without use of an organic solvent.

[0018] In another aspect, the present invention is directed to a method of extracting lipid from a microorganism by lysing a microorganism containing a lipid to produce a lysate

comprising a lipid:aqueous emulsion, and cooling the emulsion below 25°C to separate the lipid from the emulsion.

[0019] In another aspect, the present invention is directed to a method of extracting lipid from microbial biomass generated by culturing a microorganism that produces a lipid. Extraction comprises lysing the microorganisms in the biomass to produce a lysate comprising a lipid:aqueous emulsion. In some cases, separation of a lipid layer from the emulsion comprises destabilizing the emulsion with the addition of a surfactant to the emulsion. In some cases, the lysate can be treated with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent, and the lysate can be separated into two or more layers, including a lipid:organic solvent layer and at least one aqueous layer. In some cases, the method can further include removing the lipid:organic solvent composition from the other layer(s). Optionally, removing the lipid:organic solvent composition from the other layer(s) can be performed without substantially separating the lipid from the organic solvent. In some cases, the method can further include transesterifying the lipid:organic solvent composition to produce a fatty acid alkyl ester, or treating the lipid:organic solvent composition to produce a straight chain alkane. In the latter case, the method can further include cracking the straight chain alkane.

[0020] In another aspect, the present invention is directed to a composition of fatty acid esters comprising esters derived from a microorganism, and esters derived from a nonmicrobial oil. In some embodiments, no more than 20%, no more than 10%, or no more than 6%, of the esters in the composition are derived from a nonmicrobial oil. In some cases, the esters are selected from the group consisting of methyl, ethyl and alkyl esters. In some embodiments, the cultured microorganism has not been separated from liquid medium used to culture the microorganism, and/or is present at a ratio of less than 1:1 v/v of cultured microorganism to extracellular liquid media.

[0021] In various embodiments of the method of extracting lipid from a microorganism, the emulsion is cooled below 10°C, or to a temperature at or below 0°C. In one embodiment, the method further comprises centrifuging the lysate to produce layers comprising the emulsion and an aqueous layer. In some cases, the method further comprises separating the emulsion from the aqueous layer. In one embodiment, the lipid is extracted from the lipid:aqueous emulsion without use of an organic solvent. In some cases, the method further comprises

treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent.

[0022] Microorganisms useful in accordance with the present invention can be selected from the group of microorganisms consisting of a bacterium, a cyanobacterium, a eukaryotic microalgae, an oleaginous yeast, and a fungus. In some cases, the microorganisms can be selected from Tables 1, 2 or 3. Such microorganisms include microorganisms of the genus *Chlorella*. In one embodiment, the microorganism is *Chlorella protothecoides*. In some cases, the microorganisms of the present invention produce a lipid that comprises at least 10% w/v of total cellular lipid as C18 triacylglycerols, or at least 10% w/v of total cellular lipid as C16 triacylglycerols. In some cases, the microorganisms produce a lipid that comprises at least 10% w/v of total cellular lipid as C14 triacylglycerols, at least 10% w/v of total cellular lipid as C12 triacylglycerols, or at least 10% w/v of total cellular lipid as C10 triacylglycerols. In some cases, the microorganisms of the present invention have at least 85% 23S rRNA genomic sequence identity to one or more sequences selected from the group consisting of SEQ ID NOs:7-33.

[0023] In some cases, the microorganism is of the genus *Prototheca*. In one embodiment, the microorganism is *Prototheca moriformis*. In some cases, the microorganisms of the present invention have at least 75% sequence identity to one or more sequences selected from the group consisting of SEQ ID NOs: 33, 34, 35, 16, 36, 17, 37, 38, 39, and 32.

[0024] In various embodiments in accordance with the present invention, lysing the cultured microorganisms is performed at an acidic pH, at a pH of no more than 5, at a pH of no more than 4, at a pH of no more than 3, or at a pH of no more than 2. In some embodiments, lysing the cultured microorganisms is performed at a pH of at least 9. In some cases, lysing the cultured microorganism comprises one or more methods of lysing selected from the list consisting of heating, sonication, mechanical lysis, osmotic shock, pressure oscillation, expression of an autolysis gene, exposure to pH above 8, exposure to pH below 6, and digestion with an enzyme. In some cases, lysing the microorganism comprises acidic lysis and heating. In one embodiment, the microorganism is lysed by digestion with a polysaccharide-degrading enzyme, which can be a polysaccharide-degrading enzyme from *Chlorella* or a *Chlorella* virus. In some cases, the microorganism is lysed by digestion with a protease. In other cases, the microorganism is digested with a combination of at least one protease and at least one polysaccharide-degrading enzyme. In some cases, the protease is

alcalase and/or the polysaccharide-degrading enzyme is mannaway. In some cases, other combinations of the foregoing are used, for example, contacting the biomass with a protease and a polysaccharide-degrading enzymes in combination with heating the biomass to a temperature of at least 30 degrees Celsius.

[0025] In various embodiments in accordance with the present invention, treating the lysate comprises treating with more than about 5% v/v of an organic solvent to the lysate. In other cases, the treating step comprises treating with more than about 6% v/v of an organic solvent to the lysate, or with more than about 7% v/v of an organic solvent to the lysate. In still other cases, the v/v of the organic solvent to the lysate is between greater than about 5% and greater than about 25%. In one embodiment, treatment of the lysate can be facilitated by agitating the lysate.

[0026] In embodiments of the present invention the organic solvent can comprise an oil. The oil can be selected from the group consisting of oil from soy, rapeseed, canola, palm, palm kernel, coconut, corn, waste vegetable, Chinese tallow, olive, sunflower, cotton seed, chicken fat, beef tallow, porcine tallow, microalgae, macroalgae, *Cuphea*, flax, peanut, choice white grease, lard, *Camelina sativa*, mustard seed cashew nut, oats, lupine, kenaf, calendula, hemp, coffee, linseed, hazelnut, euphorbia, pumpkin seed, coriander, camellia, sesame, safflower, rice, tung oil tree, cocoa, copra, pium poppy, castor beans, pecan, jojoba, jatropha, macadamia, Brazil nuts, avocado, a fossil oil, or a distillate fraction thereof. In some cases the oil is soy oil. In some cases the oil is palm oil. In other cases the oil is coconut oil. In still other cases the oil is canola oil. In yet other cases the oil is jatropha oil. In one embodiment, the time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent is between 0.1 and 30 minutes.

[0027] In various embodiments in accordance with the present invention, separating the lysate comprises centrifugation of the treated lysate, whereby the lysate is separated into a light layer comprising the lipid:organic solvent composition and a heavy layer comprising the aqueous composition and, optionally, an emulsified composition and/or cell pellet composition. In some cases, separating the lysate comprises settling of the treated lysate, whereby the lysate is separated into a light layer comprising the lipid:organic solvent composition and a heavy layer comprising the aqueous composition and, optionally, an emulsified composition and/or cell pellet composition. Separation of the lysate can include reducing the temperature of the mixture below 25°C, below 10°C, or to a temperature at or

below 0°C. In some aspects of the invention, the layers further comprise a lipid:aqueous emulsion layer and/or a cell pellet.

[0028] In some cases, methods of the invention further comprise removing the lipid:organic solvent composition from the other layer(s). In one embodiment, removal is performed without substantially separating the lipid from the organic solvent.

[0029] In some methods of the invention, the microorganism is produced in a culturing process and then optionally stored for a period of time between termination of the culturing process and undertaking additional steps to lyse the microorganism. In some cases, lysing the microorganisms in the biomass produces a lysate comprising a lipid:aqueous emulsion. In some cases, the microorganism is stored for at least one hour between termination of the culturing process and undertaking additional steps to lyse the cultured microorganism. In some cases, the microorganism is stored for at least twenty-four hours between termination of the culturing process and undertaking additional steps to lyse the cultured microorganism. In some cases, the microorganism is stored for at least forty-eight hours between termination of the culturing process and undertaking additional steps to lyse the cultured microorganism. In other cases, the fermentation broth is stored. In some cases, the fermentation broth is concentrated, for example, by centrifugation or filtration, and the cells are resuspended in an aqueous media before storage. In some cases the aqueous media is deionized or distilled water.

[0030] In some embodiments of the present invention, microorganisms (biomass) prepared in a culture process are optionally stored at a temperature below 15 degrees Celsius between termination of the culturing process and undertaking additional steps to lyse the cultured microorganism. In some cases, the biomass is stored at a temperature below 5 degrees Celsius between termination of the culturing process and undertaking additional steps to lyse the cultured microorganisms. In some cases, the microorganism is stored at a temperature above 30 degrees Celsius between termination of the culturing process and undertaking additional steps to lyse the cultured microorganism. In some cases, the microorganism is stored at a temperature above 40 degrees Celsius between termination of the culturing process and undertaking additional steps to lyse the cultured microorganism. In some cases, the microorganism is subjected to agitation during storage.

BRIEF DESCRIPTION OF THE DRAWINGS

[0031] Figure 1 depicts two lipid profiles for pure oils prepared according to the method of the present invention. Fatty acids C10:1, C10, C12, C18:3, C14, C18:2, C16:0, C18:1, and C18:0 of pure *Chlorella* oil (light-shaded bars) and coconut oil (dark-shaded bars) are shown. The majority of pure *Chlorella* oil comprises C18:1 fatty acids (59%), while the majority of pure coconut oil comprises C12 fatty acids (54%). Details are described in Example 6.

[0032] Figure 2 depicts the ratio of C18:1 to C12 fatty acids and C18:1 to C14 fatty acids in pure algal/coconut oil mixture. Details are described in Example 6.

[0033] Figure 3 depicts results of chemical/heat treatment of *Chlorella protothecoides*. Details are described in Example 7.

[0034] Figure 4 depicts an image of oil layer recovered from a frozen, heat treated emulsion. Details are described in Example 7 (Enzyme Treatment #1).

[0035] Figure 5 depicts a result of a thin layer chromatography (TLC) analysis of control, basic, and acid generated oil samples. Lane 1, Reference @ 100 µg/each (FAME/TAG/FFA); Lane 2, No chemical treatment; Lane 3, treatment with 120 mN H₂SO₄; Lane 4, treatment with 160 mN KOH; Lane 5, no chemical treatment; Lane 6, treatment with 120 mN H₂SO₄; Lane 7, treatment with 160mN KOH. FAME, fatty acid methyl ester; TAG, triacylglycerides; FFA, free fatty acid; DAG, diacylglycerol (1,3 – diolein); MAG, monoacylglycerol (1-monoolein) Details are described in Example 7.

[0036] Figure 6 depicts an image of layers recovered from enzyme treated (left tube) versus untreated (right tube) cell culture material. Details are described in Example 7 (Enzyme Treatment #2).

[0037] Figure 7 shows maps of the cassettes used in *Prototheca* transformations, as described in Example 9.

[0038] Figure 8 shows the results of Southern blot analysis on three transformants of UTEX strain 1435, as described in Example 9.

[0039] Figure 9 shows a schematic of the codon optimized and non-codon optimized *suc2* (yeast sucrose invertase (yInv)) transgene construct. The relevant restriction cloning sites are indicated and arrows indicate the direction of transcription.

DETAILED DESCRIPTION OF THE INVENTION

I. DEFINITIONS

[0040] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton *et al.*, *Dictionary of Microbiology and Molecular Biology* (2nd ed. 1994); *The Cambridge Dictionary of Science and Technology* (Walker ed., 1988); *The Glossary of Genetics*, 5th Ed., R. Rieger *et al.* (eds.), Springer Verlag (1991); and Hale & Marham, *The Harper Collins Dictionary of Biology* (1991). As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

[0041] As used with reference to a nucleic acid, the phrase "active in microalgae" refers to a nucleic acid that is functional in microalgae. For example, a promoter that has been used to drive an antibiotic resistance gene to impart antibiotic resistance to a transgenic microalgae is active in microalgae. Examples of promoters active in microalgae are promoters endogenous to certain algae species and promoters found in plant viruses.

[0042] As used herein, an "acyl carrier protein" or "ACP" is a protein which binds a growing acyl chain during fatty acid synthesis as a thiol ester at the distal thiol of the 4'-phosphopantetheine moiety and comprises a component of the fatty acid synthase complex.

[0043] An "acyl-CoA molecule" or "acyl-CoA" is a molecule comprising an acyl moiety covalently attached to coenzyme A through a thiol ester linkage at the distal thiol of the 4'-phosphopantetheine moiety of coenzyme A.

[0044] As used herein, the term "alkyl" refers to a straight or branched chain hydrocarbon radical, and can include di- and multivalent radicals, having the number of carbon atoms designated (i.e. C1-C10 means one to ten carbons). Examples of saturated hydrocarbon radicals include groups such as methyl, ethyl, n-propyl, isopropyl, n-butyl, t-butyl, isobutyl, sec-butyl, homologs and isomers of, for example, n-pentyl, n-hexyl, n-heptyl, n-octyl, and the like.

[0045] As used herein, the term "alkenyl" refers to an unsaturated alkyl group one having one or more double bonds. Examples of alkenyl groups include vinyl, 2-propenyl, crotyl, 2-isopentenyl, 2-(butadienyl), 2,4-pentadienyl and 3-(1,4-pentadienyl), and the higher homologs and isomers.

[0046] As used herein, the term “alkynyl” refers to an unsaturated alkyl group one having one or more triple bonds. Examples of alkynyl groups include ethynyl (acetylenyl), 1-propynyl, 1- and 2-butynyl, and the higher homologs and isomers.

[0047] As used herein, the phrase “aqueous or emulsified composition” refers to microbial biomass that contains lipid.

[0048] As used herein, the phrase “area percent” refers to the area of peaks observed using FAME GC/FID detection methods in which every fatty acid in the sample is converted into a fatty acid methyl ester (FAME) prior to detection. For examples, a separate peak is observed for a fatty acid of 14 carbon atoms with no unsaturation (C14:0) compared to any other fatty acid such as a C14:1. The peak area for each class of FAME is directly proportional to its percent composition in the mixture and is calculated based on the sum of all peaks present in the sample (*i.e.*, [area under specific peak/total area of all measured peaks] x 100). When referring to lipid profiles of oils and cells of the invention, “at least 4% C8-C14,” for example, means that at least 4% of the total fatty acids in the cell or in the extracted glycerolipid composition have a chain length that includes 8, 10, 12, or 14 carbon atoms.

[0049] As used herein, the term “aryl” refers to a polyunsaturated, aromatic, hydrocarbon substituent having 5-12 ring members, which can be a single ring or multiple rings (up to three rings) which are fused together or linked covalently. Non-limiting examples of aryl groups include phenyl, 1-naphthyl, 2-naphthyl, 4-biphenyl, and benzyl. Other aryl groups are also useful in the present invention.

[0050] As used herein, the term “axenic” refers to a culture of an organism that is free from contamination by other living organisms.

[0051] As used herein, the term “base” refers to any compound whose pKa is greater than that of water.

[0052] As used herein, the term “biodiesel” refers to a fatty acid ester produced from transesterification of a lipid.

[0053] As used herein, the term “biomass” refers to material produced by growth and/or propagation of cells. Biomass may contain cells and/or intracellular contents as well as extracellular material. Extracellular material includes, but is not limited to, compounds secreted by a cell.

[0054] As used herein, the term "bioreactor" refers to an enclosure or partial enclosure in which cells, e.g., microorganisms, are cultured, optionally in suspension.

[0055] As used herein, the term "catalyst" refers to an agent, such as a molecule or macromolecular complex, capable of facilitating or promoting a chemical reaction of a reactant to a product without becoming a part of the product. A catalyst thus increases the rate of a reaction, after which, the catalyst may act on another reactant to form the product. A catalyst generally lowers the overall activation energy required for the reaction such that it proceeds more quickly or at a lower temperature. Thus a reaction equilibrium may be more quickly attained. Examples of catalysts include enzymes, which are biological catalysts, and heat, which is a non-biological catalyst.

[0056] As used herein, the term "cellulosic material" means the products of digestion of cellulose, such as glucose, xylose, arabinose, disaccharides, oligosaccharides, lignin, furfurals, and other molecules.

[0057] As used herein, the term "co-culture" and variants thereof, such as "co-cultivate," refer to the presence of two or more types of cells in the same bioreactor. The two or more types of cells may both be microorganisms, such as microalgae, or may be a microalgal cell cultured with a different cell type. The culture conditions may be those that foster growth and/or propagation of the two or more cell types or those that facilitate growth and/or proliferation of one, or a subset, of the two or more cells while maintaining cellular growth for the remainder.

[0058] As used herein, the term "cofactor" refers to any molecule, other than the substrate, that is required for an enzyme to carry out its enzymatic activity.

[0059] A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions.

[0060] As used herein, the term "cultivated" and variants thereof, refer to the intentional fostering of growth (increases in cell size, cellular contents, and/or cellular activity) and/or propagation (increases in cell numbers via mitosis) of one or more cells by use of intended culture conditions. The combination of both growth and propagation may be termed proliferation. The one or more cells may be those of a microorganism, such as microalgae. Examples of intended conditions include the use of a defined medium (with known characteristics such as pH, ionic strength, and carbon source), specified temperature, oxygen

tension, carbon dioxide levels, and growth in a bioreactor. The term does not refer to the growth or propagation of microorganisms in nature or otherwise without direct human intervention, such as natural growth of an organism that ultimately becomes fossilized to produce geological crude oil.

[0061] As used herein, the term "cycloalkyl" refers to a saturated cyclic hydrocarbon having 3 to 15 carbons, and 1 to 3 rings that can be fused or linked covalently. Cycloalkyl groups useful in the present invention include, but are not limited to, cyclopentyl, cyclohexyl, cycloheptyl and cyclooctyl. Bicycloalkyl groups useful in the present invention include, but are not limited to, [3.3.0]bicyclooctanyl, [2.2.2]bicyclooctanyl, [4.3.0]bicyclononane, [4.4.0]bicyclodecane (decalin), spiro[3.4]octanyl, spiro[2.5]octanyl, and so forth.

[0062] As used herein, the term "cycloalkenyl" refers to an unsaturated cyclic hydrocarbon having 3 to 15 carbons, and 1 to 3 rings that can be fused or linked covalently. Cycloalkenyl groups useful in the present invention include, but are not limited to, cyclopentenyl, cyclohexenyl, cycloheptenyl and cyclooctenyl. Bicycloalkenyl groups are also useful in the present invention.

[0063] As used herein, the term "cytolysis" refers to the lysis of cells in a hypotonic environment. Cytolysis is caused by excessive osmosis, or movement of water, towards the inside of a cell (hyperhydration). The cell membrane cannot withstand the osmotic pressure of the water inside, and so it explodes.

[0064] As used herein, the term "exogenous gene" refers to a nucleic acid transformed into a cell. A transformed cell may be referred to as a recombinant cell, into which additional exogenous gene(s) may be introduced. The exogenous gene may be from a different species (and so heterologous), or from the same species (and so homologous) relative to the cell being transformed. In the case of a homologous gene, it occupies a different location in the genome of the cell relative to the endogenous copy of the gene. The exogenous gene may be present in more than one copy in the cell. The exogenous gene may be maintained in a cell as an insertion into the genome or as an episomal molecule.

[0065] As used herein, the term "exogenously provided" in the context of culturing a cell, refers to a molecule provided to a culture media of a cell culture.

[0066] As used herein, the terms "expression vector" or "expression construct" refer to a nucleic acid construct, generated recombinantly or synthetically, with a series of specified

nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

[0067] As used herein, a "fatty acyl-ACP thioesterase" is an enzyme that catalyzes the cleavage of a fatty acid from an acyl carrier protein (ACP) during lipid synthesis.

[0068] As used herein, a "fatty acyl-CoA/aldehyde reductase" is an enzyme that catalyzes the reduction of an acyl-CoA molecule to a primary alcohol.

[0069] As used herein, a "fatty acyl-CoA reductase" is an enzyme that catalyzes the reduction of an acyl-CoA molecule to an aldehyde.

[0070] As used herein, a "fatty aldehyde decarbonylase" is an enzyme that catalyzes the conversion of a fatty aldehyde to an alkane.

[0071] As used herein, a "fatty aldehyde reductase" is an enzyme that catalyzes the reduction of an aldehyde to a primary alcohol.

[0072] As used herein, the term "fixed carbon source" means molecule(s) containing carbon, preferably organic, that are present at ambient temperature and pressure in solid or liquid form.

[0073] As used herein, the term "fungus," means heterotrophic organisms characterized by a chitinous cell wall from the kingdom of fungi.

[0074] As used herein, the term "heteroaryl" refers to a polyunsaturated, aromatic, hydrocarbon substituent having 5-12 ring members, which can be a single ring or multiple rings (up to three rings) which are fused together or linked covalently, and which has at least one heteroatom in the ring, such as N, O, or S. A heteroaryl group can be attached to the remainder of the molecule through a heteroatom. Non-limiting examples of heteroaryl groups include 1-pyrrolyl, 2-pyrrolyl, 3-pyrrolyl, 3-pyrazolyl, 2-imidazolyl, 4-imidazolyl, pyrazinyl, 2-oxazolyl, 4-oxazolyl, 2-phenyl-4-oxazolyl, 5-oxazolyl, 3-isoxazolyl, 4-isoxazolyl, 5-isoxazolyl, 2-thiazolyl, 4-thiazolyl, 5-thiazolyl, 2-furyl, 3-furyl, 2-thienyl, 3-thienyl, 2-pyridyl, 3-pyridyl, 4-pyridyl, 2-pyrimidyl, 4-pyrimidyl, 5-benzothiazolyl, purinyl, 2-benzimidazolyl, 5-indolyl, 1-isoquinolyl, 5-isoquinolyl, 2-quinoxalyl, 5-quinoxalyl, 3-quinolyl, and 6-quinolyl. Additional heteroaryl groups useful in the present invention include

pyridyl N-oxide, tetrazolyl, benzofuranyl, benzothienyl, indazolyl, or any of the radicals substituted, especially mono- or di-substituted.

[0075] As used herein, the term "heteroatom" means any atom that is not carbon or hydrogen. Examples of heteroatoms include magnesium, calcium, potassium, sodium, sulfur, phosphorus, iron and copper.

[0076] As used herein, the term "heterocycloalkyl" refers to a saturated cyclic hydrocarbon having 3 to 15 ring members, and 1 to 3 rings that can be fused or linked covalently, and which has at least one heteroatom in the ring, such as N, O, or S. Additionally, a heteroatom can occupy the position at which the heterocycle is attached to the remainder of the molecule. Examples of heterocycloalkyl include 1 (1,2,5,6-tetrahydropyridyl), 1-piperidinyl, 2-piperidinyl, 3-piperidinyl, 4-morpholinyl, 3-morpholinyl, tetrahydrofuran-2-yl, tetrahydrofuran-3-yl, tetrahydrothien-2-yl, tetrahydrothien-3-yl, 1-piperazinyl, 2-piperazinyl, and the like.

[0077] As used herein, in the context of biomass, the term "homogenate" means biomass that has been physically disrupted.

[0078] As used herein, "hydrocarbon" refers to: (a) a molecule containing only hydrogen and carbon atoms wherein the carbon atoms are covalently linked to form a linear, branched, cyclic, or partially cyclic backbone to which the hydrogen atoms are attached; or (b) a molecule that only primarily contains hydrogen and carbon atoms and that can be converted to contain only hydrogen and carbon atoms by one to four chemical reactions. Nonlimiting examples of the latter include hydrocarbons containing an oxygen atom between one carbon and one hydrogen atom to form an alcohol molecule, as well as aldehydes containing a single oxygen atom. Methods for the reduction of alcohols to hydrocarbons containing only carbon and hydrogen atoms are well known. Another example of a hydrocarbon is an ester, in which an organic group replaces a hydrogen atom (or more than one) in an oxygen acid. The molecular structure of hydrocarbon compounds varies from the simplest, in the form of methane (CH₄), which is a constituent of natural gas, to the very heavy and very complex, such as some molecules such as asphaltenes found in crude oil, petroleum, and bitumens. Hydrocarbons may be in gaseous, liquid, or solid form, or any combination of these forms, and may have one or more double or triple bonds between adjacent carbon atoms in the backbone. Accordingly, the term includes linear, branched, cyclic, or partially cyclic alkanes,

alkenes, lipids, and paraffin. Examples include propane, butane, pentane, hexane, octane, squalene and carotenoids.

[0079] As used herein, the term "hydrocarbon modification enzyme" refers to an enzyme that alters the covalent structure of a hydrocarbon. Examples of hydrocarbon modification enzymes include a lipase, a fatty acyl-ACP thioesterase, a fatty acyl-CoA/aldehyde reductase, a fatty acyl-CoA reductase, a fatty aldehyde reductase, and a fatty aldehyde decarbonylase.

[0080] As used herein, the term "hydrogen:carbon ratio" refers to the ratio of hydrogen atoms to carbon atoms in a molecule on an atom-to-atom basis. The ratio may be used to refer to the number of carbon and hydrogen atoms in a hydrocarbon molecule. For example, the hydrocarbon with the highest ratio is methane CH₄ (4:1).

[0081] As used herein, the term "hydrophobic fraction" refers to a portion, or fraction, of a material that is more soluble in a hydrophobic phase in comparison to an aqueous phase. A hydrophobic fraction is substantially insoluble in water and usually non-polar.

[0082] As used herein, the phrase "increase lipid yield" refers to an increase in the productivity of a microbial culture by, for example, increasing dry weight of cells per liter of culture, increasing the percentage of cells that constitute lipid, or increasing the overall amount of lipid per liter of culture volume per unit time.

[0083] An "inducible promoter" is one that mediates transcription of an operably linked gene in response to a particular stimulus.

[0084] As used herein, the term "in situ" means "in place" or "in its original position." For example, a culture may contain a first microalgae secreting a catalyst and a second microorganism secreting a substrate, wherein the first and second cell types produce the components necessary for a particular chemical reaction to occur in situ in the co-culture without requiring further separation or processing of the materials.

[0085] As used herein, the term "isomers" refers to compounds of the present invention that possess asymmetric carbon atoms (optical centers) or double bonds. The racemates, diastereomers, geometric isomers and individual isomers are all intended to be encompassed within the scope of the present invention.

[0086] As used herein, the phrase "limiting concentration of a nutrient" refers to a concentration in a culture that limits the propagation of a cultured organism. A "non-limiting

concentration of a nutrient” is a concentration that supports maximal propagation during a given culture period. Thus, the number of cells produced during a given culture period is lower in the presence of a limiting concentration of a nutrient than when the nutrient is non-limiting. A nutrient is said to be “in excess” in a culture, when the nutrient is present at a concentration greater than that which supports maximal propagation.

[0087] As used herein, a “lipase” is an enzyme that catalyzes the hydrolysis of ester bonds in water-insoluble, lipid substrates. Lipases catalyze the hydrolysis of lipids into glycerols and fatty acids.

[0088] “Lipids” are a class of molecules that are soluble in nonpolar solvents (such as ether and chloroform) and are relatively or completely insoluble in water. Lipid molecules have these properties because they consist largely of long hydrocarbon tails that are hydrophobic in nature. Examples of lipids include fatty acids (saturated and unsaturated); glycerides or glycerolipids (such as monoglycerides, diglycerides, triglyceries or neutral fats and phosphoglycerides or glycerophospholipids); nonglycerides (sphingolipids, sterol lipids including cholesterol and steroid hormones, prenol lipids including terpenoids, fatty alcohols, waxes and polyketides); and complex lipid derivatives (sugar-linked lipids, or glycolipids, and protein-linked lipids). “Fats” are a subgroup of lipids called “triacylglycerides”.

[0089] As used herein, the phrase “lipid:organic solvent composition” refers to a mixture of lipid and organic solvent.

[0090] As used herein, a “lipid pathway enzyme” is any enzyme that plays a role in lipid metabolism, i.e., either lipid synthesis, modification, or degradation. This term encompasses proteins that chemically modify lipids, as well as carrier proteins.

[0091] As used herein, the term “lysate” refers to a solution containing the contents of lysed cells.

[0092] As used herein, the term “lysis” refers to the breakage of the plasma membrane and optionally the cell wall of a biological organism sufficient to release at least some intracellular content, often by mechanical, viral or osmotic mechanisms that compromise its integrity.

[0093] As used herein, the term “lysing” refers to disrupting the cellular membrane and optionally the cell wall of a biological organism or cell sufficient to release at least some intracellular content.

[0094] As used herein, the term “microalgae” means a microbial organism that is either (a) eukaryotic and contains a chloroplast or chloroplast remnant, or (b) is a cyanobacteria. Microalgae include obligate photoautotrophs, which cannot metabolize a fixed carbon source as energy, as well as heterotrophs, which can live solely off of a fixed carbon source. Microalgae can refer to unicellular organisms that separate from sister cells shortly after cell division, such as *Chlamydomonas*, and can also refer to microbes such as, for example, *Volvox*, which is a simple multicellular photosynthetic microbe of two distinct cell types. “Microalgae” can also refer to cells such as *Chlorella* and *Dunaliella*. “Microalgae” also includes other microbial photosynthetic organisms that exhibit cell-cell adhesion, such as *Agmenellum*, *Anabaena*, and *Pyrobotrys*, as well as organisms that contain chloroplast-like structures that are no longer capable of performing photosynthesis, such as microalgae of the genus *Prototheca* and some dinoflagellates. “Microalgae” also includes obligate heterotrophic microorganisms that have lost the ability to perform photosynthesis, such as certain dinoflagellate species.

[0095] The terms “microorganism” and “microbe” are used interchangeably herein to refer to microscopic unicellular organisms.

[0096] As used herein, the term “oil” means a hydrophobic, lipophilic, nonpolar carbon-containing substance including but not limited to geologically-derived crude oil, distillate fractions of geologically-derived crude oil, vegetable oil, algal oil, and microbial lipids.

[0097] As used herein, the term “oleaginous yeast,” means yeast that can accumulate more than 10% of its dry cell weight as lipid. Oleaginous yeast includes organisms such as *Yarrowia lipolytica*, as well as engineered strains of yeast such as *Saccharomyces cerevisiae* that have been engineered to accumulate more than 10% of the dry cell weight as lipid.

[0098] As used herein, the terms “operably linked,” “in operable linkage,” or grammatical equivalents thereof refer to a functional linkage between two sequences, such a control sequence (typically a promoter) and the linked sequence. A promoter is in operable linkage with an exogenous gene if it can mediate transcription of the gene.

[0099] As used herein, the term “organic solvent” refers to a carbon-containing material that dissolves a solid, liquid, or gaseous solute, resulting in a solution.

[0100] As used herein, the term "osmotic shock" refers to the rupture of bacterial, algal, or other cells in a solution following a sudden reduction in osmotic pressure. Osmotic shock is sometimes induced to release cellular components of such cells into a solution.

[0101] As used herein, the term "photobioreactor" refers to a container, at least part of which is at least partially transparent or partially open, thereby allowing light to pass through, in which, e.g., one or more microalgae cells are cultured. Photobioreactors may be closed, as in the instance of a polyethylene bag or Erlenmeyer flask, or may be open to the environment, as in the instance of an outdoor pond.

[0102] As used herein, the term "polysaccharide" (also called "glycan") refers to carbohydrate made up of monosaccharides joined together by glycosidic linkages. Cellulose is an example of a polysaccharide that makes up certain plant cell walls. Cellulose can be depolymerized by enzymes to yield monosaccharides such as xylose and glucose, as well as larger disaccharides and oligosaccharides.

[0103] As used herein, the term "polysaccharide-degrading enzyme" refers to any enzyme capable of catalyzing the hydrolysis, or depolymerization, of any polysaccharide. For example, cellulase catalyzes the hydrolysis of cellulose.

[0104] As used herein, the term "port," in the context of a bioreactor, refers to an opening in the bioreactor that allows influx or efflux of materials such as gases, liquids, and cells. Ports are usually connected to tubing leading from the photobioreactor.

[0105] A "promoter" is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription.

[0106] As used herein, the term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, e.g., recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under

expressed or not expressed at all. By the term “recombinant nucleic acid” herein is meant nucleic acid, originally formed *in vitro*, in general, by the manipulation of nucleic acid, e.g., using polymerases and endonucleases, in a form not normally found in nature. In this manner, operably linkage of different sequences is achieved. Thus an isolated nucleic acid, in a linear form, or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined, are both considered recombinant for the purposes of this invention. It is understood that once a recombinant nucleic acid is made and reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the *in vivo* cellular machinery of the host cell rather than *in vitro* manipulations; however, such nucleic acids, once produced recombinantly, although subsequently replicated non-recombinantly, are still considered recombinant for the purposes of the invention. Similarly, a “recombinant protein” is a protein made using recombinant techniques, i.e., through the expression of a recombinant nucleic acid as depicted above.

[0107] As used herein, the term “renewable diesel” refers to a mixture of alkanes (such as C10:0, C12:0, C14:0, C16:0 and C18:0) produced through hydrogenation and deoxygenation of lipids. Renewable diesel also includes diesel fuel derived from biomass as defined in Section 45K(c)(3), using the process of thermal depolymerization (EPA 2005).

[0108] As used herein, the term “sonication” refers to a process of disrupting biologic materials, such as a cell, by use of sound wave energy.

[0109] As used herein, the term “wastewater” refers to a watery waste which typically contains washing water, laundry waste, faeces, urine, and other liquid or semi-liquid wastes. It includes some forms of municipal waste as well as secondarily treated sewage.

II. GENERAL

[0110] U.S. Patent application Nos: 60/941,581; 60/959,174; 60/968,291; and 61/024,069 are hereby incorporated by reference in their entirety for all purposes.

[0111] The invention generally relates to the production of hydrocarbon compositions, such as a lipid, in microorganisms. In particular, the invention provides methods for extracting, recovering, isolating and obtaining a lipid from a microorganism and compositions comprising the lipid. The invention also discloses methods for producing hydrocarbon compositions for use as biodiesel, renewable diesel, jet fuel, and for producing a lipid surfactant having a carbon chain length of C12 and C14.

[0112] The invention is premised in part on the insight that certain microorganisms can be used to produce hydrocarbon compositions economically and in large quantities for use in the transportation fuel and petrochemical industry among other applications. Suitable microorganisms include microalgae, oleaginous yeast, and fungi. A preferred genus of microalgae for use in the invention is the lipid-producing microalgae *Chlorella*. Acidic transesterification of lipids yields long-chain fatty acid esters useful as biodiesel. Other enzymatic processes can be tailored to yield fatty acids, aldehydes, alcohols, and alkanes. The present application describes methods for genetic modification of multiple species and strains of microorganisms, including *Chlorella* and similar microbes to provide organisms having characteristics that facilitate the production of lipids suitable for conversion into biodiesel or other hydrocarbon compounds. The present application also describes methods of cultivating microalgae for increased productivity and increased lipid yield.

[0113] Microorganisms useful in the invention produce lipids or hydrocarbons suitable for biodiesel production or as feedstock for industrial applications. Suitable hydrocarbons for biodiesel production include triacylglycerides (TAGs) containing long-chain fatty acid molecules. Suitable hydrocarbons for industrial applications, such as surfactant manufacturing, include fatty acids, aldehydes, alcohols, and alkanes. In some embodiments, suitable fatty acids, or the corresponding primary alcohols, aldehydes, or alkanes, generated by the methods described herein, contain at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, or at least 35 carbon atoms or more. Preferred long-chain fatty acids for biodiesel generally contain at least 14 carbon atoms or more.

[0114] Preferred fatty acids, or the corresponding primary alcohols, aldehydes, and alkanes, for industrial applications contain at least 8 carbon atoms or more. In certain embodiments, the above fatty acids, as well as the other corresponding hydrocarbon molecules, are saturated (with no carbon-carbon double or triple bonds); mono unsaturated (single double bond); poly unsaturated (two or more double bonds); are linear (not cyclic); and/or have little or no branching in their structures.

[0115] Triacylglycerols containing carbon chain lengths in the C8 to C22 range are preferred. Preferred for surfactants are C10-C14. Preferred for biodiesel or renewable diesel

are C16 to C18. Preferred for jet fuel are C8-C10. Preferred for nutrition are C22 polyunsaturated fatty acids (such as DHA) and carotenoids (such as astaxanthin).

III. MICROORGANISMS USEFUL FOR PRODUCING LIPIDS

[0116] Any species of organism that produces suitable lipid or hydrocarbon can be used, although microorganisms that naturally produce high levels of suitable lipid or hydrocarbon are preferred. Production of hydrocarbons by microorganisms is reviewed by Metzger *et al. Appl Microbiol Biotechnol* (2005) 66: 486–496 and A Look Back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae, NREL/TP-580-24190, John Sheehan, Terri Dunahay, John Benemann and Paul Roessler (1998).

[0117] Considerations affecting the selection of a microorganism for use in the invention include, in addition to production of suitable hydrocarbons for biodiesel or for industrial applications: (1) high lipid content as a percentage of cell weight; (2) ease of growth; (3) ease of genetic engineering; and (3) ease of processing. In particular embodiments, the wild-type or genetically engineered microorganism yields cells that are at least: about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, or about 70% or more lipid. Preferred organisms grow heterotrophically (on sugar in the absence of light) or can be engineered to do so using, for example, methods disclosed in commonly-owned U.S. Patent Application Nos. 60/837,839 and 60/968,291, which are incorporated herein by reference in their entireties. The ease of transformation and availability of selectable markers and promoters, constitutive and/or inducible, that are functional in the microorganism affect the ease of genetic engineering. Processing considerations can include, for example, the availability of effective means for lysing the cells.

A. Algae

[0118] In a preferred embodiment of the present invention, a microorganism producing a lipid or a microorganism from which a lipid can be extracted, recovered, or obtained, is an algae.

[0119] Examples of algae that can be used to practice the present invention include, but are not limited to the following algae listed in Table 1.

[0120] Table 1. Examples of algae.

<i>Achnanthes orientalis</i> , <i>Agmenellum</i> , <i>Amphiprora hyaline</i> , <i>Amphora coffeiformis</i> , <i>Amphora</i>

coffeiformis lineae, *Amphora coffeiformis punctata*, *Amphora coffeiformis taylori*, *Amphora coffeiformis tenuis*, *Amphora delicatissima*, *Amphora delicatissima capitata*, *Amphora* sp., *Anabaena*, *Ankistrodesmus*, *Ankistrodesmus falcatus*, *Boekelovia hooglandii*, *Borodinella* sp., *Botryococcus braunii*, *Botryococcus sudeticus*, *Carteria*, *Chaetoceros gracilis*, *Chaetoceros muelleri*, *Chaetoceros muelleri subsalsum*, *Chaetoceros* sp., *Chlorella anitrata*, *Chlorella Antarctica*, *Chlorella aureoviridis*, *Chlorella candida*, *Chlorella capsulate*, *Chlorella desiccata*, *Chlorella ellipsoidea*, *Chlorella emersonii*, *Chlorella fusca*, *Chlorella fusca* var. *vacuolata*, *Chlorella glucotropha*, *Chlorella infusionum*, *Chlorella infusionum* var. *actophila*, *Chlorella infusionum* var. *auxenophila*, *Chlorella kessleri*, *Chlorella lobophora* (strain SAG 37.88), *Chlorella luteoviridis*, *Chlorella luteoviridis* var. *aureoviridis*, *Chlorella luteoviridis* var. *lutescens*, *Chlorella miniata*, *Chlorella minutissima*, *Chlorella mutabilis*, *Chlorella nocturna*, *Chlorella parva*, *Chlorella photophila*, *Chlorella pringsheimii*, *Chlorella protothecoides* (including any of UTEX strains 1806, 411, 264, 256, 255, 250, 249, 31, 29, 25, and CCAP strains 211/17 and 211/8d), *Chlorella protothecoides* var. *acidicola*, *Chlorella regularis*, *Chlorella regularis* var. *minima*, *Chlorella regularis* var. *umbricata*, *Chlorella reisiigii*, *Chlorella saccharophila*, *Chlorella saccharophila* var. *ellipsoidea*, *Chlorella salina*, *Chlorella simplex*, *Chlorella sorokiniana*, *Chlorella* sp., *Chlorella sphaerica*, *Chlorella stigmatophora*, *Chlorella vanniellii*, *Chlorella vulgaris*, *Chlorella vulgaris*, *Chlorella vulgaris* f. *tertia*, *Chlorella vulgaris* var. *autotrophica*, *Chlorella vulgaris* var. *viridis*, *Chlorella vulgaris* var. *vulgaris*, *Chlorella vulgaris* var. *vulgaris* f. *tertia*, *Chlorella vulgaris* var. *vulgaris* f. *viridis*, *Chlorella xanthella*, *Chlorella zofingiensis*, *Chlorella trebouxioidea*, *Chlorella vulgaris*, *Chlorococcum infusionum*, *Chlorococcum* sp., *Chlorogonium*, *Chroomonas* sp., *Chrysosphaera* sp., *Cricosphaera* sp., *Cryptomonas* sp., *Cyclotella cryptica*, *Cyclotella meneghiniana*, *Cyclotella* sp., *Dunaliella* sp., *Dunaliella bardawil*, *Dunaliella bioculata*, *Dunaliella granulate*, *Dunaliella maritime*, *Dunaliella minuta*, *Dunaliella parva*, *Dunaliella peircei*, *Dunaliella primolecta*, *Dunaliella salina*, *Dunaliella terricola*, *Dunaliella tertiolecta*, *Dunaliella viridis*, *Dunaliella tertiolecta*, *Eremosphaera viridis*, *Eremosphaera* sp., *Ellipsoidon* sp., *Euglena*, *Franceia* sp., *Fragilaria crotonensis*, *Fragilaria* sp., *Gleocapsa* sp., *Gloeothamnion* sp., *Hymenomonas* sp., *Isochrysis* aff. *galbana*, *Isochrysis galbana*, *Lepocinclis*, *Micractinium*, *Micractinium* (UTEX LB 2614), *Monoraphidium minutum*, *Monoraphidium* sp., *Nannochloris* sp., *Nannochloropsis salina*, *Nannochloropsis* sp., *Navicula acceptata*, *Navicula biskanterae*, *Navicula pseudotenelloides*, *Navicula pelliculosa*, *Navicula saprophila*, *Navicula* sp., *Nephrochloris* sp., *Nephroselmis* sp., *Nitzschia communis*,

Nitzschia alexandrina, *Nitzschia communis*, *Nitzschia dissipata*, *Nitzschia frustulum*, *Nitzschia hantzschiana*, *Nitzschia inconspicua*, *Nitzschia intermedia*, *Nitzschia microcephala*, *Nitzschia pusilla*, *Nitzschia pusilla elliptica*, *Nitzschia pusilla monoensis*, *Nitzschia quadrangular*, *Nitzschia sp.*, *Ochromonas sp.*, *Oocystis parva*, *Oocystis pusilla*, *Oocystis sp.*, *Oscillatoria limnetica*, *Oscillatoria sp.*, *Oscillatoria subbrevis*, *Pascheria acidophila*, *Pavlova sp.*, *Phagus*, *Phormidium*, *Platymonas sp.*, *Pleurochrysis carterae*, *Pleurochrysis dentate*, *Pleurochrysis sp.*, *Prototheca wickerhamii*, *Prototheca stagnora*, *Prototheca portoricensis*, *Prototheca moriformis*, *Prototheca zopfii*, *Pyramimonas sp.*, *Pyrobotrys*, *Sarcinoid chrysophyte*, *Scenedesmus armatus*, *Spirogyra*, *Spirulina platensis*, *Stichococcus sp.*, *Synechococcus sp.*, *Tetraedron*, *Tetraselmis sp.*, *Tetraselmis suecica*, *Thalassiosira weissflogii*, and *Viridiella fridericana*

1. *Chlorella* and *Prototheca*

[0121] In a preferred embodiment of the present invention, the microorganism producing a lipid or a microorganism from which a lipid can be extracted, recovered, or obtained, is of the genus *Chlorella*, preferably, *Chlorella protothecoides*.

[0122] *Chlorella* is a genus of single-celled green algae, belonging to the phylum Chlorophyta. It is spherical in shape, about 2 to 10 μm in diameter, and is without flagella. Some species of *Chlorella* are naturally heterotrophic.

[0123] *Chlorella*, particularly *Chlorella protothecoides*, is a preferred microorganism for use in the invention because of its high composition (at least about 55% by weight) of lipid, particularly long-chain lipid suitable for biodiesel. In addition, this microalgae grows heterotrophically, can be genetically engineered as demonstrated in the Examples herein, and can be lysed by *Chlorella* virus.

[0124] In a preferred embodiment of the present invention, the microorganism used for expression of a transgene is of the genus *Chlorella*, preferably, *Chlorella protothecoides*. Examples of expression of transgenes in, e.g., *Chlorella*, can be found in the literature (see for example *Current Microbiology* Vol. 35 (1997), pp. 356–362; Sheng Wu Gong Cheng Xue Bao. 2000 Jul;16(4):443-6; *Current Microbiology* Vol. 38 (1999), pp. 335–341; *Appl Microbiol Biotechnol* (2006) 72: 197–205; *Marine Biotechnology* 4, 63–73, 2002; *Current Genetics* 39:5, 365-370 (2001); *Plant Cell Reports* 18:9, 778-780, (1999); *Biologia Plantarum* 42(2): 209-216, (1999); *Plant Pathol. J* 21(1): 13-20, (2005)). Also see

Examples herein. Other lipid-producing microbes can be engineered as well, including prokaryotic microbes (see Kalscheuer et al., Applied Microbiology and Biotechnology, volume 52, number 4/October, 1999).

[0125] In another preferred embodiment of the present invention, the microorganism producing a lipid or a microorganism from which a lipid can be extracted, recovered, or obtained, is of the genus *Prototheca*, preferably, *Prototheca moriformis*. Species of the genus *Prototheca* are suited for the production of lipid because they can produce high levels of lipids, particularly lipids suitable for fuel and chemical production. The lipid produced by *Prototheca* has fatty acid chains of shorter chain length and a higher degree of saturation than the lipid produced by most microalgae. Moreover, *Prototheca* lipid is generally free of pigment (low to undetectable levels of chlorophyll and certain carotenoids) and in any event contains much less pigment than lipid from other microalgae. Illustrative *Prototheca* strains for use in the methods of the invention include *Prototheca wickerhamii*, *Prototheca stagnora* (including UTEX 327), *Prototheca portoricensis*, *Prototheca moriformis* (including UTEX strains 1441, 1435), and *Prototheca zopfii*. Species of the genus *Prototheca* are obligate heterotrophs.

2. Identification of Microalgal Species

[0126] Species of microalgae, including *Chlorella* and *Prototheca*, for use in the invention can be identified by amplification of certain target regions of the genome. For example, identification of a specific *Chlorella* species can be achieved through amplification and sequencing of nuclear and/or chloroplast DNA using primers and methodology described in Wu *et al.*, Identification of *Chlorella spp.* isolates using any region of the genome. Examples include ribosomal DNA sequences, *Bot. Bull. Acad. Sin.* (2001) 42:115-121. Well established methods of phylogenetic analysis, such as amplification and sequencing of ribosomal internal transcribed spacer (ITS1 and ITS2 rDNA), 18S rRNA, and other conserved genomic regions can be used by those skilled in the art to identify species of not only *Chlorella* or *Prototheca*, but other hydrocarbon and lipid producing organisms capable of using the methods disclosed herein. For examples of methods of identification and classification of algae also see for example *Genetics*, 2005 Aug;170(4):1601-10 and *RNA*, 2005 Apr;11(4):361-4.

[0127] Genomic DNA comparison can be used to identify species of microalgae to be used in the present invention. Regions of conserved genomic DNA, such as but not limited to

DNA encoding 23S rRNA, can be amplified from microalgal species and compared to consensus sequences in order to screen for microalgal species that are taxonomically related to the preferred microalgae used in the present invention. Examples of such DNA sequence comparison for species within the *Chlorella* genus are shown below.

[0128] Genomic DNA comparison can also be useful to identify microalgal species that have been misidentified in a strain collection. Often a strain collection will identify species of microalgae based on phenotypic and morphological characteristics. The use of these characteristics may lead to miscategorization of the species or the genus of a microalgae. The use of genomic DNA comparison can be a better method of categorizing microalgae species based on their phylogenetic relationship. Specific examples of using genotyping data to establish phylogenetic relationships of possibly misidentified microalgal strains are described below in the Examples.

[0129] In some cases, microalgae that are preferred for use in the present invention have genomic DNA sequences encoding for 23S rRNA that have at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 89%, at least 88%, at least 87%, or at least 86% nucleotide identity to at least one of the sequences listed in SEQ ID NOs: 7-31. In other cases, microalgae that are preferred for use in the present invention have genomic DNA sequences encoding for 23S rRNA that have at least 85%, at least 80%, at least 75% at least 70% at least 65% or at least 60% nucleotide identity to at least one of the sequences listed in SEQ ID NOs.7-31, 32, 33, 34, 35, 36, 37, 38, and 39.

B. Oleaginous Yeast

[0130] In a preferred embodiment of the present invention, a microorganism producing a lipid or a microorganism from which a lipid can be extracted, recovered, or obtained, is an oleaginous yeast.

[0131] Examples of oleaginous yeast that can be used to practice the present invention include, but are not limited to the following oleaginous yeast listed in Table 2.

[0132] Table 2. Examples of oleaginous yeast.

Candida apicola, *Candida sp.*, *Cryptococcus curvatus*, *Cryptococcus terricolus*,
Debaromyces hansenii, *Endomycopsis vernalis*, *Geotrichum carabidarum*, *Geotrichum cucujoidarum*, *Geotrichum histeridarum*, *Geotrichum silvicola*, *Geotrichum vulgare*,

Hyphopichia burtonii, *Lipomyces lipofer*, *Lypomyces orentalis*, *Lipomyces starkeyi*, *Lipomyces tetrasporous*, *Pichia mexicana*, *Rodosporidium sphaerocarpum*, *Rhodosporidium toruloides*, *Rhodotorula aurantiaca*, *Rhodotorula dairenensis*, *Rhodotorula diffluens*, *Rhodotorula glutinus*, *Rhodotorula glutinis* var. *glutinis*, *Rhodotorula gracilis*, *Rhodotorula graminis*, *Rhodotorula minuta*, *Rhodotorula mucilaginosa*, *Rhodotorula mucilaginosa* var. *mucilaginosa*, *Rhodotorula terpenoidalis*, *Rhodotorula toruloides*, *Sporobolomyces alborubescens*, *Starmerella bombicola*, *Torulaspora delbruekii*, *Torulaspora pretoriensis*, *Trichosporon behrend*, *Trichosporon brassicae*, *Trichosporon domesticum*, *Trichosporon laibachii*, *Trichosporon loubieri*, *Trichosporon loubieri* var. *loubieri*, *Trichosporon montevidense*, *Trichosporon pullulans*, *Trichosporon sp.*, *Wickerhamomyces Canadensis*, *Yarrowia lipolytica*, and *Zygoascus meyeriae*.

[0133] In a preferred embodiment of the present invention, the microorganism used for expression of a transgene is an oleaginous yeast. Examples of expression of transgenes in oleaginous yeast (e.g., *Yarrowia lipolytica*) can be found in the literature (see, for example, Bordes *et al.*, *J Microbiol Methods*, Jun 27 (2007)).

C. Other Fungi

[0134] In a preferred embodiment of the present invention, a microorganism producing a lipid or a microorganism from which a lipid can be extracted, recovered, or obtained, is a fungus.

[0135] Examples of fungi that can be used to practice the present invention include, but are not limited to the following fungi listed in Table 3.

[0136] Table 3. Examples of non-yeast fungi.

Mortierella, *Mortierella vinacea*, *Mortierella alpine*, *Pythium debaryanum*, *Mucor circinelloides*, *Aspergillus ochraceus*, *Aspergillus terreus*, *Penicillium lilacinum*, *Hansenula*, *Chaetomium*, *Cladosporium*, *Malbranchea*, *Rhizopus*, and *Pythium*

[0137] In a preferred embodiment of the present invention, the microorganism used for expression of a transgene is a fungus. Examples of expression of transgenes in fungi (e.g., *Mortierella alpine*, *Mucor circinelloides*, and *Aspergillus ochraceus*) can also be found in the literature (see, for example, *Microbiology*, Jul; 153(Pt. 7):2013-25 (2007); *Mol Genet Genomics*, Jun; 271(5):595-602 (2004); *Curr Genet*, Mar; 21(3):215-23 (1992); *Current*

Microbiology, 30(2):83-86 (1995); Sakuradani, NISR Research Grant, "Studies of Metabolic Engineering of Useful Lipid-producing Microorganisms" (2004); and PCT/JP2004/012021).

D. Bacteria

[0138] In a preferred embodiment of the present invention, a microorganism producing a lipid or a microorganism from which a lipid can be extracted, recovered, or obtained, is a bacterium.

[0139] Examples of expression of exogenous genes in bacteria, such as *E. coli*, are well known; see for example *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.* (3d edition, 2001, Cold Spring Harbor Press).

IV. MICROBE ENGINEERING

[0140] In certain embodiments of the present invention, it is desirable to genetically modify a microorganism. Thus, the present application describes genetically engineering strains of microalgae, oleaginous yeast, bacteria, or fungi with one or more exogenous genes to produce various hydrocarbon compounds.

[0141] Promoters, cDNAs, and 3'UTRs, as well as other elements of expression vectors, can be generated through cloning techniques using fragments isolated from native sources (see for example *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.* (3d edition, 2001, Cold Spring Harbor Press; and U.S. Patent 4,683,202). Alternatively, elements can be generated synthetically using known methods (see for example *Gene*, 1995 Oct 16;164(1):49-53).

A. Codon-Optimization For Expression

[0142] DNA encoding a polypeptide to be expressed in a microorganism, e.g., a lipase and selectable marker are preferably codon-optimized cDNA. Methods of recoding genes for expression in microalgae are described in U.S. Patent 7,135,290. Additional information for codon optimization is available, e.g., at the codon usage database of GenBank. As non-limiting examples, codon usage in *Prototheca* species, *Dunaliella salina*, and *Chlorella protothecoides* are shown in Tables 4, 5, and 6, respectively.

[0143] Table 4. Codon usage in *Prototheca* species

Ala	GCG	345 (0.36)	Asn	AAT	8 (0.04)
	GCA	66 (0.07)		AAC	201 (0.96)

	GCT	101 (0.11)			
	GCC	442 (0.46)		Pro	CCG 161 (0.29)
					CCA 49 (0.09)
Cys	TGT	12 (0.10)			CCT 71 (0.13)
	TGC	105 (0.90)			CCC 267 (0.49)
Asp	GAT	43 (0.12)		Gln	CAG 226 (0.82)
	GAC	316 (0.88)			CAA 48 (0.18)
Glu	GAG	377 (0.96)		Arg	AGG 33 (0.06)
	GAA	14 (0.04)			AGA 14 (0.02)
					CGG 102 (0.18)
Phe	TTT	89 (0.29)			CGA 49 (0.08)
	TTC	216 (0.71)			CGT 51 (0.09)
					CGC 331 (0.57)
Gly	GGG	92 (0.12)			
	GGA	56 (0.07)		Ser	AGT 16 (0.03)
	GGT	76 (0.10)			AGC 123 (0.22)
	GGC	559 (0.71)			TCG 152 (0.28)
					TCA 31 (0.06)
His	CAT	42 (0.21)			TCT 55 (0.10)
	CAC	154 (0.79)			TCC 173 (0.31)
Ile	ATA	4 (0.01)		Thr	ACG 184 (0.38)
	ATT	30 (0.08)			ACA 24 (0.05)
	ATC	338 (0.91)			ACT 21 (0.05)
					ACC 249 (0.52)
Lys	AAG	284 (0.98)			
	AAA	7 (0.02)		Val	GTG 308 (0.50)
					GTA 9 (0.01)
Leu	TTG	26 (0.04)			GTT 35 (0.06)
	TTA	3 (0.00)			GTC 262 (0.43)
	CTG	447 (0.61)			
	CTA	20 (0.03)		Trp	TGG 107 (1.00)
	CTT	45 (0.06)			
	CTC	190 (0.26)		Tyr	TAT 10 (0.05)
					TAC 180 (0.95)
Met	ATG	191 (1.00)			
				Stop	TGA/TAG/TAA

[0144] Table 5. Preferred codon usage in *Dunaliella salina*.

TTC (Phe)	TAC (Tyr)	TGC (Cys)	TAA (Stop)
TGG (Trp)	CCC (Pro)	CAC (His)	CGC (Arg)
CTG (Leu)	CAG (Gln)	ATC (Ile)	ACC (Thr)
AAC (Asn)	AGC (Ser)	ATG (Met)	AAG (Lys)
GCC (Ala)	GAC (Asp)	GGC (Gly)	GTG (Val)
GAG (Glu)			

[0145] Table 6. Preferred codon usage in *Chlorella protothecoides*.

TTC (Phe)	TAC (Tyr)	TGC (Cys)	TGA (Stop)
TGG (Trp)	CCC (Pro)	CAC (His)	CGC (Arg)
CTG (Leu)	CAG (Gln)	ATC (Ile)	ACC (Thr)
GAC (Asp)	TCC (Ser)	ATG (Met)	AAG (Lys)
GCC (Ala)	AAC (Asn)	GGC (Gly)	GTG (Val)
GAG (Glu)			

B. Promoters

[0146] Many promoters are active in microalgae, including promoters that are endogenous to the algae being transformed, as well as promoters that are not endogenous to the algae being transformed (i.e., promoters from other algae, promoters from higher plants, and promoters from plant viruses or algae viruses). Exogenous and/or endogenous promoters that are active in microalgae, and antibiotic resistance genes functional in microalgae are described by e.g., *Curr Microbiol.* 1997 Dec;35(6):356-62 (*Chlorella vulgaris*); *Mar Biotechnol* (NY). 2002 Jan;4(1):63-73 (*Chlorella ellipsoidea*); *Mol Gen Genet.* 1996 Oct 16;252(5):572-9 (*Phaeodactylum tricornutum*); *Plant Mol Biol.* 1996 Apr;31(1):1-12 (*Volvox carteri*); *Proc Natl Acad Sci U S A.* 1994 Nov 22;91(24):11562-6 (*Volvox carteri*); Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C, PMID: 10383998, 1999 May;1(3):239-251 (Laboratory of Molecular Plant Biology, Stazione Zoologica, Villa Comunale, I-80121 Naples, Italy) (*Phaeodactylum tricornutum* and *Thalassiosira weissflogii*); *Plant Physiol.* 2002 May;129(1):7-12. (*Porphyridium sp.*); *Proc Natl Acad Sci U S A.* 2003 Jan 21;100(2):438-42. (*Chlamydomonas reinhardtii*); *Proc Natl Acad Sci U S A.* 1990 Feb;87(3):1228-32. (*Chlamydomonas reinhardtii*); *Nucleic Acids Res.* 1992 Jun 25;20(12):2959-65; *Mar Biotechnol* (NY). 2002 Jan;4(1):63-73 (*Chlorella*); *Biochem Mol Biol Int.* 1995 Aug;36(5):1025-35 (*Chlamydomonas reinhardtii*); *J Microbiol.* 2005 Aug;43(4):361-5 (*Dunaliella*); Yi Chuan Xue Bao. 2005 Apr;32(4):424-33 (*Dunaliella*); *Mar Biotechnol* (NY). 1999 May;1(3):239-251. (*Thalassiosira* and *Phaeodactylum*); Koksharova, *Appl Microbiol Biotechnol* 2002 Feb;58(2):123-37 (various species); *Mol Genet Genomics* 2004 Feb;271(1):50-9 (*Thermosynechococcus elongates*); *J. Bacteriol.* (2000), 182, 211-215; *FEMS Microbiol Lett.* 2003 Apr 25;221(2):155-9; *Plant Physiol.* 1994 Jun;105(2):635-41; *Plant Mol Biol.* 1995 Dec;29(5):897-907 (*Synechococcus* PCC 7942); *Mar Pollut Bull.* 2002;45(1-12):163-7 (*Anabaena* PCC 7120); *Proc Natl Acad Sci U S A.* 1984

Mar;81(5):1561-5 (*Anabaena* (various strains)); *Proc Natl Acad Sci U S A*. 2001 Mar 27;98(7):4243-8 (*Synechocystis*); Wirth, *Mol Gen Genet* 1989 Mar;216(1):175-7 (various species); *Mol Microbiol*, 2002 Jun;44(6):1517-31; *Plasmid*, 1993 Sep;30(2):90-105 (*Fremyella diplosiphon*); Hall *et al.* (1993) *Gene* 124: 75-81 (*Chlamydomonas reinhardtii*); Gruber *et al.* (1991). *Current Micro.* 22: 15-20; Jarvis *et al.* (1991) *Current Genet.* 19: 317-322 (*Chlorella*); for additional promoters see also Table 1 from U.S. Patent 6,027,900.

[0147] The promoter used to express an exogenous gene can be the promoter naturally linked to that gene or can be a heterologous gene. Some promoters are active in more than one species of microalgae. Other promoters are species-specific. Preferred promoters include promoters such as RBCS2 from *Chlamydomonas reinhardtii* and viral promoters, such as cauliflower mosaic virus (CMV) and chlorella virus, which have been shown to be active in multiple species of microalgae (see for example *Plant Cell Rep.* 2005 Mar;23(10-11):727-35; *J Microbiol.* 2005 Aug;43(4):361-5; *Mar Biotechnol* (NY). 2002 Jan;4(1):63-73). In other embodiments, the *Botryococcus malate* dehydrogenase promoter, such a nucleic acid comprising any part of SEQ ID NO:3, or the *Chlamydomonas reinhardtii* RBCS2 promoter (SEQ ID NO:4) can be used. Optionally, at least 10, 20, 30, 40, 50, or 60 nucleotides or more of these sequences containing a promoter are used. Preferred promoters endogenous to species of the genus *Chlorella* are SEQ ID NOs: 1 and 2.

[0148] Preferred promoters useful for expression of exogenous genes in *Chlorella* are listed in the sequence listing of this application, such as the promoter of the *Chlorella* HUP1 gene (SEQ ID NO:1) and the *Chlorella ellipsoidea* nitrate reductase promoter (SEQ ID NO:2). *Chlorella* virus promoters can also be used to express genes in *Chlorella*, such as SEQ ID NOs: 1-7 of U.S. Patent 6,395,965. Additional promoters active in *Chlorella* can be found, for example, in *Biochem Biophys Res Commun*, 1994 Oct 14;204(1):187-94; *Plant Mol Biol*, 1994 Oct;26(1):85-93; *Virology*, 2004 Aug 15;326(1):150-9; and *Virology*, 2004 Jan 5;318(1):214-23.

C. Selectable Markers

[0149] Any of a wide variety of selectable markers can be employed in a transgene construct useful for transforming a microorganism, e.g., *Chlorella*. Examples of suitable selectable markers include the nitrate reductase gene, the hygromycin phosphotransferase gene (HPT), the neomycin phosphotransferase gene, and the ble gene, which confers

resistance to phleomycin. Methods of determining sensitivity of microalgae to antibiotics are well known. For example, *Mol Gen Genet*, 1996 Oct 16;252(5):572-9.

[0150] More specifically, Dawson *et al.* (1997), *Current Microbiology* 35:356-362 (incorporated by reference herein in its entirety), described the use of the nitrate reductase (NR) gene from *Chlorella vulgaris* as a selectable marker for NR-deficient *Chlorella sorokiniana* mutants. Kim *et al.* (2002), *Mar. Biotechnol.* 4:63-73 (incorporated by reference herein in its entirety), disclosed the use of the HPT gene as a selectable marker for transforming *Chlorella ellipsoidea*. Huang *et al.* (2007), *Appl. Microbiol. Biotechnol.* 72:197-205 (incorporated by reference herein in its entirety), reported on the use of Sh ble as a selectable marker for *Chlorella sp. DT*.

D. Inducible Expression

[0151] The present invention also provides for the use of an inducible promoter to express a gene of interest. In particular, the use of an inducible promoter to express a lipase gene permits production of the lipase after growth of the microorganism when conditions have been adjusted, if necessary, to enhance transesterification, for example, after disruption of the cells, reduction of the water content of the reaction mixture, and/or addition sufficient alcohol to drive conversion of TAGs to fatty acid esters.

[0152] Inducible promoters useful in the invention include those that mediate transcription of an operably linked gene in response to a stimulus, such as an exogenously provided small molecule (e.g., glucose, as in SEQ ID NO:1), temperature (heat or cold), light, etc. Suitable promoters can activate transcription of an essentially silent gene or upregulate, preferably substantially, transcription of an operably linked gene that is transcribed at a low level. In the latter case, the level of transcription of the gene of interest, e.g., the lipase gene, preferably does not significantly interfere with the growth of the microorganism in which it is expressed.

[0153] Expression of a transgene in *Chlorella* can be performed under inducible conditions, e.g., using promoters such as the promoter for the *Chlorella* hexose transporter gene (SEQ ID NO:1). This promoter is strongly activated by the presence of glucose in the culture media.

E. Expression Of Two Or More Exogenous Genes

[0154] Further, a genetically engineered microorganism, such as a microalgae, may comprise and express two or more exogenous genes, such as, for example, a lipase and a lytic gene, e.g., one encoding a polysaccharide-degrading enzyme. One or both genes can be

expressed using an inducible promoter, which allows the relative timing of expression of these genes to be controlled to enhance the lipid yield and conversion to fatty acid esters. Expression of the two or more exogenous genes may be under control of the same inducible promoter or under control of a different inducible promoters. In the latter situation, expression of a first exogenous gene can be induced for a first period of time (during which expression of a second exogenous gene may or may not be induced) and expression of a second exogenous gene can be induced for a second period of time (during which expression of a first exogenous gene may or may not be induced). Provided herein are vectors and methods for engineering lipid-producing microbes to metabolize sucrose, which is an advantageous trait because it allows the engineered cells to convert sugar cane feedstocks into lipids appropriate for biodiesel production.

[0155] Also provided herein are genetically engineered strains of microbes (e.g., microalgae, oleaginous yeast, bacteria, or fungi) that express two or more exogenous genes, such as, for example, a fatty acyl-ACP thioesterase and a fatty acyl-CoA/aldehyde reductase, the combined action of which yields an alcohol product. Further provided are other combinations of exogenous genes, including without limitation, a fatty acyl-ACP thioesterase and a fatty acyl-CoA reductase to generate aldehydes. In addition, this application provides for the combination of a fatty acyl-ACP thioesterase, a fatty acyl-CoA reductase, and a fatty aldehyde decarbonylase to generate alkanes. One or more of the exogenous genes can be expressed using an inducible promoter.

[0156] Examples of further modifications suitable for use in the present invention are described in co-pending, commonly owned Application No. 60/837,839, which is incorporated herein by reference. This application discloses genetically engineering strains of microalgae to express two or more exogenous genes, one encoding a transporter of a fixed carbon source (such as sucrose) and a second encoding a sucrose invertase enzyme. The resulting fermentable organisms produce hydrocarbons at lower manufacturing cost than what has been obtainable by previously known methods of biological hydrocarbon production. This co-pending application also teaches that the insertion of the two exogenous genes described above can be combined with the disruption of polysaccharide biosynthesis through directed and/or random mutagenesis, which steers ever greater carbon flux into hydrocarbon production. Individually and in combination, trophic conversion, engineering to alter hydrocarbon production and treatment with exogenous enzymes alter the hydrocarbon composition produced by a microorganism. The alteration can be a change in the amount of

hydrocarbons produced, the amount of one or more hydrocarbon species produced relative to other hydrocarbons, and/or the types of hydrocarbon species produced in the microorganism. For example, microalgae can be engineered to produce a higher amount and/or percentage of TAGs.

F. Compartmentalized Expression

[0157] The present invention also provides for compartmentalized expression a gene of interest. In particular, it can be advantageous, in particular embodiments, to target expression of a lipase gene, to one or more cellular compartments, where it is sequestered from the majority of cellular lipids until initiation of the transesterification reaction. Preferred organelles for targeting are chloroplasts, mitochondria, and endoplasmic reticulum.

1. Expression in Chloroplasts

[0158] In a preferred embodiment of the present invention, the expression of a polypeptide in a microorganism is targeted to chloroplasts. Methods for targeting expression of a heterologous gene to the chloroplast are known and can be employed in the present invention. Methods for targeting foreign gene products into chloroplasts are described in Shrier *et al.*, *EMBO J.* (1985) 4:25 32. See also Tomai *et al.*, *Gen. Biol. Chem.* (1988) 263:15104 15109 and U.S. Pat. No. 4,940,835 for the use of transit peptides for translocating nuclear gene products into the chloroplast. Methods for directing the transport of proteins to the chloroplast are also reviewed in Kenauf *TIBTECH* (1987) 5:40 47. Chloroplast targeting sequences endogenous to *Chlorella* are known, such as genes in the *Chlorella* nuclear genome that encode proteins that are targeted to the chloroplast; see for example GenBank Accession numbers AY646197 and AF499684.

[0159] Wageningen UR-Plant Research International sells an IMPACTVECTOR1.4 vector, which uses the secretion signal of the *Chrysanthemum morifolium* small subunit protein to deliver a heterologous protein into the chloroplast stroma (cytoplasmic) environment, shuttling across a double membrane system. The protein is fused to the first 11 amino acids of the mature rubisco protein in order to allow proper processing of the signal peptide (Wong *et al.*, *Plant Molecular Biology* 20: 81-93 (1992)). The signal peptide contains a natural intron from the RbcS gene.

[0160] In another approach, the chloroplast genome is genetically engineered to express the heterologous protein. Stable transformation of chloroplasts of *Chlamydomonas reinhardtii* (a

green alga) using bombardment of recipient cells with high-velocity tungsten microprojectiles coated with foreign DNA has been described. See, for example, Boynton *et al.*, *Science* (1988) 240: 1534 1538; Blowers *et al.* *Plant Cell* (1989) 1:123 132 and Debuchy *et al.*, *EMBO J.* (1989) 8: 2803 2809. The transformation technique, using tungsten microprojectiles, is described by Klein *et al.*, *Nature* (London) (1987) 7:70 73. Other methods of chloroplast transformation for both plants and microalgae are known. See for example U.S. Patents 5,693,507; 6,680,426; *Plant Physiol.* 2002 May;129(1):7-12; and *Plant Biotechnol J.* 2007 May;5(3):402-12.

[0161] As described in U.S. Patent No. 6,320,101 (issued November 20, 2001 to Kaplan *et al.*; which is incorporated herein by reference), cells can be chemically treated so as to reduce the number of chloroplasts per cell to about one. Then, the heterologous nucleic acid can be introduced into the cells via particle bombardment with the aim of introducing at least one heterologous nucleic acid molecule into the chloroplasts. The heterologous nucleic acid is selected such that it is integratable into the chloroplast's genome via homologous recombination which is readily effected by enzymes inherent to the chloroplast. To this end, the heterologous nucleic acid includes, in addition to a gene of interest, at least one nucleic acid sequence that is derived from the chloroplast's genome. In addition, the heterologous nucleic acid typically includes a selectable marker. Further details relating to this technique are found in U.S. Patent Nos. 4,945,050 and 5,693,507, which are incorporated herein by reference. A polypeptide can thus be produced by the protein expression system of the chloroplast.

[0162] U.S. Patent No. 7,135,620 (issued November 14, 2006 to Daniell *et al.*; incorporated herein by reference) describes chloroplast expression vectors and related methods. Expression cassettes are DNA constructs including a coding sequence and appropriate control sequences to provide for proper expression of the coding sequence in the chloroplast. Typical expression cassettes include the following components: the 5' untranslated region from a microorganism gene or chloroplast gene such as *psbA* which will provide for transcription and translation of a DNA sequence encoding a polypeptide of interest in the chloroplast; a DNA sequence encoding a polypeptide of interest; and a translational and transcriptional termination region, such as a 3' inverted repeat region of a chloroplast gene that can stabilize RNA of introduced genes, thereby enhancing foreign gene expression. The cassette can optionally include an antibiotic resistance gene.

[0163] Typically, the expression cassette is flanked by convenient restriction sites for insertion into an appropriate genome. The expression cassette can be flanked by DNA sequences from chloroplast DNA to facilitate stable integration of the expression cassette into the chloroplast genome, particularly by homologous recombination. Alternatively, the expression cassette may remain unintegrated, in which case, the expression cassette typically includes a chloroplast origin of replication, which is capable of providing for replication of the heterologous DNA in the chloroplast.

[0164] The expression cassette generally includes a promoter region from a gene capable of expression in the chloroplast. The promoter region may include promoters obtainable from chloroplast genes, such as the *psbA* gene from spinach or pea, or the *rbcL* and *atpB* promoter region from maize and rRNA promoters. Examples of promoters are described in Hanley-Bowdoin and Chua, *TIBS* (1987) 12:67 70; Mullet *et al.*, *Plant Molec Biol.* (1985) 4: 39 54; Hanley-Bowdoin (1986) PhD. Dissertation, the Rockefeller University; Krebbers *et al.*, *Nucleic Acids Res.* (1982) 10: 4985 5002; Zurawaki *et al.*, *Nucleic Acids Res.* (1981) 9:3251 3270; and Zurawski *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* (1982) 79: 7699 7703. Other promoters can be identified and the relative strength of promoters so identified evaluated, by placing a promoter of interest 5' to a promoterless marker gene and observing its effectiveness relative to transcription obtained from, for example, the promoter from the *psbA* gene, a relatively strong chloroplast promoter. The efficiency of heterologous gene expression additionally can be enhanced by any of a variety of techniques. These include the use of multiple promoters inserted in tandem 5' to the heterologous gene, for example a double *psbA* promoter, the addition of enhancer sequences and the like.

[0165] Numerous promoters active in the *Chlorella* chloroplast can be used for expression of exogenous genes in the *Chlorella* chloroplast, such as those found in GenBank accession number NC_001865 (*Chlorella vulgaris* chloroplast, complete genome).

[0166] Where it is desired to provide for inducible expression of the heterologous gene, an inducible promoter and/or a 5' untranslated region containing sequences which provide for regulation at the level of transcription and/or translation (at the 3' end) may be included in the expression cassette. For example, the 5' untranslated region can be from a gene wherein expression is regulatable by light. Similarly, 3' inverted repeat regions could be used to stabilize RNA of heterologous genes. Inducible genes may be identified by enhanced expression in response to a particular stimulus of interest and low or absent expression in the

53

absence of the stimulus. For example, a light-inducible gene can be identified where enhanced expression occurs during irradiation with light, while substantially reduced expression or no expression occurs in low or no light. Light regulated promoters from green microalgae are known (see for example *Mol Genet Genomics*, 2005 Dec;274(6):625-36).

[0167] The termination region which is employed will be primarily one of convenience, since the termination region appears to be relatively interchangeable among chloroplasts and bacteria. The termination region may be native to the transcriptional initiation region, may be native to the DNA sequence of interest, or may be obtainable from another source. See, for example, Chen and Orozco, *Nucleic Acids Res.* (1988) 16:8411.

[0168] The expression cassettes may be transformed into a plant cell of interest by any of a number of methods. These methods include, for example, biolistic methods (See, for example, Sanford, *Trends In Biotech.* (1988) 6:299 302, U.S. Patent No. 4,945,050; electroporation (Fromm *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1985) 82:5824 5828); use of a laser beam, microinjection or any other method capable of introducing DNA into a chloroplast.

[0169] Additional descriptions of chloroplast expression vectors suitable for use in microorganisms such as microalgae are found in U.S. Patent Nos. 7,081,567 (issued July 25, 2006 to Xue *et al.*); 6,680,426 (issued January 20, 2004 to Daniell *et al.*); and 5,693,507 (issued December 2, 1997 to Daniell *et al.*).

[0170] Proteins expressed in the nuclear genome of *Chlorella* can be targeted to the chloroplast using chloroplast targeting signals. Chloroplast targeting sequences endogenous to *Chlorella* are known, such as genes in the *Chlorella* nuclear genome that encode proteins that are targeted to the chloroplast; see for example GenBank Accession numbers AY646197 and AF499684. Proteins can also be expressed in the *Chlorella* chloroplast by insertion of genes directly into the chloroplast genome. Chloroplast transformation typically occurs through homologous recombination, and can be performed if chloroplast genome sequences are known for creation of targeting vectors (see for example the complete genome sequence of a *Chlorella* chloroplast; Genbank accession number NC_001865). See previous sections herein for details of chloroplast transformation.

2. Expression in Mitochondria

[0171] In another preferred embodiment of the present invention, the expression of a polypeptide in a microorganism is targeted to mitochondria. Methods for targeting foreign gene products into mitochondria (Boutry *et al.*, *Nature* (London) (1987) 328:340-342) have been described, including in green microalgae (see for example *Mol Gen Genet.*, 1993 Jan;236(2-3):235-44).

[0172] For example, an expression vector encoding a suitable secretion signal can target a heterologous protein to the mitochondria. An exemplary expression vector for mitochondria targeting is the IMPACTVECTOR1.5 vector, from Wageningen UR-Plant Research International, which uses the yeast CoxIV secretion signal. This expression vector was shown to deliver proteins in the mitochondrial matrix. The protein is fused to the first 4 amino acids of the yeast CoxIV protein in order to allow proper processing of the signal peptide (Kohler *et al.*, *Plant J.* 11: 613-621 (1997)). Other mitochondrial targeting sequences are known, including those functional in green microalgae. For example, see *FEBS Lett.* 1990 Jan 29;260(2):165-8; and *J. Biol. Chem.* 2002 Feb 22;277(8):6051-8.

[0173] Proteins expressed in the nuclear genome of *Chlorella* can be targeted to the mitochondria using mitochondrial targeting signals. Details of mitochondrial protein targeting and transformation are provided herein.

3. Expression in Endoplasmic Reticulum

[0174] In another preferred embodiment of the present invention, the expression of a polypeptide in a microorganism is targeted to the endoplasmic reticulum (ER). The inclusion of an appropriate retention or sorting signal in an expression vector ensure that proteins are retained in the endoplasmic reticulum (ER) and do not go downstream into Golgi. For example, the IMPACTVECTOR1.3 vector, from Wageningen UR-Plant Research International, includes the well known KDEL retention or sorting signal. With this vector, ER retention has a practical advantage in that it has been reported to improve expression levels 5-fold or more. The main reason for this appears to be that the ER contains lower concentrations and/or different proteases responsible for post-translational degradation of expressed proteins than are present in the cytoplasm. ER retention signals functional in green microalgae are known. For example, see *Proc. Natl Acad. Sci. U S A.* 2005 Apr 26;102(17):6225-30.

G. Transformation

[0175] Cells can be transformed by any suitable technique including, *e.g.*, biolistics, electroporation (see Maruyama et al., (2004) *Biotechnology Techniques* 8:821-826), glass bead transformation and silicon carbide whisker transformation. Another method that can be used involves forming protoplasts and using CaCl₂ and polyethylene glycol (PEG) to introduce recombinant DNA into microalgal cells (see Kim et al., (2002), *Mar. Biotechnol.* 4:63-73, which reports the use of this method for the transformation of *Chlorella ellipsoidea*). Co-transformation of microalgae can be used to introduce two distinct vector molecules into a cell simultaneously (see for example, *Protist* (2004) Dec; 155(4):381-93).

[0176] Biolistic methods (see for example, Sanford, *Trends in Biotech.* (1988) 6:299-302, U.S. Patent No. 4,945,050; electroporation (Fromm et al., *PNAS* (1985) 82:5824-5828), use of a laser beam, microinjection, or any other method capable of introducing DNA into a microalgae can also be used for transformation.

H. Lipid Pathway Engineering

[0177] In certain embodiments of the present invention, it is preferred to further modify a microorganism, such as a microalgae, for example, to provide desired growth characteristics and/or to enhance the amount and/or quality of lipids produced. For example, microalgae can be engineered to increase carbon flux into the lipid pathway and/or modify the lipid pathway to beneficially alter the proportions or properties of lipid produced by the cells. The pathway is further, or alternatively, modified to alter the properties and/or proportions of various hydrocarbon molecules produced through enzymatic processing of lipids.

1. Alteration Of Properties Or Portions Of Lipids Or Hydrocarbons Produced

[0178] In some embodiments of the present invention, it can be desirable to alter characteristics, such as lipid yield per unit volume and/or per unit time, carbon chain length (*e.g.*, for biodiesel production or for industrial applications requiring hydrocarbon feedstock), reducing the number of double or triple bonds, optionally to zero, removing or eliminating rings and cyclic structures, and increasing the hydrogen:carbon ratio of a particular species of lipid or of a population of distinct lipid. In addition, microalgae that produce appropriate hydrocarbons can also be engineered to have even more desirable hydrocarbon outputs. Examples of such microalgae include species of the genus *Chlorella*.

a) Regulation of Enzymes Controlling Branch Points in Fatty Acid Synthesis

[0179] In particular embodiments of the present invention, one or more key enzymes that control branch points in metabolism to fatty acid synthesis is up-regulated or down-regulated to improve lipid production. Up-regulation is achieved, for example, by transforming cells with expression constructs in which a gene encoding the enzyme of interest is expressed, e.g., using a strong promoter and/or enhancer elements that increase transcription. Such expression constructs can include a selectable marker such that the transformants can be subjected to selection, which can result in amplification of the construct and an increase in the expression level of the encoded enzyme. Examples of enzymes suitable for up-regulation according to the methods of the invention include pyruvate dehydrogenase, which plays a role in converting pyruvate to acetyl-CoA (examples, some from microalgae, include GenBank accession numbers NP_415392; AAA53047; Q1XDM1; and CAF05587). Up-regulation of pyruvate dehydrogenase can increase production of acetyl-CoA, and thereby increase fatty acid synthesis. Acetyl-CoA carboxylase catalyzes the initial step in fatty acid synthesis. Accordingly, in certain embodiments of the present invention, this enzyme is up-regulated to increase production of fatty acids (examples, some from microalgae, include GenBank accession numbers BAA94752; AAA75528; AAA81471; YP_537052; YP_536879; NP_045833; and BAA57908). In another embodiment, fatty acid production is increased by up-regulation of acyl carrier protein (ACP), which carries the growing acyl chains during fatty acid synthesis (examples, some from microalgae, include GenBank accession numbers A0T0F8; P51280; NP_849041; YP_874433). Glycerol-3-phosphate acyltransferase catalyzes the rate-limiting step of fatty acid synthesis. Up-regulation of this enzyme is desired to increase fatty acid production (examples, some from microalgae, include GenBank accession numbers AAA74319; AAA33122; AAA37647; P44857; and ABO94442). In some embodiments, two or more of these polypeptides (pyruvate dehydrogenase, Acetyl-CoA carboxylase, acyl carrier protein (ACP), Glycerol-3-phosphate acyltransferase) are up-regulated. In that case, the two or more genes encoding the respective polypeptides may reside on a single expression construct or, alternatively, on two or more expression constructs. The preceding proteins are candidates for expression in microalgae, including species of the genus *Chlorella* and/or *Prototheca*.

[0180] Down-regulation of an enzyme of interest can be achieved using, e.g., antisense, catalytic RNA/DNA, RNA interference (RNAi), "knock-out," "knock-down," or other mutagenesis techniques. Enzyme expression/function can also be inhibited using intrabodies.

Examples of enzymes suitable for down-regulation according to the methods of the invention include citrate synthase, which consumes acetyl-CoA as part of the tricarboxylic acid (TCA) cycle. Down-regulation of citrate synthase can force more acetyl-CoA into the fatty acid synthetic pathway.

b) Modulation of Global Regulators of Fatty Acid Synthetic Genes

[0181] Global regulators modulate the expression of the genes of the fatty acid biosynthetic pathways. Accordingly, one or more global regulators of fatty acid synthesis can be up- or down-regulated, as appropriate, to inhibit or enhance, respectively, the expression of a plurality of fatty acid synthetic genes and, ultimately, to increase lipid production. Examples include sterol regulatory element binding proteins (SREBPs), such as SREBP-1a and SREBP-1c (for examples see GenBank accession numbers NP_035610 and Q9WTN3). Global regulators can be up- or down-regulated, for example, as described above with respect to regulation of control point enzymes.

c) Regulation of Hydrocarbon Modification Enzymes

[0182] The present application describes genetically engineering strains of microalgae, oleaginous yeast, bacteria, or fungi with one or more exogenous genes to produce various hydrocarbon compounds. Thus, in certain embodiments of the present invention, the methods of the invention also comprise transforming cells with one or more genes encoding hydrocarbon modification enzymes, such as, for example, a fatty acyl-ACP thioesterase (see examples in Table 7 with accession numbers), a fatty acyl-CoA/aldehyde reductase (see examples in Table 8 with accession numbers), a fatty acyl-CoA reductase (see examples in Table 9 with accession numbers), a fatty aldehyde decarbonylase (see examples in Table 10 with accession numbers), a fatty aldehyde reductase, or a squalene synthase gene (e.g., see GenBank Accession number AF205791). Stearoyl-ACP desaturase, for example, catalyzes the conversion of stearoyl-ACP to oleoyl-ACP. Up-regulation of this gene can increase the proportion of monounsaturated fatty acids produced by a cell; whereas down-regulation can reduce the proportion of monounsaturates. Similarly, the expression of one or more glycerolipid desaturases can be controlled to alter the ratio of unsaturated to saturated fatty acids such as ω -6 fatty acid desaturase, ω -3 fatty acid desaturase, or ω -6-oleate desaturase.

[0183] For example, microalgae that would naturally, or through genetic modification, produce high levels of lipids can be engineered (or further engineered) to express an

exogenous fatty acyl-ACP thioesterase, which can facilitate the cleavage of fatty acids from the acyl carrier protein (ACP) during lipid synthesis. These fatty acids can be recovered or, through further enzymatic processing within the cell, yield other hydrocarbon compounds. Optionally, the fatty acyl-ACP thioesterase can be expressed from a gene operably linked to an inducible promoter, and/or can be expressed in an intracellular compartment.

[0184] Thus, in a preferred embodiment of the present invention, the hydrocarbon modification enzyme suitable for use with the microorganisms and methods of the invention is a fatty acyl-ACP thioesterase. Fatty acyl-ACP thioesterases include, without limitation, those listed in Table 7, each of which is hereby incorporated by reference.

[0185] Table 7. Fatty acyl-ACP thioesterases and GenBank accession numbers.

<i>Umbellularia californica</i> fatty acyl-ACP thioesterase (GenBank #AAC49001)
<i>Cinnamomum camphora</i> fatty acyl-ACP thioesterase (GenBank #Q39473)
<i>Umbellularia californica</i> fatty acyl-ACP thioesterase (GenBank #Q41635)
<i>Myristica fragrans</i> fatty acyl-ACP thioesterase (GenBank #AAB71729)
<i>Myristica fragrans</i> fatty acyl-ACP thioesterase (GenBank #AAB71730)
<i>Elaeis guineensis</i> fatty acyl-ACP thioesterase (GenBank #ABD83939)
<i>Elaeis guineensis</i> fatty acyl-ACP thioesterase (GenBank #AAD42220)
<i>Populus tomentosa</i> fatty acyl-ACP thioesterase (GenBank #ABC47311)
<i>Arabidopsis thaliana</i> fatty acyl-ACP thioesterase (GenBank #NP_172327)
<i>Arabidopsis thaliana</i> fatty acyl-ACP thioesterase (GenBank #CAA85387)
<i>Arabidopsis thaliana</i> fatty acyl-ACP thioesterase (GenBank #CAA85388)
<i>Gossypium hirsutum</i> fatty acyl-ACP thioesterase (GenBank #Q9SQI3)
<i>Cuphea lanceolata</i> fatty acyl-ACP thioesterase (GenBank #CAA54060)
<i>Cuphea hookeriana</i> fatty acyl-ACP thioesterase (GenBank #AAC72882)
<i>Cuphea calophylla</i> subsp. <i>mesostemon</i> fatty acyl-ACP thioesterase (GenBank #ABB71581)
<i>Cuphea lanceolata</i> fatty acyl-ACP thioesterase (GenBank #CAC19933)
<i>Elaeis guineensis</i> fatty acyl-ACP thioesterase (GenBank #AAL15645)
<i>Cuphea hookeriana</i> fatty acyl-ACP thioesterase (GenBank #Q39513)
<i>Gossypium hirsutum</i> fatty acyl-ACP thioesterase (GenBank #AAD01982)
<i>Vitis vinifera</i> fatty acyl-ACP thioesterase (GenBank #CAN81819)
<i>Garcinia mangostana</i> fatty acyl-ACP thioesterase (GenBank #AAB51525)

Brassica juncea fatty acyl-ACP thioesterase (GenBank #ABI18986)
Madhuca longifolia fatty acyl-ACP thioesterase (GenBank #AAX51637)
Brassica napus fatty acyl-ACP thioesterase (GenBank #ABH11710)
Oryza sativa (indica cultivar-group) fatty acyl-ACP thioesterase (GenBank #EAY86877)
Oryza sativa (japonica cultivar-group) fatty acyl-ACP thioesterase (GenBank #NP_001068400)
Oryza sativa (indica cultivar-group) fatty acyl-ACP thioesterase (GenBank #EAY99617)
Cuphea hookeriana fatty acyl-ACP thioesterase (GenBank #AAC49269)

[0186] A fatty acyl-ACP thioesterase can be chosen based on its specificity for a growing (during fatty acid synthesis) fatty acid having a particular carbon chain length. For example, the fatty acyl-ACP thioesterase can have a specificity for a carbon chain length ranging from 8 to 34 carbon atoms, preferably from 8 to 18 carbon atoms, and more preferably from 10 to 14 carbon atoms. A specificity for a fatty acid with 12 carbon atoms is most preferred. A specificity for a fatty acid with 14 carbon atoms is also preferred.

[0187] In another preferred embodiment of the present invention, the hydrocarbon modification enzyme suitable for use with the microorganisms and methods of the invention is a fatty acyl-CoA/aldehyde reductase. Fatty acyl-CoA/aldehyde reductases include, without limitation, those listed in Table 8.

[0188] Table 8. Fatty acyl-CoA/aldehyde reductases listed by GenBank accession numbers.

AAC45217, YP_047869, BAB85476, YP_001086217, YP_580344, YP_001280274, YP_264583, YP_436109, YP_959769, ZP_01736962, ZP_01900335, ZP_01892096, ZP_01103974, ZP_01915077, YP_924106, YP_130411, ZP_01222731, YP_550815, YP_983712, YP_001019688, YP_524762, YP_856798, ZP_01115500, YP_001141848, NP_336047, NP_216059, YP_882409, YP_706156, YP_001136150, YP_952365, ZP_01221833, YP_130076, NP_567936, AAR88762, ABK28586, NP_197634, CAD30694, NP_001063962, BAD46254, NP_001030809, EAZ10132, EAZ43639, EAZ07989, NP_001062488, CAB88537, NP_001052541, CAH66597, CAE02214, CAH66590, CAB88538, EAZ39844, AAZ06658, CAA68190, CAA52019, and BAC84377

[0189] In another preferred embodiment of the present invention, the hydrocarbon modification enzyme suitable for use with the microorganisms and methods of the invention is a fatty acyl-CoA reductase. Fatty acyl-CoA reductases include, without limitation, those listed in Table 9.

[0190] Table 9. Fatty acyl-CoA reductases listed by GenBank accession numbers.

NP_187805, ABO14927, NP_001049083, CAN83375, NP_191229, EAZ42242, EAZ06453, CAD30696, BAD31814, NP_190040, AAD38039, CAD30692, CAN81280, NP_197642, NP_190041, AAL15288, and NP_190042
--

[0191] In another preferred embodiment of the present invention, the hydrocarbon modification enzyme suitable for use with the microorganisms and methods of the invention is a fatty aldehyde decarbonylase. Fatty aldehyde decarbonylases include, without limitation, those listed in Table 10.

[0192] Table 10. Fatty aldehyde decarbonylases listed by GenBank accession numbers.

NP_850932, ABN07985, CAN60676, AAC23640, CAA65199, AAC24373, CAE03390, ABD28319, NP_181306, EAZ31322, CAN63491, EAY94825, EAY86731, CAL55686, XP_001420263, EAZ23849, NP_200588, NP_001063227, CAN83072, AAR90847, and AAR97643

[0193] Additional examples of amino acid sequences for hydrocarbon modification enzymes or nucleic acids encoding them are described in U.S. Patent Nos. 6,610,527, 6,451,576, 6,429,014, 6,342,380, 6,265,639, 6,194,185, 6,114,160, 6,083,731, 6,043,072, 5,994,114, 5,891,697, 5,871,988, and 6,265,639, and further described in GenBank Accession numbers AAO18435, ZP_00513891, Q38710, AAK60613, AAK60610, AAK60611, NP_113747, CAB75874, AAK60612, AAF20201, BAA11024, AF205791, and CAA03710.

[0194] In particular embodiments, microorganisms of the present invention are genetically engineered to express one or more exogenous genes selected from a fatty acyl-ACP thioesterase, a fatty acyl-CoA/aldehyde reductase, a fatty acyl-CoA reductase, a fatty aldehyde reductase, or a fatty aldehyde decarbonylase. Suitable expression methods are described above with respect to the expression of a lipase gene, including, among other methods, inducible expression and compartmentalized expression.

[0195] Other suitable enzymes for use with the microorganisms and the methods of the invention include those that have at least 70% amino acid identity with one of the proteins listed in Tables 7-10, and that exhibit the corresponding desired enzymatic activity (e.g., cleavage of a fatty acid from an acyl carrier protein, reduction of an acyl-CoA to an aldehyde or an alcohol, or conversion of an aldehyde to an alkane). In additional embodiments, the enzymatic activity is present in a sequence that has at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% identity with one of the above described sequences, all of which are hereby incorporated by reference as if fully set forth.

[0196] The hydrocarbon modification enzymes described herein are useful in the production of various hydrocarbons from a microorganism (e.g., a microalgae, an oleaginous yeast, or a fungus) or population of microorganisms, whereby a fatty acyl-ACP thioesterase cleaves a fatty acid from an acyl carrier protein (ACP) during lipid synthesis. Through further enzymatic processing, the cleaved fatty acid is then combined with a coenzyme to yield an acyl-CoA molecule. This acyl-CoA is the substrate for the enzymatic activity of a fatty acyl-CoA reductase to yield an aldehyde, as well as for a fatty acyl-CoA/aldehyde reductase to yield an alcohol. The aldehyde produced by the action of the fatty acyl-CoA reductase identified above is the substrate for further enzymatic activity by either a fatty aldehyde reductase to yield an alcohol, or a fatty aldehyde decarbonylase to yield an alkane.

[0197] The hydrocarbon modification enzymes have a specificity for acting on a substrate which includes a specific number of carbon atoms. For example, a fatty acyl-ACP thioesterase may have a specificity for cleaving a fatty acid having 12 carbon atoms from the ACP. Therefore, in various embodiments, the microorganism can contain an exogenous gene that encodes a protein with specificity for catalyzing an enzymatic activity (e.g., cleavage of a fatty acid from an ACP, reduction of an acyl-CoA to an aldehyde or an alcohol, or conversion of an aldehyde to an alkane) with regard to the number of carbon atoms contained in the substrate. The enzymatic specificity can, in various embodiments, be for a substrate having from 8 to 34 carbon atoms, preferably from 8 to 18 carbon atoms, and more preferably from 10 to 14 carbon atoms. The most preferred specificity is for a substrate having 12 carbon atoms. In yet another embodiment, the preferred specificity is for a substrate having 14 carbon atoms. In other embodiments the specificity can be for 20 to 30 carbon atoms.

[0198] By selecting the desired combination of exogenous genes to be expressed, one can tailor the product generated by the microorganism, which may then be extracted from the aqueous biomass. For example, in certain embodiments, the microorganism contains: (i) an exogenous gene encoding a fatty acyl-ACP thioesterase; and, optionally, (ii) an exogenous gene encoding a fatty acyl-CoA/aldehyde reductase or a fatty acyl-CoA reductase; and, optionally, (iii) an exogenous gene encoding a fatty aldehyde reductase or a fatty aldehyde decarbonylase. The microorganism, when cultured as described herein, synthesizes a fatty acid linked to an ACP and the fatty acyl-ACP thioesterase catalyzes the cleavage of the fatty acid from the ACP to yield, through further enzymatic processing, a fatty acyl-CoA molecule. When present, the fatty acyl-CoA/aldehyde reductase catalyzes the reduction of the acyl-CoA to an alcohol. Similarly, the fatty acyl-CoA reductase, when present, catalyzes the reduction of the acyl CoA to an aldehyde. In those embodiments in which an exogenous gene encoding a fatty acyl-CoA reductase is present and expressed to yield an aldehyde product, a fatty aldehyde reductase, encoded by the third exogenous gene, catalyzes the reduction of the aldehyde to an alcohol. Similarly, a fatty aldehyde decarbonylase catalyzes the conversion of the aldehyde to an alkane, when present in the microorganism.

[0199] Genes encoding such enzymes can be obtained from cells already known to exhibit significant lipid production such as *Chlorella protothecoides*. Genes already known to have a role in lipid production, e.g., a gene encoding an enzyme that saturates double bonds, can be transformed individually into recipient cells. However, to practice the invention it is not necessary to make a priori assumptions as to which genes are required. A library of DNA containing different genes, such as cDNAs from a good lipid-production organism, can be transformed into recipient cells. The cDNA is preferably in operable linkage with a promoter active in microalgae. Different recipient microalgae cells transformed by a library receive different genes from the library. Transformants having improved lipid production are identified through screening methods known in the art, such as, for example, HPLC, gas chromatography, and mass spectrometry methods of hydrocarbon analysis (for examples of such analysis, see *Biomass and Bioenergy* Vol. 6. No. 4. pp. 269-274 (1994); *Experientia* 38; 47-49 (1982); and *Phytochemistry* 65 (2004) 3159-3165). These transformants are then subjected to further transformation with the original library and/or optionally interbred to generate a further round of organisms having improved lipid production. General procedures for evolving whole organisms to acquire a desired property are described in, e.g., US 6,716,631. Such methods entail, e.g., introducing a library of DNA fragments into a plurality

of cells, whereby at least one of the fragments undergoes recombination with a segment in the genome or an episome of the cells to produce modified cells. The modified cells are then screened for modified cells that have evolved toward acquisition of the desired function. Vectors and methods for transformation are analogous to those discussed in connection with expression of lipase genes.

[0200] Furthermore, subtractive libraries can be used to identify genes whose transcription is induced under different conditions, especially conditions employed in culturing microorganisms for biodiesel production, or for the production of hydrocarbons useful as a feedstock for industrial applications. Subtractive libraries contain nucleotide sequences reflecting the differences between two different samples. Such libraries are prepared by procedures that include the steps of denaturing and hybridizing populations of polynucleotides (e.g., mRNA, cDNA, amplified sequences) from each sample. Sequences common to both samples hybridize and are removed, leaving the sequences that differ between the samples. In this manner, sequences that are induced under particular conditions can be identified. This technique can be employed, for example, to identify genes useful for increasing lipid (e.g., fatty acid) production and, in particular, lipid production under any desired culture conditions. The subtractive hybridization technique can also be employed to identify promoters, e.g., inducible promoters, useful in expression constructs according to the invention.

[0201] Thus, for example, subtractive libraries can be prepared from microorganism cultures grown autotrophically (in the light without a fixed carbon source) or heterotrophically (in the dark in the presence of a fixed carbon source). In particular, heterotrophic genes may be induced during dark growth in the presence of a fixed carbon source and may therefore be present in a library generated by subtracting sequences from autotrophic cells from sequences from dark heterotrophic cells. Subtractive libraries can also be prepared from cultures to which a particular carbon substrate, such as glucose, has been added to identify genes that play a role in metabolizing the substrate. Subtractive libraries prepared from cultures grown in the presence of excess versus limited nitrogen can be used to identify genes that control cell division as opposed to hydrocarbon accumulation production. The preparation of a subtractive library from a culture to which lipids (e.g., fatty acids) have been added can help identify genes whose overexpression increases fatty acid production. More specifically, the addition of fatty acids to a culture of cells that can use the added fatty

acids will lead to the down-regulation of fatty acid synthetic genes to down-regulate fatty acid production. The overexpression of one or more such genes will have the opposite effect.

2. Increased Carbon Flux Into Lipid Pathway

[0202] Some microalgae produce significant quantities of non-lipid metabolites, such as, for example, polysaccharides. Because polysaccharide biosynthesis can use a significant proportion of the total metabolic energy available to cells, mutagenesis of lipid-producing cells followed by screening for reduced or eliminated polysaccharide production generates novel strains that are capable of producing higher yields of lipids.

[0203] The phenol: sulfuric acid assay detects carbohydrates (see Hellebust, *Handbook of Phycological Methods*, Cambridge University Press, 1978; and Cuesta G., *et al.*, *J. Microbiol. Methods*, 2003 Jan;52(1):69-73). The 1,6 dimethylmethylene blue assay detects anionic polysaccharides. (see for example *Braz. J. Med. Biol. Res.* 1999 May;32(5):545-50; *Clin. Chem.* 1986 Nov;32(11):2073-6).

[0204] Polysaccharides can also be analyzed through methods such as HPLC, size exclusion chromatography, and anion exchange chromatography (see for example Prosky L, Asp N, Schweizer TF, DeVries JW & Furda I (1988) Determination of insoluble, soluble and total dietary fiber in food and food products: Interlaboratory study. *Journal of the Association of Official Analytical Chemists* 71, 1017±1023; *Int J Biol Macromol.* 2003 Nov;33(1-3):9-18). Polysaccharides can also be detected using gel electrophoresis (see for example, *Anal Biochem.* 2003 Oct 15;321(2):174-82; *Anal Biochem.* 2002 Jan 1;300(1):53-68).

V. METHODS OF CULTURING MICROORGANISMS

A. Bioreactor

[0205] Microorganisms are cultured both for purposes of conducting genetic manipulations and for subsequent production of hydrocarbons (*e.g.*, lipids, fatty acids, aldehydes, alcohols, and alkanes). The former type of culture is conducted on a small scale and initially, at least, under conditions in which the starting microorganism can grow. For example, if the starting microorganism is a photoautotroph the initial culture is conducted in the presence of light. The culture conditions can be changed if the microorganism is evolved or engineered to grow independently of light. Culture for purposes of hydrocarbon production is preferentially conducted on a large scale (*e.g.*, 10,000L, 40,000L, 100,000L or larger bioreactors) in a bioreactor. Microorganisms (*e.g.*, microalgae) of the invention are typically cultured in the

methods of the invention in liquid media within a bioreactor, typically in the absence of light (heterotrophic growth).

[0206] The bioreactor or fermentor is used to culture microalgal cells through various phases of their physiological cycle. Bioreactors offer many advantages for use in the heterotrophic growth and propagation. To produce biomass for lipid production, microalgae are preferably grown in large quantities in liquid, such as in suspension cultures as an example. Bioreactors such as stainless steel fermentors can accommodate very large culture volumes (40,000 liter and greater capacity bioreactors are used in various embodiments of the invention). Bioreactors also typically allow for the control of culture conditions such as temperature, pH, oxygen tension, and carbon dioxide levels. For example, bioreactors are typically configurable, for example, using ports attached to tubing, to allow gaseous components, like oxygen or nitrogen, to be bubbled through a liquid culture. Other culture parameters, such as pH of the culture media, the identity and concentration of trace elements, and other media constituents can also be more readily manipulated using a bioreactor.

[0207] Bioreactors can be configured to flow culture media through the bioreactor throughout the time period during which the microalgae reproduce and increase in number. In some embodiments, for example, media can be infused into the bioreactor after inoculation but before the cells reach a desired density. In other instances, a bioreactor is filled with culture media at the beginning of a culture, and no more culture media is infused after the culture is inoculated. In other words, the microalgal biomass is cultured in an aqueous medium for a period of time during which the microalgae reproduce and increase in number; however, quantities of aqueous culture medium are not flowed through the bioreactor throughout the time period. Thus, in some embodiments, aqueous culture medium is not flowed through the bioreactor after inoculation.

[0208] Bioreactor ports can be used to introduce, or extract gases, solids, semisolids, and liquids, into the bioreactor chamber containing microalgae. While many bioreactors have more than one port (for example, one for media entry, and another for sampling), it is not necessary that only one substance enter or leave a port. For example, a port can be used to flow culture media or additional carbon source into the bioreactor and later be used for sampling, gas entry, gas exit, or other purposes. Preferably, a sampling port can be used repeatedly without altering or compromising the axenic nature of the culture. A sampling port can be configured with a valve or other device that allows the flow of sample to be

stopped or started or to provide a means of continuous sampling. Bioreactors typically have at least one port that allows inoculation of a culture, and such a port can also be used for other purposes such as media or gas entry.

[0209] Bioreactors allow the gas content of the culture of microorganism (*e.g.*, microalgae) to be manipulated. To illustrate, part of the volume of a bioreactor can be gas rather than liquid, and the gas inlets of the bioreactor allow pumping of gases into the bioreactor. Gases that can be beneficially pumped into a bioreactor include air, air/CO₂ mixtures, noble gases, such as argon, and other gases. Bioreactors are typically equipped to enable the user to control the rate of entry of a gas into the bioreactor. As noted above, increasing gas flow into a bioreactor can be used to increase mixing of the culture.

[0210] Increased gas flow affects the turbidity of the culture as well. Turbulence can be achieved by placing a gas entry port below the level of the aqueous culture media so that gas entering the bioreactor bubbles to the surface of the culture. One or more gas exit ports allow gas to escape, thereby preventing pressure buildup in the bioreactor. Preferably a gas exit port leads to a "one-way" valve that prevents contaminating microorganisms from entering the bioreactor.

B. Media

[0211] Culture media for the cultivation of microorganisms, including microalgae, typically contains components such as a fixed nitrogen source, a fixed carbon source, trace elements, optionally a buffer for pH maintenance, and phosphate (typically provided as a phosphate salt). Other components can include salts such as sodium chloride, particularly for seawater microalgae. Nitrogen sources include organic and inorganic nitrogen sources, including, for example, but without limitation, molecular nitrogen, nitrate, nitrate salts, ammonia (pure or in salt form, such as (NH₄)₂SO₄ and NH₄OH), protein (and amino acids), soybean meal, cornsteep liquor, and yeast extract. Examples of trace elements include zinc, boron, cobalt, copper, manganese, and molybdenum, in for example, the respective forms of ZnCl₂, H₃BO₃, CoCl₂·6H₂O, CuCl₂·2H₂O, MnCl₂·4H₂O and (NH₄)₆Mo₇O₂₄·4H₂O.

[0212] Solid and liquid growth media are generally available from a wide variety of sources, and instructions for the preparation of particular media that is suitable for a wide variety of strains of microorganisms can be found, for example, online at <http://www.utex.org/>, a site maintained by the University of Texas at Austin, 1 University Station A6700, Austin, Texas 78712-0183, for its culture collection of algae (UTEX). For

example, various fresh water and salt water media include those described in PCT Pub No. 2008/151149, incorporated herein by reference.

[0213] In a particular example, Proteose Medium is suitable for axenic cultures, and a 1 liter volume of Proteose Medium (pH ~6.8) can be prepared by the addition of 1g proteose peptone to 1 liter of Bristol Medium. Bristol Medium comprises 2.94 mM NaNO₃, 0.17 mM CaCl₂·2H₂O, 0.3 mM MgSO₄·7H₂O, 0.43 mM K₂HPO₄, 1.29 mM KH₂PO₄ and 1.43 mM NaCl in an aqueous solution. The solution is covered and autoclaved, and the stored at a refrigerated temperature prior to use. Another example is the *Prototheca* isolateion medium (PIM), which comprises 10g/L potassium hydrogen phthalate (KHP), 0.9g/L sodium hydroxide, 0.1g/L magnesium sulfate, 0.2g/L potassium hydrogen phosphate, 0.3g/L ammonium chloride, 10g/L glucose, 0.001g/L thiamine hydrochloride, 20g/L agar, 0.25g/L 5-fluorocytosine, at a pH in the range of 5.0 to 5.2 (see Pore (1973) *App. Microbiology*, 26:648-649). Other suitable media for use with the methods of the invention can be readily identified by consulting the URL identified above, or by consulting other organizations that maintain cultures of microorganisms, such as SAG, CCAP or CCALA. SAG refers to the Culture Collection of Algae at the University of Göttingen (Göttingen, Germany), CCAP refers to the culture collection of algae and protozoa managed by the Scottish Association for Marine Science (Scotland, United Kingdom), and CCLA refers to the culture collection of algal laboratory at the Institute of Botany (Třeboň, Czech Republic). Additionally, US Patent No. 5,900,370 describes media formulations and conditions suitable for heterotrophic fermentation of *Prototheca* species.

[0214] For oil (lipid) production, selection of fixed carbon source is important, as the cost of the fixed carbon source must be sufficiently low to make oil production economical. Thus, while suitable carbon sources include, for example, acetate, folidoside, fructose, galactose, glucuronic acid, glucose, glycerol, lactose, mannose, N-acetylglucosamine, rhamnase, sucrose, glucose, and/or xylose, selection of feedstocks containing these compounds is an important aspect of the methods of the invention. Some microorganism (*e.g.*, microalgae) species can grow by utilizing a fixed carbon source such as glucose or acetate in the absence of light. Such growth is known as heterotrophic growth. For *Chlorella protothecoides*, for example, heterotrophic growth results in high production of biomass and accumulation of high lipid content in cells. Other suitable feedstocks include, for eample, black liquor, corn starch, depolymerized cellulosic material, milk whey, molasses, thick cane juice, potato, sorghum, sucrose, sugar beet, sugar cane, rice and wheat. Carbon sources can also be

provided as a mixture, such as a mixture of sucrose and depolymerized sugar beet pulp. The one or more carbon source(s) can be supplied at a concentration of at least about 50 μ M, at least about 100 μ M, at least about 500 μ M, at least about 5mM, at least about 50 mM, at least about 500mM, of one or more exogenously provided fixed carbon source(s). Carbon sources of particular interest for purposes of the present invention include cellulose (in a depolymerized form), glycerol, sucrose (in the form of cane juice or molasses) and sorghum.

[0215] Some microorganisms naturally grow on or can be engineered to grow on a fixed carbon source that is a heterogeneous source of compounds such as municipal waste, secondarily treated sewage, wastewater, and other sources of fixed carbon and other nutrients such as sulfates, phosphates, and nitrates. The sewage component serves as a nutrient source in the production of hydrocarbons, and the culture provides an inexpensive source of hydrocarbons.

[0216] In some heterotrophic growth methods, microorganisms can be cultured using cellulosic biomass as a feedstock. Cellulosic biomass (e.g., stover, such as corn stover) is inexpensive and readily available; however, attempts to use this material as a feedstock for yeast have failed. In particular, such feedstock have been found to be inhibitory to yeast growth, and yeast cannot use the 5-carbon sugars produced from cellulosic materials (e.g., xylose from hemi-cellulose). By contrast, microalgae can grow on processed cellulosic material. Accordingly, heterotrophic growth methods include a method of culturing a microalgae in the presence of a cellulosic material and/or a 5-carbon sugar. Cellulosic materials generally include:

Component	Percent Dry Weight
Cellulose	40-60%
Hemicellulose	20-40%
Lignin	10-30%

[0217] Suitable cellulosic materials include residues from herbaceous and woody energy crops, as well as agricultural crops, i.e., the plant parts, primarily stalks and leaves, not removed from the fields with the primary food or fiber product. Examples include corn stover (stalks, leaves, husks, and cobs), wheat straw, and rice straw. Five-carbon sugars that are produced from such materials include xylose.

[0218] *Chlorella protothecoides*, for example, has been shown to exhibit higher levels of productivity when cultured on a combination of glucose and xylose than when cultured on either glucose or xylose alone. This synergistic effect provides a significant advantage in that it allows cultivation of *Chlorella* on combinations of xylose and glucose, such as cellulosic material.

[0219] In another embodiment of the methods of the invention, the carbon source is sucrose, including a complex feedstock containing sucrose, such as thick cane juice from sugar cane processing or molasses. In one embodiment, the culture medium further includes at least one sucrose utilization enzyme. In some cases, the culture medium includes a sucrose invertase. In one embodiment, the sucrose invertase enzyme is a secretable sucrose invertase enzyme encoded by an exogenous sucrose invertase gene expressed by the population of microorganisms. Thus, in some cases, the microalgae has been genetically engineered to express a sucrose utilization enzyme, such as a sucrose transporter, a sucrose invertase, a hexokinase, a glucokinase, or a fructokinase.

[0220] Complex feedstocks containing sucrose include waste molasses from sugar cane processing; the use of this low-value product of sugar cane processing can provide significant cost savings in the production of lipids/oil. Another complex feedstock containing sucrose that is useful in the the methods of the invention is sorghum, including sorghum syrup and pure sorghum. Sorghum syrup is produced from the juice of sweet sorghum cane; its sugar profile consists of mainly glucose (dextrose), fructose, and sucrose.

C. Increasing Yield Of Lipids

[0221] For the production of lipids/oil in accordance with the methods of the invention, it is preferable to culture cells in the dark, as is the case, for example, when using extremely large (40,000 liter or greater capacity) fermentors that do not allow light to strike the culture. As an example, an inoculum of lipid-producing microalgal cells are introduced into the medium; there is a lag period (lag phase) before the cells begin to propagate. Following the lag phase, the propagation rate increases steadily and enters the log, or exponential phase. The exponential phase is in turn followed by a slowing of propagation due to decreases in nutrients such as nitrogen, increases in toxic substances, and quorum sensing mechanisms. After this slowing, propagation stops, and cells enter a stationary phase or steady growth state, depending on the particular environment provided to the cells. For obtaining lipid-rich biomass, the culture is typically harvested well after the end of the exponential phase, which

70

may be terminated early by allowing nitrogen or another key nutrient (other than carbon) to become depleted, forcing the cells to convert the carbon sources, which are present in excess, to lipid. Culture parameters can be manipulated to optimize total oil production, the combination of lipid species produced, and/or production of a specific oil.

[0222] Process conditions can be adjusted to increase the yield of lipids suitable for use as biodiesel or other target molecules, and/or to reduce production cost. For example, in certain embodiments, a microbe (e.g., a microalgae) is cultured in the presence of a limiting concentration of one or more nutrients, such as, for example, nitrogen. This condition tends to increase microbial lipid yield over microbial lipid yield in a culture in which nitrogen is provided in excess. In particular embodiments, the increase in lipid yield is at least about: 10%, 20%, 30%, 40%, 50%, 75%, 100%, 200%, 300%, 400%, or 500%. The microbe can be cultured in the presence of a limiting amount of the nutrient for a portion of the total culture period or for the entire period. In particular embodiments, the nutrient concentration is cycled between a limiting concentration and a non-limiting concentration at least twice during the total culture period. In addition, as described above, certain fixed carbon feedstocks, such as glycerol, can be employed to increase the percentage of cell weight that is lipid, in relation to comparable quantities of other fixed carbon feedstocks.

[0223] To increase lipid yield, acetic acid can be employed in the feedstock for a lipid-producing microbe (e.g., a microalgae). Acetic acid feeds directly into the point of metabolism that initiates fatty acid synthesis (i.e., acetyl-CoA); thus providing acetic acid in the culture can increase fatty acid production. Generally, the microbe is cultured in the presence of a sufficient amount of acetic acid to increase microbial lipid yield, and/or microbial fatty acid yield, specifically, over microbial lipid (e.g., fatty acid) yield in the absence of acetic acid.

[0224] In another embodiment, lipid yield is increased by culturing a lipid-producing microbe (e.g., microalgae) in the presence of one or more cofactor(s) for a lipid pathway enzyme (e.g., a fatty acid synthetic enzyme). Generally, the concentration of the cofactor(s) is sufficient to increase microbial lipid (e.g., fatty acid) yield over microbial lipid yield in the absence of the cofactor(s). In a particular embodiment, the cofactor(s) are provided to the culture by including in the culture a microbe (e.g., microalgae) containing an exogenous gene encoding the cofactor(s). Alternatively, cofactor(s) may be provided to a culture by including a microbe (e.g., microalgae) containing an exogenous gene that encodes a protein that

participates in the synthesis of the cofactor. In certain embodiments, suitable cofactors include any vitamin required by a lipid pathway enzyme, such as, for example: biotin, pantothenate. Genes encoding cofactors suitable for use in the invention or that participate in the synthesis of such cofactors are well known and can be introduced into microbes (e.g., microalgae), using constructs and techniques such as those described above.

D. Microalgal Biomass with High Oil Content

[0225] Microalgal biomass with a high percentage of oil/lipid accumulation by dry cell weight has been generated using different methods of culture, which are known in the art. Microalgal biomass with a higher percentage of accumulated oil/lipid is useful in accordance with the present invention. Li et al. describe *Chlorella vulgaris* cultures with up to 56.6% lipid by dry cell weight (DCW) in stationary cultures grown under autotrophic conditions using high iron concentrations (Li et al., *Bioresource Technology* 99(11):4717-22 (2008)). Rodolfi et al., describe *Nanochloropsis sp.* and *Chaetoceros calcitrans* cultures with 60% lipid by DCW and 39.8% lipid by DCW, respectively, grown in a photobioreactor under nitrogen starvation conditions (Rodolfi et al., *Biotechnology & Bioengineering* (2008) [Jun18 Epub ahead of print]). Solovchenko et al., describe *Parietochloris incise* cultures with approximately 30% lipid accumulation (by DCW) when grown phototrophically and under low nitrogen conditions (Solovchenko et al., *Journal of Applied Phycology* 20:245-251 (2008)). *Chlorella protothecoides* can produce up to 55% lipid (DCW) grown under certain heterotrophic conditions with nitrogen starvation (Miao and Wu, *Bioresource Technology* 97:841-846 (2006)). Other *Chlorella* species including *Chlorella emersonii*, *Chlorella sorokiniana* and *Chlorella minutissima* have been described to have accumulated up to 63% oil (DCW) when grown in stirred tank bioreactors under low-nitrogen media conditions (Tilman et al., *Enzyme and Microbial Technology* 27:631-635 (2000)). Still higher percent lipid accumulation by dry cell weight has been reported, including 70% lipid (DCW) accumulation in *Dumaliella tertiolecta* cultures grown in increased NaCl conditions (Takagi et al., *Journal of Bioscience and Bioengineering* 101(3): 223-226 (2006)) and 75% lipid accumulation in *Botryococcus braunii* cultures (Banerjee et al., *Critical Reviews in Biotechnology* 22(3): 245-279 (2002)).

[0226] Microalgal biomass generated by the culture methods described herein and useful in accordance with the present invention comprises at least 10% microalgal oil by dry weight. In some embodiments, the microalgal biomass comprises at least 15%, at least 25%, at least

35%, at least 45%, at least 55%, or at least 60% microalgal oil by dry weight. In some embodiments, the microalgal biomass contains from 10-90% microalgal oil, from 25-75% microalgal oil, from 40-75% microalgal oil, or from 50-70% microalgal oil by dry weight.

E. Culturing Microorganisms Under Induced Conditions

[0227] As described herein, in certain embodiments of the present invention, microorganisms are cultured under induced conditions, i.e., by providing a stimulus. Generally, this is carried out as follows: The microorganism is cultured for a first period of time sufficient to increase the cell density. Then, the stimulus is provided and the microorganisms are cultured for a second period of time. During the second period of time the desired effect of the stimulus takes place, e.g., induction of an exogenous gene or increased lipid production. Culturing the microorganisms for the second period of time can be in the continued presence of the stimulus. Alternatively, the stimulus may not be provided or may only partially provided (i.e., for a limited time) during the culturing of the microorganism for the second period of time.

F. Storing Microorganisms Prior to Extraction of Lipid

[0228] In some cases it is desirable to store cultured microorganisms for a period of time prior to subjecting them to the extraction processes described below. In some methods of the invention, the microorganisms, produced via a culturing process as described herein are optionally stored for a period of time between termination of the culturing process and lysing the microorganism. In some cases, the microorganism is stored for at least one hour between termination of the culturing process and lysing the cultured microorganism. In other cases, the microorganism is stored for at least two hours, at least three hours, at least four hours, at least five hours, at least six hours, at least seven hours, at least eight hours, at least nine hours, at least ten hours, at least eleven hours, at least twelve hours, at least thirteen hours, at least fourteen hours, at least fifteen hours, at least sixteen hours, at least seventeen hours, at least eighteen hours, at least nineteen hours, at least twenty hours, at least twenty-one hours, at least twenty-two hours, at least twenty-three hours, or for at least twenty-four hours between termination of the culturing process and lysing the cultured microorganism. In some cases, the microorganism is stored for at least thirty-six hours between termination of the culturing process and lysing the cultured microorganism. In some cases, the microorganism is stored for at least forty-eight hours between termination of the culturing process and lysing the cultured microorganism, or for longer periods of time. In other cases, the microorganism

is stored for at least sixty or seventy-two hours, or longer, between termination of the culturing process and lysing the cultured microorganism.

[0229] Microorganisms prepared in a culture process are optionally stored at a temperature below 15 degrees Celsius between termination of the culturing process and lysing the cultured microorganism. In some cases the microorganisms are stored at a temperature below 14°C, below 13°C, below 12°C, below 11°C, below 10°C, below 9°C, below 8°C, below 7°C, below 6°C, below 5°C, below 4°C, below 3°C, below 2°C, or below 1°C between termination of the culturing process and lysing the cultured microorganism. In some cases, the microorganism is stored at a temperature above 30 degrees Celsius between termination of the culturing process and lysing the cultured microorganism. In some cases, the microorganism is stored at a temperature above 40 degrees Celsius between termination of the culturing process and lysing the cultured microorganism. In other cases, the microorganism is stored at a temperature above 31°C, above 32°C, above 33°C, above 34°C, above 35°C, above 36°C, above 37°C, above 38°C, above 39°C, above 41°C, above 42°C, above 43°C, above 44°C, above 45°C, above 46°C, above 47°C, above 48°C, above 49°C, or above 50°C between termination of the culturing process and lysing the cultured microorganism.

[0230] In some cases, storage or ageing of the cultured microorganisms is used to disrupt the cells to facilitate oil extraction. With storage, the cell structure may weaken sufficiently to cause the contents of the cells to begin leaking, or to permit passage of reagents or other materials into the cells to facilitate extraction of the lipid contents. Storage can be used in this context in combination with other lysing methods described below.

[0231] In some cases, the microorganism is subjected to agitation during storage. Agitation can be in addition to any combination of storage conditions set forth above. Agitation can be done on a shaker, vortexer, or the like. Agitation can also result from the shear forces present during high g force centrifugation. Alternatively, the microorganism is not agitated during storage. The microorganisms can be stored in a bioreactor or other culture vessel, or optionally transferred to a separate storage container.

VI. METHODS OF EXTRACTION OF LIPID FROM MICROORGANISM

[0232] In one aspect, the present invention is directed to a process for extracting, recovering, isolating or obtaining lipids from microorganisms. The process of the present invention is applicable to extracting a variety of lipids from a variety of microorganisms.

A. Lysing Cells

[0233] Intracellular lipids produced in microorganisms are extracted after lysing the cells of the microorganism. Once extracted, the lipids can be further refined to produce a high purity lipids.

[0234] After completion of culturing, the microorganisms can be separated from the fermentation broth, preferably without a drying step such as drum drying, spray drying, tray drying, vacuum drying and other steps that remove substantially all of the extracellular and intracellular water from the broth. Optionally, the separation is effected by centrifugation to generate a concentrated paste. Centrifugation does not remove significant amounts of intracellular water from the microorganisms and is not a drying step. The biomass can then be washed with a washing solution (e.g., DI water) to get rid of the fermentation broth and debris. Optionally, the washed microbial biomass may also be dried (oven dried, lyophilized, etc.) prior to cell disruption. Alternatively, cells can be lysed without separation from some or all of the fermentation broth when the fermentation is complete. For example, the cells can be at a ratio of less than 1:1 v:v cells to extracellular liquid when the cells are lysed.

[0235] Microorganism containing a lipid can be lysed to produce a lysate. As detailed herein, the step of lysing a microorganism (also referred to as cell lysis) can be achieved by any convenient means, including heat-induced lysis, adding a base, adding an acid, using enzymes such as proteases and polysaccharide degradation enzymes such as amylases, using ultrasound, mechanical lysis, using osmotic shock, infection with a lytic virus, and/or expression of one or more lytic genes. Lysis is performed to release intracellular molecules which have been produced by the microorganism. Each of these methods for lysing a microorganism can be used as a single method or in combination.

[0236] The extent of cell disruption can be observed by microscopic analysis. Using one or more of the methods described herein, typically more than 70% cell breakage is observed. Preferably, cell breakage is more than 80%, more preferably more than 90% and most preferred about 100%.

[0237] In particular embodiments, the microorganism is lysed after growth, for example to increase the exposure of cellular lipid to a catalyst for transesterification such as a lipase or a chemical catalyst, expressed as described above. The timing of lipase expression (e.g., via an inducible promoter), cell lysis, and the adjustment of transesterification reaction conditions

(e.g., removal of water, addition of alcohol, etc.) can be adjusted to optimize the yield of fatty acid esters from lipase-mediated transesterification. Below are described a number of lysis techniques. These techniques can be used individually or in combination.

1. Heat-Induced Lysis

[0238] In a preferred embodiment of the present invention, the step of lysing a microorganism comprises heating of a cellular suspension containing the microorganism. In this embodiment, the fermentation broth containing the microorganisms (or a suspension of microorganisms isolated from the fermentation broth) is heated until the microorganisms, i.e., the cell walls and membranes of microorganisms degrade or breakdown. Typically, temperatures applied are at least 50°C. Higher temperatures, such as, at least 60°C, at least 70°C, at least 80°C, at least 90°C, at least 100°C, at least 110°C, at least 120°C, at least 130°C or higher are used for more efficient cell lysis. Example 7 describes an embodiment of lysis using heat treatment.

[0239] Lysing cells by heat treatment can be performed by boiling the microorganism. Alternatively, heat treatment (without boiling) can be performed in an autoclave (see Example 6). The heat treated lysate may be cooled for further treatment.

[0240] Cell disruption can also be performed by steam treatment, i.e., through addition of pressurized steam. Steam treatment of microalgae for cell disruption is described, for example, in U.S. Patent No. 6,750,048.

2. Lysis Using A Base

[0241] In another preferred embodiment of the present invention, the step of lysing a microorganism comprises adding a base to a cellular suspension containing the microorganism.

[0242] The base should be strong enough to hydrolyze at least a portion of the proteinaceous compounds of the microorganisms used. Bases which are useful for solubilizing proteins are known in the art of chemistry. Exemplary bases which are useful in the methods of the present invention include, but are not limited to, hydroxides, carbonates and bicarbonates of lithium, sodium, potassium, calcium, and mixtures thereof. A preferred base is KOH. Examples 6 and 7 describe embodiments of cell lysis using KOH.

3. Acidic Lysis

[0243] In another preferred embodiment of the present invention, the step of lysing a microorganism comprises adding an acid to a cellular suspension containing the microorganism. Acid lysis can be effected using an acid at a concentration of 10-500 mM or preferably 40-160 mM. Acid lysis is preferably performed at above room temperature (e.g., at 40-160°, and preferably a temperature of 50-130°. For moderate temperatures (e.g., room temperature to 100°C and particularly room temperature to 65°, acid treatment can usefully be combined with sonication or other cell disruption methods. Example 7 describes embodiments of cell lysis using acidic lysis.

4. Lysing Cells Using Enzymes

[0244] In another preferred embodiment of the present invention, the step of lysing a microorganism comprises lysing the microorganism by using an enzyme. Preferred enzymes for lysing a microorganism are proteases and polysaccharide-degrading enzymes such as hemicellulase (e.g., hemicellulase from *Aspergillus niger*; Sigma Aldrich, St. Louis, MO; #H2125), pectinase (e.g., pectinase from *Rhizopus sp.*; Sigma Aldrich, St. Louis, MO; #P2401), Mannaway 4.0 L (Novozymes), cellulase (e.g., cellulase from *Trichoderma viride*; Sigma Aldrich, St. Louis, MO; #C9422), and driselase (e.g., driselase from *Basidiomycetes sp.*; Sigma Aldrich, St. Louis, MO; #D9515). Example 7 describes embodiments of cell lysis using enzymes.

a) Cellulases

[0245] In a preferred embodiment of the present invention, a cellulase for lysing a microorganism is a polysaccharide-degrading enzyme, optionally from *Chlorella* or a *Chlorella* virus. Example 7 describes embodiments of cell lysis using a cellulase. Another example of a polysaccharide-degrading enzyme that can be used is mannaway, as described in the Examples.

b) Proteases

[0246] Proteases such as *Streptomyces griseus* protease, chymotrypsin, proteinase K, proteases listed in Degradation of Polylactide by Commercial Proteases, Oda Yet al., Journal of Polymers and the Environment, Volume 8, Number 1, January 2000 , pp. 29-32(4), and other proteases can be used to lyse microorganisms. Other proteases that can be used include

Alcalase 2.4 FG (Novozymes) and Flavourzyme 100 L (Novozymes), as described in the Examples.

c) Combinations

[0247] Any combination of a protease and a polysaccharide-degrading enzyme can also be used, including any combination of the preceding proteases and polysaccharide-degrading enzymes.

5. Lysis Cells Using Ultrasound

[0248] In another preferred embodiment of the present invention, the step of lysing a microorganism is performed by using ultrasound, i.e., sonication. Thus, cells can also be lysed with high frequency sound. The sound can be produced electronically and transported through a metallic tip to an appropriately concentrated cellular suspension. This sonication (or ultrasonication) disrupts cellular integrity based on the creation of cavities in cell suspension. Example 6 describes a method for cell lysis using ultrasound. Example 7 describes embodiments of cell lysis using sonication.

6. Mechanical Lysis

[0249] In another preferred embodiment of the present invention, the step of lysing a microorganism is performed by mechanical lysis. Cells can be lysed mechanically and optionally homogenized to facilitate hydrocarbon (e.g., lipid) collection. For example, a pressure disrupter can be used to pump a cell containing slurry through a restricted orifice valve. High pressure (up to 1500 bar) is applied, followed by an instant expansion through an exiting nozzle. Cell disruption is accomplished by three different mechanisms: impingement on the valve, high liquid shear in the orifice, and sudden pressure drop upon discharge, causing an explosion of the cell. The method releases intracellular molecules.

[0250] Alternatively, a ball mill can be used. In a ball mill, cells are agitated in suspension with small abrasive particles, such as beads. Cells break because of shear forces, grinding between beads, and collisions with beads. The beads disrupt the cells to release cellular contents. Cells can also be disrupted by shear forces, such as with the use of blending (such as with a high speed or Waring blender as examples), the french press, or even centrifugation in case of weak cell walls, to disrupt cells.

7. Lysing Cells By Osmotic Shock (Cytolysis)

[0251] In another preferred embodiment of the present invention, the step of lysing a microorganism is performed by applying an osmotic shock. This can be achieved, for example, by centrifuging fermentation broth and resuspending the cell paste in deionized water.

8. Infection With A Lytic Virus

[0252] In a preferred embodiment of the present invention, the step of lysing a microorganism comprises infection of the microorganism with a lytic virus. A wide variety of viruses are known to lyse microorganisms suitable for use in the present invention, and the selection and use of a particular lytic virus for a particular microorganism is within the level of skill in the art.

[0253] For example, *paramecium bursaria chlorella* virus (PBCV-1) is the prototype of a group (family Phycodnaviridae, genus Chlorovirus) of large, icosahedral, plaque-forming, double-stranded DNA viruses that replicate in, and lyse, certain unicellular, eukaryotic chlorella-like green algae. Accordingly, any susceptible microalgae, such as *C. protothecoides*, can be lysed by infecting the culture with a suitable chlorella virus. Methods of infecting species of *Chlorella* with a chlorella virus are known. See for example *Adv. Virus Res.* 2006;66:293-336; *Virology*, 1999 Apr 25;257(1):15-23; *Virology*, 2004 Jan 5;318(1):214-23; *Nucleic Acids Symp. Ser.* 2000;(44):161-2; *J. Virol.* 2006 Mar;80(5):2437-44; and *Annu. Rev. Microbiol.* 1999;53:447-94.

9. Autolysis (Expression of a Lytic Gene)

[0254] In another preferred embodiment of the present invention, the step of lysing a microorganism comprises autolysis. In this embodiment, a microorganism according to the invention is genetically engineered to produce a lytic gene that will lyse the microorganism. This lytic gene can be expressed using an inducible promoter, so that the cells can first be grown to a desirable density in a fermentor and then harvested, followed by induction of the promoter to express the lytic gene to lyse the cells. In one embodiment, the lytic gene encodes a polysaccharide-degrading enzyme.

[0255] In certain other embodiments, the lytic gene is a gene from a lytic virus. Thus, for example, a lytic gene from a *Chlorella* virus can be expressed in a *Chlorella*, such as *C. protothecoides*.

[0256] Suitable expression methods are described herein with respect to the expression of a lipase gene. Expression of lytic genes is preferably done using an inducible promoter, such as a promoter active in microalgae that is induced by a stimulus such as the presence of a small molecule, light, heat, and other stimuli. Lytic genes from chlorella viruses are known. For example, see *Virology* 260, 308-315 (1999); *FEMS Microbiology Letters* 180 (1999) 45-53; *Virology* 263, 376-387 (1999); and *Virology* 230, 361-368 (1997).

10. Pressure Oscillation

[0257] In another preferred embodiment of the present invention, the step of lysing a microorganism comprises subjecting the microorganism to rapid increases and decreases in pressure. Such rapid increases and decreases are preferably performed across a wide enough differential in pressure that the cells are not able to self-regulate, and burst as a result.

11. Additional Consideration

[0258] When lipids are not extracted immediately after isolating the microorganism, the isolated microorganisms are typically dried. Drying a microorganism can be done, e.g., on a drum dryer. Dried microorganisms can be packaged in vacuum-sealed containers to prevent degradation of lipids.

B. Treatment Of Cell Lysate With Organic Solvent

[0259] A lipid produced by a microorganism is also referred herein as a "first lipid." A first lipid may comprise one or more lipids produced by the microorganism.

[0260] A lipid used to extract a first lipid from a microorganism is referred to as a "second lipid." A preferred second lipid is an oil. An organic solvent of the present invention comprises a second lipid.

[0261] Some methods of the present invention comprise the step of treating a lysate with an organic solvent. Typically, the organic solvent is added directly to the lysate without prior separation of the lysate components. After addition of the solvent, the lysate separates either of its own accord or as a result of centrifugation or the like into different layers. The layers can include in order of decreasing density: a pellet of heavy solids, an aqueous phase, an emulsion phase, and an oil phase. The emulsion phase is an emulsion of lipids and aqueous phase. Depending on the percentage of organic solvent added with respect to the lysate (w/w or v/v), the force of centrifugation if any, volume of aqueous media and other factors, either or both of the emulsion and oil phases can be present.

[0262] Incubation or treatment of the cell lysate or the emulsion phase with the organic solvent is performed for a time sufficient to allow the lipid produced by the microorganism to become solubilized in the organic solvent to form a heterogeneous mixture.

1. Oils

[0263] In a preferred embodiment of the present invention, an organic solvent is an oil selected from the group consisting of oil from soy, rapeseed, canola, palm, palm kernel, coconut, corn, waste vegetable oil, Chinese tallow, olive, sunflower, cotton seed, chicken fat, beef tallow, porcine tallow, microalgae, macroalgae, *Cuphea*, flax, peanut, choice white grease (lard), *Camelina sativa* mustard seed, cashew nut, oats, lupine, kenaf, calendula, hemp, coffee, linseed, hazelnut, euphorbia, pumpkin seed, coriander, camellia, sesame, safflower, rice, tung oil tree, cocoa, copra, pium poppy, castor beans, pecan, jojoba, jatropha, macadamia, Brazil nuts, and avocado. Also included are fossil oils such as crude oil or a distillate fraction of a fossil oil.

[0264] The amount of organic solvent added to the lysate is typically greater than 5% (measured by v/v and/or w/w) of the the lysate with which the solvent is being combined. Thus, a preferred v/v or w/w of the organic solvent is greater than 5%, at least 6%, at least 7%, at least 10%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, and at least 95% of the the cell lysate.

2. Other Organic Solvents

[0265] Other non-limiting examples of organic solvents which can be used to practice the methods of the present invention include hexane, isohexane, methanol, dodecane, fossil-derived crude oil and distillate fractions thereof, and supercritical carbon dioxide. Methanol can be an advantageous organic solvent to use in biodiesel manufacturing because it can be used in the transesterification process as well as extraction of oil.

C. Solventless Extraction

[0266] Lipids can also be extracted from a lysate without substantial or any use of organic solvents by cooling the lysate. In such methods, the lysate is preferably produced by acid treatment in combination with above room temperature. Sonication can also be used, particularly if the temperature is between room temperature and 65°C. Such a lysate on centrifugation or settling can be separated into layers, one of which is an aqueous:lipid layer. Other layers can include a solid pellet, an aqueous layer, and a lipid layer. Lipid can be

extracted from the emulsion layer by freeze thawing or otherwise cooling the emulsion as described further below. In such methods, it is not necessary to add any organic solvent. If any organic solvent is added, the organic solvent can be below 5% v/v or w/w of the lysate.

D. Agitation or No Agitation of The Heterogeneous Mixture

[0267] In a preferred embodiment of the present invention, after producing a cell lysate and optionally adding the organic solvent to the cell lysate or to the emulsion phase as described above, the heterogeneous mixture is agitated. Agitation can be done on a shaker, vortexer, or the like. Agitation can also result from the shear forces present during high g force centrifugation. Alternatively, the heterogeneous mixture is not agitated.

E. Separation Of The Heterogeneous Mixture Into A Lipid:Organic Solvent Composition And An Aqueous Composition And, Optionally An Emulsified Composition Or Cell Pellet Composition

[0268] The cell lysate, optionally treated with an organic solvent, produced by the above methods is a heterogeneous mixture including lipids, aqueous solutions, cell debris and the organic solvent (if added). The heterogeneous mixture can be separated into multiple layers as described below. The multiple layers can include in order of descending density, a pellet of cell debris, an aqueous layer, a lipid:aqueous emulsion layer and a lipid layer. The presence and relative proportions of the different layers depends on the lysis technique, the separation technique including centrifugal force, concentration of lipid in the microorganism and whether an organic solvent is used. The desired lipid can occur both in a separate lipid layer and in a lipid: aqueous layer. The use of an organic solvent favors formation of a separate lipid layer. The higher the concentration of the organic solvent relative to the lysis, the more lipid is likely to be found in a separate lipid layer rather than in an aqueous:lipid emulsion layer. The organic solvent, if present, associates with the lipid either in the lipid layer or the emulsion or both. The lipid can be obtained both from a separate lipid layer and from an emulsion layer, as described below.

[0269] Cell lysates can be separated into different layers as described above by techniques, such as, centrifugation and settling, i.e., allowing the layers to form spontaneously with time.

F. Separation of Lipid

[0270] The separation of lipid from other components present in the lysate depends on which layers are formed as a result of the separation. If a separate lipid layer is formed, this layer forms on top of other layers and can be suctioned off, pipetted or decanted or the like

from the top of the vessel containing the separated layers. This lipid layer can then be used in various applications described below. Likewise solid and/or aqueous layers can be drained off and discarded from the bottom of the vessel. If an emulsion layer is formed, the emulsion can be subjected to further extraction to separate lipids in the emulsion from aqueous fluid. In such a circumstance, separate aqueous and/or pellet layers, if present, can be drained off before separating the emulsion. A separate lipid layer if present can also be decanted off; however, such is not necessary, and the separate lipid layer can facilitate separation of the emulsion into its components.

[0271] The emulsion can be separated into its components by washing with a further volume of organic solvent to extract lipids from the emulsion. Alternatively, the emulsion can be washed with an aqueous solution to remove aqueous fluids from the emulsion. Alternatively, the emulsion can be subject to cooling below room temperature. Preferably, the emulsion is frozen and rethawed. Freezing is preferably to a temperature of -5°C to -30°C, and is maintained for at least one hour.

[0272] After any of the above treatments, the emulsion is re-separated into component layers by the same techniques as previously described. The component layers can include a cell pellet, aqueous layer, emulsion layer and lipid layer as previously described. The component layers can be separated as previously described. If an emulsion layer is present, the emulsion layer can be subjected to a further round of extraction using the same procedures described above.

[0273] After one or more rounds of separation, the lysate eventually yields a lipid layer. The lipid layer can be used in various applications as described below. When any of the rounds of extraction is performed with an organic solvent, that organic solvent remains associated with the lipid layer. Typically, it is not necessary to separate the lipid layer from the organic solvent. For example, when the organic solvent is an oil or lipid, the lipid present in the microorganisms can be referred to as a first lipid and the lipid or oil used for extraction can be referred to as a second lipid. For example, the first lipid can be triacylglycerol from a microalgae and the second lipid can be soy oil. After separation of a lipid:organic solvent layer, (in this case microalgal lipid:soy oil) the mixture can for example be transesterified or hydrotreated to yield biodiesel or renewable diesel, respectively.

[0274] One method of separation that can be used on a lysate, including a lysate that contains an emulsion, is the application of shear forces. Following cell lysis, oil may be

present as an oil-in-water emulsion with a small (<5 or <10 micron) droplet size. The emulsion may be stabilized by any number of ampipathic emulsifiers (e.g. oleosins, denatured proteins, phospholipids, fatty acids etc.). All else being equal, the stability of an emulsion is inversely related to the droplet size of the included phase. There are a number of procedures provided herein that may be applied to cause oil droplets to coalesce and increase in size, thus rendering the emulsion more amenable to breakage (e.g. gentle agitation in the presence of certain detergents, or acid or base). Once an oil-in-water emulsion has been treated so as to weaken the emulsion (by an increase in droplet size, as described above, and/or by some chemical or physical treatment that decreases the efficacy of the emulsifying agent(s)), the droplets need to be physically forced together in order that they may coalesce and phase separation may occur.

[0275] One method of forcing the included oil droplets together is by centrifugation that causes physical crowding by forcing the emulsion through a thin film and introducing concomitant shear that may be required to break the emulsion. A device that can be used in the methods provided herein to introduce g-force, thin film and shear is a stacked disk centrifuge in which the light phase is ejected as a thin film between rapidly rotating disks. The film is preferably between 1 micrometer and 1000 micrometers. The film is preferably less than or similar to the diameter of the included oil droplets.

[0276] In one embodiment, microbial oil-bearing cells in fermentation broth are centrifuged to reduce the water content of the composition. Cells are ruptured by homegenization and/or chemical / enzymatic treatment. Optionally, the oil droplets are forced to coalesce by gentle agitation of the lysate. Phase separation of the composition is then achieved by stacked disk centrifugation.

[0277] Another method of separation that can be used alone or in combination with other methods described herein is the use of one or more surfactants to destabilize an emulsion. Destabilization results in coalescence of oil globules into larger clusters and eventually into a phase separated composition of a light layer of oil, and a heavier aqueous phase. Preferred surfactants for destabilizing emulsions of oil-bearing microbial biomass is oleamide DEA (diethanolamine) (Stepan Chemical Co.) including but not limited to NINOL 201, laurelamide DEA including but not limited to NINOL 96L and NINOL 55L and cocoamide DEA including but not limited to NINOL 40CO. Methods of oil recovery from a lipid emulsion using surfactants are provided in the Examples below.

[0278] The lipid composition obtained can be analyzed by a number of methods, including HPLC as described in Example 6.

VII. METHOD OF PRODUCING FUELS SUITABLE FOR USE IN DIESEL VEHICLES AND JET ENGINES

[0279] Increasing interest is directed to the use of hydrocarbon components of biological origin in fuels, such as biodiesel, renewable diesel, and jet fuel, since renewable biological starting materials that may replace fossil ones are available, and the use thereof is desirable. There is an urgent need for methods for producing hydrocarbon components from biological materials. The present invention fulfills this need by providing methods for production of biodiesel, renewable diesel, and jet fuel using the lipid:organic solvent composition or the lipids described herein as a biological material to produce biodiesel, renewable diesel, and jet fuel.

[0280] After extraction, the present invention provides the advantage that the microbial oil and the organic solvent (preferably a plant oil) can together be subjected to chemical treatment to manufacture a fuel for use in diesel vehicles and jet engines. The ability to avoid separation of the organic solvent used for extraction from the microbial oil provides a significant advantage over traditional methods of oil extraction such as hexane extraction in which hexane must be distilled away from the microbial oil prior to any further processing steps.

[0281] Traditional diesel fuels are petroleum distillates rich in paraffinic hydrocarbons. They have boiling ranges as broad as 370° to 780°F, which are suitable for combustion in a compression ignition engine, such as a diesel engine vehicle. The American Society of Testing and Materials (ASTM) establishes the grade of diesel according to the boiling range, along with allowable ranges of other fuel properties, such as cetane number, cloud point, flash point, viscosity, aniline point, sulfur content, water content, ash content, copper strip corrosion, and carbon residue. Technically, any hydrocarbon distillate material derived from biomass that meets the appropriate ASTM specification can be defined as diesel, or as biodiesel.

[0282] Diesel fuel can be produced from biomass via several types of technologies. Feedstocks for diesel fuels derived from biomass include, but are not limited to, soybean, rape seed, canola, palm, and waste cooking oils, along with animal fats. Starting oils can also

be of algal origin. The lipid:organic solvent layer produced by the method of the present invention can serve as feedstock to produce biodiesel and renewable diesel.

A. Biodiesel

[0283] Biodiesel is a liquid which varies in color - between golden and dark brown - depending on the production feedstock. It is practically immiscible with water, has a high boiling point and low vapor pressure. Biodiesel refers to a diesel-equivalent processed fuel for use in diesel-engine vehicles. Biodiesel is biodegradable and non-toxic. An additional benefit of biodiesel over conventional diesel fuel is lower engine wear.

[0284] Typically, biodiesel comprises short chain alkyl esters. Various processes convert biomass or a lipid produced and isolated as described herein to diesel fuels. A preferred method to produce biodiesel is by transesterification of a lipid as described herein. A preferred short chain alkyl ester for use as biodiesel is a methyl ester or ethyl ester.

[0285] Biodiesel produced by a method described herein can be used alone or blended with conventional diesel fuel at any concentration in most modern diesel-engine vehicles. When blended with conventional diesel fuel (petroleum diesel), biodiesel may be present from about 0.1% to about 99.9%. Much of the world uses a system known as the "B" factor to state the amount of biodiesel in any fuel mix. For example, fuel containing 20% biodiesel is labeled B20. Pure biodiesel is referred to as B100.

[0286] Biodiesel can also be used as a heating fuel in domestic and commercial boilers. Existing oil boilers may contain rubber parts and may require conversion to run on biodiesel. The conversion process is usually relatively simple, involving the exchange of rubber parts for synthetic parts due to biodiesel being a strong solvent. Due to its strong solvent power, burning biodiesel will increase the efficiency of boilers.

[0287] Biodiesel can be used as an additive in formulations of diesel to increase the lubricity of pure Ultra-Low Sulfur Diesel (ULSD) fuel, which is advantageous because it has virtually no sulfur content.

[0288] Biodiesel is a better solvent than petrodiesel and can be used to break down deposits of residues in the fuel lines of vehicles that have previously been run on petrodiesel.

1. Production Of Biodiesel

[0289] Biodiesel can be produced by transesterification of triglycerides contained in oil-rich biomass and animal fats. The lipid:organic solvent layer or lipids produced by the method of the present invention can serve as feedstock to produce biodiesel. Thus, in another aspect of the present invention a method for producing biodiesel is provided. In a preferred embodiment, the method for producing biodiesel comprises the steps of (a) lysing a lipid-containing microorganism to produce a lysate; (b) treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent, wherein by the organic solvent-treated lysate forms a heterogeneous mixture; (c) separating the heterogeneous mixture into layers comprising a lipid:organic solvent layer and an aqueous layer; (d) removing the lipid:organic solvent composition from the aqueous composition, emulsion composition, or cell pellet composition; and (e) transesterifying the lipid:organic solvent composition, whereby biodiesel is produced.. The lipid:organic solvent composition comprises the organic solvent and the microbial lipids.

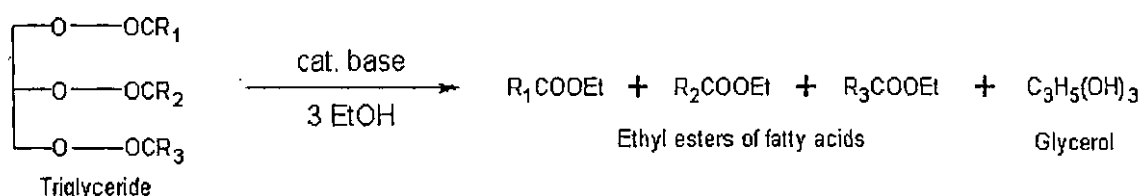
[0290] Methods for growth of a microorganism, lysing a microorganism to produce a lysate, treating the lysate in a medium comprising an organic solvent to form a heterogeneous mixture and separating the treated lysate into a lipid:organic solvent composition and an aqueous or emulsified composition have been described above and can also be used in the method of producing biodiesel.

[0291] Lipids and lipid:organic solvent composition, where the organic solvent is a triacylglyceride selected as described above can be subjected to transesterification to yield long-chain fatty acid esters useful as biodiesel. Preferred transesterification reactions are outlined below and include base catalyzed transesterification and transesterification using recombinant lipases.

[0292] In a base-catalyzed transesterification process, the triacylglycerides are reacted with an alcohol, such as methanol or ethanol, in the presence of an alkaline catalyst, typically potassium hydroxide. This reaction forms methyl or ethyl esters and glycerin (glycerol) as a byproduct.

a) General Chemical Process

[0293] Animal and plant oils are typically made of triglycerides which are esters of free fatty acids with the trihydric alcohol, glycerol. In transesterification, the glycerol in a triacylglyceride (TAG) is replaced with a short-chain alcohol such as methanol or ethanol. A typical reaction scheme is as follows:



[0294] In this scheme, the alcohol is deprotonated with a base to make it a stronger nucleophile. Commonly, ethanol or methanol is used in vast excess (up to 50-fold). Normally, this reaction will proceed either exceedingly slowly or not at all. Heat, as well as an acid or base can be used to help the reaction proceed more quickly. The acid or base are not consumed by the transesterification reaction, thus they are not reactants but catalysts. Almost all biodiesel has been produced using the base-catalyzed technique as it requires only low temperatures and pressures and produces over 98% conversion yield (provided the starting oil is low in moisture and free fatty acids).

[0295] Any free fatty acids in the base oil are either converted to soap and removed from the process, or they are esterified (yielding more biodiesel) using an acidic catalyst.

[0296] The most common form of transesterification uses methanol to produce methyl esters as it is the cheapest alcohol available. Ethanol is used to produce ethyl ester biodiesel. Higher alcohols, such as isopropanol and butanol can also be used.

[0297] A byproduct of the transesterification process is the production of glycerol. Approximately for every ton of biodiesel produced, 100 kg of glycerol are produced. This glycerol may be used as a chemical building block, and may also be used as a carbon source to ferment microorganisms.

b) Using Recombinant Lipases

[0298] Transesterification has also been carried out using an enzyme, such as a lipase instead of a base. Lipase-catalyzed transesterification can be carried out, for example, at a temperature between the room temperature and 80°C, and a mole ratio of the TAG to the

lower alcohol of greater than 1:1, preferably 2:1, more preferably 3:1, and most preferably about 3:1.

[0299] Lipases suitable for use in transesterification include, but are not limited to, those listed in Table 11 below. Other examples of lipases useful for transesterification are found in, e.g. U.S. Patent Nos. 4,798,793; 4,940,845 5,156,963; 5,342,768; 5,776,741 and WO89/01032.

[0300] Table 11. Lipases suitable for use in transesterification.

Aspergillus niger lipase ABG73614, *Candida antarctica* lipase B (novozym-435) CAA83122, *Candida cylindracea* lipase AAR24090, *Candida lipolytica* lipase (Lipase L; Amano Pharmaceutical Co., Ltd.), *Candida rugosa* lipase (e.g., Lipase-OF; Meito Sangyo Co., Ltd.), *Mucor miehei* lipase (Lipozyme IM 20), *Pseudomonas fluorescens* lipase AAA25882, *Rhizopus japonicas* lipase (Lilipase A-10FG) Q7M4U7_1, *Rhizomucor miehei* lipase B34959, *Rhizopus oryzae* lipase (Lipase F) AAF32408, *Serratia marcescens* lipase (SM Enzyme) ABI13521, *Thermomyces lanuginosa* lipase CAB58509, Lipase P (Nagase ChemteX Corporation), and Lipase QLM (Meito Sangyo Co., Ltd., Nagoya, Japan)

[0301] One challenge to using a lipase for the production of fatty acid esters suitable for biodiesel is that the price of lipase is much higher than the price of sodium hydroxide (NaOH) used by the strong base process. This challenge has been addressed by using an immobilized lipase, which can be recycled. However, the activity of the immobilized lipase must be maintained after being recycled for a minimum number of cycles to allow a lipase-based process to compete with the strong base process in terms of the production cost. Immobilized lipases are subject to poisoning by the lower alcohols typically used in transesterification. U.S. Patent No. 6,398,707 (issued June 4, 2002 to Wu et al.) describes methods for enhancing the activity of immobilized lipases and regenerating immobilized lipases having reduced activity.

2. Standards

[0302] The common international standard for biodiesel is EN 14214. ASTM D6751 is the most common biodiesel standard referenced in the United States and Canada. Germany uses DIN EN 14214 and the UK requires compliance with BS EN 14214.

[0303] Basic industrial tests to determine whether the products conform to these standards typically include gas chromatography, HPLC, and others. Biodiesel meeting the quality standards is very non-toxic, with a toxicity rating (LD₅₀) of greater than 50 mL/kg.

B. Renewable Diesel

[0304] Renewable diesel comprises a mixture of alkanes, such as C10:0, C12:0, C14:0, C16:0 and C18:0 and thus, are distinguishable from biodiesel. High quality renewable diesel conforms to the ASTM D 975 standard.

[0305] The lipid:organic solvent layer or lipids produced by the method of the present invention can serve as feedstock to produce renewable diesel. Thus, in another aspect of the present invention, a method for producing renewable diesel is provided. Renewable diesel can be produced by at least three processes, hydrothermal processing (hydrotreating), hydroprocessing, and indirect liquefaction. These processes yield non-ester distillates. During these processes, triacylglycerides produced and isolated as described herein, are broken down to alkenes of C16 and C18.

[0306] Thus, in another aspect of the present invention a method for producing renewable diesel is provided. In a preferred embodiment, the method for producing renewable diesel comprises the steps of A method of producing renewable diesel comprising the steps of: (a) lysing a lipid-containing microorganism to produce a lysate; (b) treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent to form a heterogeneous mixture; (c) separating the heterogeneous mixture into a lipid:organic solvent composition and an aqueous composition and, optionally, an emulsified composition or cell pellet composition; (d) removing the lipid:organic solvent composition from the aqueous composition, emulsion composition, or cell pellet composition; and (e) treating the lipid:organic to produce a straight chain alkane, whereby renewable diesel is produced. The lipid:organic solvent composition comprises the organic solvent and the cellular lipids.

1. Hydrotreating

[0307] In a preferred embodiment of the method for producing renewable diesel, treating the lipid:organic solvent composition or the lipids produced and isolated to produce a straight chain alkane, is performed by hydrotreating of the lipid:organic solvent composition. In hydrothermal processing, typically, biomass is reacted in water at an elevated temperature

and pressure to form oils and residual solids. Conversion temperatures are typically 570° to 660°F, with pressure sufficient to keep the water primarily as a liquid, 100 to 170 standard atmosphere (atm). Reaction times are on the order of 15 to 30 minutes. After the reaction is completed, the organics are separated from the water. Thereby a distillate suitable for diesel is produced.

2. Hydroprocessing

[0308] A renewable diesel, referred to as "green diesel" can be produced from fatty acids by traditional hydroprocessing technology. The triglyceride-containing oils can be hydroprocessed either as co-feed with petroleum or as a dedicated feed. The product is a premium diesel fuel containing no sulfur and having a cetane number of 90-100. Thus, in another preferred embodiment of the method for producing renewable diesel, treating the lipid:organic solvent composition or the lipids produced and isolated to produce a straight chain alkane, is performed by hydroprocessing of the lipid:organic solvent composition.

[0309] Petroleum refiners use hydroprocessing to remove impurities by treating feeds with hydrogen. Hydroprocessing conversion temperatures are typically 600° to 700°F. Pressures are typically 40 to 100 atm. The reaction times are on the order of 10 to 60 minutes.

[0310] Solid catalysts are employed to increase certain reaction rates, improve selectivity for certain products, and optimize hydrogen consumption.

[0311] Hydrotreating and hydroprocessing can ultimately lead to a reduction in the molecular weight of the feed. In the case of triglyceride-containing oils, the triglyceride molecule is reduced to four hydrocarbon molecules under hydroprocessing conditions: a propane molecule and three hydrocarbon molecules, typically in the C12 to C18 range. In some methods, the first step of treating a triglyceride is hydroprocessing to saturate double bonds, followed by deoxygenation at elevated temperature in the presence of hydrogen and a catalyst. In some methods, hydrogenation and deoxygenation occur in the same reaction. In other methods deoxygenation occurs before hydrogenation. Isomerization is then optionally performed, also in the presence of hydrogen and a catalyst. Finally, gases and naphtha components can be removed if desired. For example, see U.S. Patents 5,475,160 (hydrogenation of triglycerides); 5,091,116 (deoxygenation, hydrogenation and gas removal); 6,391,815 (hydrogenation); and 5,888,947 (isomerization).

3. Indirect Liquefaction

[0312] A traditional ultra-low sulfur diesel can be produced from any form of biomass by a two-step process. First, the biomass is converted to a syngas, a gaseous mixture rich in hydrogen and carbon monoxide. Then, the syngas is catalytically converted to liquids. Typically, the production of liquids is accomplished using Fischer-Tropsch (FT) synthesis. This technology applies to coal, natural gas, and heavy oils. Thus, in yet another preferred embodiment of the method for producing renewable diesel, treating the lipid:organic solvent composition or the lipids produced and isolated to produce a straight chain alkane, is performed by indirect liquefaction of the lipid:organic solvent composition.

C. Jet Fuel

[0313] The annual U.S. usage of jet fuel in 2006 was about 21 billion gallons (about 80 billion liters). Aeroplane fuel is clear to straw colored. The most common fuel is an unleaded/paraffin oil-based fuel classified as Aeroplane A-1, which is produced to an internationally standardized set of specifications. Aeroplane fuel is a mixture of a large number of different hydrocarbons, possibly as many as a thousand or more. The range of their sizes (molecular weights or carbon numbers) is restricted by the requirements for the product, for example, freezing point or smoke point. Kerosene-type Aeroplane fuel (including Jet A and Jet A-1) has a carbon number distribution between about 8 and 16 carbon numbers. Wide-cut or naphta-type Aeroplane fuel (including Jet B) has a carbon number distribution between about 5 and 15 carbon numbers.

[0314] Both Aeroplanes (Jet A and jet B) may contain a number of additives. Useful additives include, but are not limited to, antioxidants, antistatic agents, corrosion inhibitors, and fuel system icing inhibitor (FSII) agents. Antioxidants prevent gumming and usually, are based on alkylated phenols, for example, AO-30, AO-31, or AO-37. Antistatic agents dissipate static electricity and prevent sparking. Stadis 450 with dinonylnaphthylsulfonic acid (DINNSA) as the active ingredient, is an example. Corrosion inhibitors, e.g., DCI-4A is used for civilian and military fuels and DCI-6A is used for military fuels. FSII agents, include, e.g., Di-EGME.

[0315] A solution is blending algae fuels with existing jet fuel. The present invention provides such a solution. The lipid:organic solvent layer or lipids produced by the method of the present invention can serve as feedstock to produce jet fuel. Thus, in another aspect of the present invention, a method for producing jet fuel is provided. Herewith two methods for

producing jet fuel from the lipid:organic solvent layer or lipids produced by the method of the present invention are provided, fluid catalytic cracking (FCC) and hydrodeoxygenation (HDO).

1. Fluid Catalytic Cracking

[0316] Fluid Catalytic Cracking (FCC) is one method which is used to produce olefins, especially propylene from heavy crude fractions. There are reports in the literature that vegetable oils such as canola oil could be processed using FCC to give a hydrocarbon stream useful as a gasoline fuel.

[0317] The lipid:organic solvent layer or lipids produced by the method of the present invention can be converted to C₂-C₅ olefins. The process involves flowing the lipid:organic solvent layer or lipids produced through an FCC zone and collecting a product stream comprised of olefins, which is useful as a jet fuel. The lipid:organic solvent layer or lipids produced are contacted with a cracking catalyst at cracking conditions to provide a product stream comprising C₂-C₅ olefins and hydrocarbons useful as jet fuel.

[0318] Thus, in yet another aspect of the present invention a method for producing jet fuel is provided. In a preferred embodiment, the method for producing jet fuel comprises the steps of (a) lysing a lipid-containing microorganism to produce a lysate; (b) treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent to form a heterogeneous mixture; (c) separating the heterogeneous mixture into a lipid:organic solvent composition and an aqueous composition and, optionally, an emulsified composition or cell pellet composition; (d) removing the lipid:organic solvent composition from the aqueous composition, emulsion composition, or cell pellet composition; and (e) treating the lipid:organic solvent composition to produce a C₂-C₅ olefine, whereby jet fuel is produced.. The lipid:organic solvent composition comprises the organic solvent and cellular lipids.

[0319] In a preferred embodiment of the method for producing a jet fuel, step (f) is performed by flowing the lipid:organic solvent composition through a fluid catalytic cracking zone. Step (f) may further comprise contacting the lipid:organic solvent composition with a cracking catalyst at cracking conditions to provide a product stream comprising C₂-C₅ olefins.

[0320] In certain embodiments of this method it may be desirable to remove any contaminants that may be present in the lipid:organic solvent composition. Thus, prior to step (f) flowing the lipid:organic solvent composition through a fluid catalytic cracking zone, the lipid:organic solvent composition is pretreated. Pretreatment may involve contacting the lipid:organic solvent composition with an ion-exchange resin. The ion exchange resin is an acidic ion exchange resin, such as AmberlystTM-15 and can be used as a bed in a reactor through which the lipid:organic solvent composition is flowed through, either upflow or downflow. Other pretreatments may include mild acid washes by contacting the lipid:organic solvent composition with an acid, such as sulfuric, acetic, nitric, or hydrochloric acid. Contacting is done with a dilute acid solution usually at ambient temperature and atmospheric pressure.

[0321] The lipid:organic solvent composition, optionally pretreated, is flowed to an FCC zone where the hydrocarbonaceous components are cracked to olefins. Catalytic cracking is accomplished by contacting the lipid:organic solvent composition in a reaction zone with a catalyst composed of finely divided particulate material. The reaction is catalytic cracking, as opposed to hydrocracking, and is carried out in the absence of added hydrogen or the consumption of hydrogen. As the cracking reaction proceeds, substantial amounts of coke are deposited on the catalyst. The catalyst is regenerated at high temperatures by burning coke from the catalyst in a regeneration zone. Coke-containing catalyst, referred to herein as "coked catalyst", is continually transported from the reaction zone to the regeneration zone to be regenerated and replaced by essentially coke-free regenerated catalyst from the regeneration zone. Fluidization of the catalyst particles by various gaseous streams allows the transport of catalyst between the reaction zone and regeneration zone. Methods for cracking hydrocarbons, such as those of the lipid:organic solvent composition described herein, in a fluidized stream of catalyst, transporting catalyst between reaction and regeneration zones, and combusting coke in the regenerator are well known by those skilled in the art of FCC processes. Exemplary FCC applications and catalysts useful for cracking the lipid:organic solvent composition to produce C₂-C₅ olefins are described in U.S. Pat. Nos. 6,538,169, 7,288,685, which are incorporated in their entirety by reference.

[0322] In one embodiment, cracking the lipid:organic solvent composition of the present invention, takes place in the riser section or, alternatively, the lift section, of the FCC zone. The lipid:organic solvent composition is introduced into the riser by a nozzle resulting in the rapid vaporization of the lipid:organic solvent composition. Before contacting the catalyst,

the lipid:organic solvent composition will ordinarily have a temperature of about 149°C to about 316°C (300°F to 600°F). The catalyst is flowed from a blending vessel to the riser where it contacts the lipid:organic solvent composition for a time of about 2 seconds or less.

[0323] The blended catalyst and reacted lipid:organic solvent composition vapors are then discharged from the top of the riser through an outlet and separated into a cracked product vapor stream including olefins and a collection of catalyst particles covered with substantial quantities of coke and generally referred to as "coked catalyst." In an effort to minimize the contact time of the lipid:organic solvent composition and the catalyst which may promote further conversion of desired products to undesirable other products, any arrangement of separators such as a swirl arm arrangement can be used to remove coked catalyst from the product stream quickly. The separator, e.g. swirl arm separator, is located in an upper portion of a chamber with a stripping zone situated in the lower portion of the chamber. Catalyst separated by the swirl arm arrangement drops down into the stripping zone. The cracked product vapor stream comprising cracked hydrocarbons including light olefins and some catalyst exit the chamber via a conduit which is in communication with cyclones. The cyclones remove remaining catalyst particles from the product vapor stream to reduce particle concentrations to very low levels. The product vapor stream then exits the top of the separating vessel. Catalyst separated by the cyclones is returned to the separating vessel and then to the stripping zone. The stripping zone removes adsorbed hydrocarbons from the surface of the catalyst by counter-current contact with steam.

[0324] Low hydrocarbon partial pressure operates to favor the production of light olefins. Accordingly, the riser pressure is set at about 172 to 241 kPa (25 to 35 psia) with a hydrocarbon partial pressure of about 35 to 172 kPa (5 to 25 psia), with a preferred hydrocarbon partial pressure of about 69 to 138 kPa (10 to 20 psia). This relatively low partial pressure for hydrocarbon is achieved by using steam as a diluent to the extent that the diluent is 10 to 55 wt-% of lipid:organic solvent composition and preferably about 15 wt-% of lipid:organic solvent composition. Other diluents such as dry gas can be used to reach equivalent hydrocarbon partial pressures.

[0325] The temperature of the cracked stream at the riser outlet will be about 510°C to 621°C (950°F to 1150°F). However, riser outlet temperatures above 566°C (1050°F) make more dry gas and more olefins. Whereas, riser outlet temperatures below 566°C (1050°F) make less ethylene and propylene. Accordingly, it is preferred to run the FCC process at a

preferred temperature of about 566°C to about 630°C, preferred pressure of about 138 kPa to about 240 kPa (20 to 35 psia). Another condition for the process is the catalyst to lipid:organic solvent composition ratio which can vary from about 5 to about 20 and preferably from about 10 to about 15.

[0326] In one embodiment of the method for producing a jet fuel, the lipid:organic solvent composition is introduced into the lift section of an FCC reactor. The temperature in the lift section will be very hot and range from about 700°C (1292°F) to about 760°C (1400°F) with a catalyst to lipid:organic solvent composition ratio of about 100 to about 150. It is anticipated that introducing the lipid:organic solvent composition into the lift section will produce considerable amounts of propylene and ethylene.

[0327] Gas and liquid hydrocarbon products produced can be analyzed by gas chromatography, HPLC, etc.

2. Hydrodeoxygenation

[0328] In another embodiment of the method for producing a jet fuel using the lipid:organic solvent composition or the lipids produced as described herein, the structure of the lipid:organic solvent composition or the lipids is broken by a process referred to as hydrodeoxygenation (HDO). As such step (f) is performed by hydrodeoxygenating the lipid:organic solvent composition.

[0329] HDO means removal of oxygen by means of hydrogen, that is, oxygen is removed while breaking the structure of the material. Olefinic double bonds are hydrogenated and any sulphur and nitrogen compounds are removed. Sulphur removal is called hydrodesulphurization (HDS). Pretreatment and purity of the raw materials (lipid:organic solvent composition or the lipids) contribute to the service life of the catalyst.

[0330] Generally in the HDO/HDS step, hydrogen is mixed with the feed stock (lipid:organic solvent composition or the lipids) and then the mixture is passed through a catalyst bed as a co-current flow, either as a single phase or a two phase feed stock. After the HDO/HDS step, the product fraction is separated and passed to a separate isomerization reactor. An isomerization reactor for biological starting material is described in the literature (FI 100 248) as a co-current reactor.

[0331] The process for producing a fuel by hydrogenating a hydrocarbon feed, e.g., the lipid:organic solvent composition or the lipids herein, can also be performed by passing the

lipid:organic solvent composition or the lipids as a co-current flow with hydrogen gas through a first hydrogenation zone, and thereafter the hydrocarbon effluent is further hydrogenated in a second hydrogenation zone by passing hydrogen gas to the second hydrogenation zone as a counter-current flow relative to the hydrocarbon effluent. Exemplary HDO applications and catalysts useful for cracking the lipid:organic solvent composition to produce C₂-C₅ olefins are described in U.S. Pat. No. 7,232,935 which is incorporated in its entirety by reference.

[0332] Typically, in the hydrodeoxygenation step, the structure of the biological component, such as the lipid:organic solvent composition or lipids herein, is decomposed, oxygen, nitrogen, phosphorus and sulphur compounds, and light hydrocarbons as gas are removed, and the olefinic bonds are hydrogenated. In the second step of the process, i.e. in the so-called isomerization step, isomerization is carried out for branching the hydrocarbon chain and improving the performance of the paraffin at low temperatures.

[0333] In the first step i.e. HDO step of the cracking process, hydrogen gas and the lipid:organic solvent composition or lipids herein which are to be hydrogenated are passed to a HDO catalyst bed system either as co-current or counter-current flows, said catalyst bed system comprising one or more catalyst bed(s), preferably 1-3 catalyst beds. The HDO step is typically operated in a co-current manner. In case of a HDO catalyst bed system comprising two or more catalyst beds, one or more of the beds may be operated using the counter-current flow principle.

[0334] In the HDO step, the pressure varies between 20 and 150 bar, preferably between 50 and 100 bar, and the temperature varies between 200 and 500°C, preferably in the range of 300-400°C.

[0335] In the HDO step, known hydrogenation catalysts containing metals from Group VII and/or VIB of the Periodic System may be used. Preferably, the hydrogenation catalysts are supported Pd, Pt, Ni, NiMo or a CoMo catalysts, the support being alumina and/or silica. Typically, NiMo/Al₂O₃ and CoMo/Al₂O₃ catalysts are used.

[0336] Prior to the HDO step, the lipid:organic solvent composition or lipids herein may optionally be treated by prehydrogenation under milder conditions thus avoiding side reactions of the double bonds. Such prehydrogenation is carried out in the presence of a prehydrogenation catalyst at temperatures of 50-400°C and at hydrogen pressures of 1-200 bar, preferably at a temperature between 150 and 250°C and at a hydrogen pressure between

10 and 100 bar. The catalyst may contain metals from Group VIII and/or VIB of the Periodic System. Preferably, the prehydrogenation catalyst is a supported Pd, Pt, Ni, NiMo or a CoMo catalyst, the support being alumina and/or silica.

[0337] A gaseous stream from the HDO step containing hydrogen is cooled and then carbon monoxide, carbon dioxide, nitrogen, phosphorus and sulphur compounds, gaseous light hydrocarbons and other impurities are removed therefrom. After compressing, the purified hydrogen or recycled hydrogen is returned back to the first catalyst bed and/or between the catalyst beds to make up for the withdrawn gas stream. Water is removed from the condensed liquid. The liquid is passed to the first catalyst bed or between the catalyst beds.

[0338] After the HDO step, the product is subjected to an isomerization step. It is substantial for the process that the impurities are removed as completely as possible before the hydrocarbons are contacted with the isomerization catalyst. The isomerization step comprises an optional stripping step, wherein the reaction product from the HDO step may be purified by stripping with water vapour or a suitable gas such as light hydrocarbon, nitrogen or hydrogen. The optional stripping step is carried out in counter-current manner in a unit upstream of the isomerization catalyst, wherein the gas and liquid are contacted with each other, or before the actual isomerization reactor in a separate stripping unit utilizing counter-current principle.

[0339] After the stripping step the hydrogen gas and the hydrogenated lipid:organic solvent composition or lipids herein, and optionally an n-paraffin mixture, are passed to a reactive isomerization unit comprising one or several catalyst bed(s). The catalyst beds of the isomerization step may operate either in co-current or counter-current manner.

[0340] It is important for the process that the counter-current flow principle is applied in the isomerization step. In the isomerization step this is done by carrying out either the optional stripping step or the isomerization reaction step or both in counter-current manner.

[0341] The isomerization step and the HDO step may be carried out in the same pressure vessel or in separate pressure vessels. Optional prehydrogenation may be carried out in a separate pressure vessel or in the same pressure vessel as the HDO and isomerization steps.

[0342] In the isomerization step, the pressure varies in the range of 20 150 bar, preferably in the range of 20 100 bar, the temperature being between 200 and 500°C, preferably between 300 and 400°C.

[0343] In the isomerization step, isomerization catalysts known in the art may be used. Suitable isomerization catalysts contain molecular sieve and/or a metal from Group VII and/or a carrier. Preferably, the isomerization catalyst contains SAPO-11 or SAPO41 or ZSM-22 or ZSM-23 or ferrierite and Pt, Pd or Ni and Al₂O₃ or SiO₂. Typical isomerization catalysts are, for example, Pt/SAPO-11/Al₂O₃, Pt/ZSM-22/Al₂O₃, Pt/ZSM-23/Al₂O₃ and Pt/SAPO-11/SiO₂.

[0344] As the product, a high quality hydrocarbon component of biological origin, useful as a diesel fuel or a component thereof, is obtained, the density, cetane number and performance at low temperate of said hydrocarbon component being excellent.

VIII. COMPOSITIONS

[0345] Another object of the present invention is to provide compositions comprising lipids isolated by using the methods described herein. A preferred composition comprises (i) a first lipid isolated from a microorganism and (ii) a second lipid, wherein the second lipid is obtained from a source other than the microorganism.

[0346] The first lipid can be isolated from any of the microorganisms described herein.

A. First Lipids

[0347] Methods of the present invention are applicable to extracting a variety of lipids from a variety of microorganisms. Microorganisms described herein produce a variety of lipids, such as phospholipids, free fatty acids, esters of fatty acids, including triglycerides of fatty acids, sterols; pigments (e.g., carotenoids and oxycarotenoids) and other lipids, and lipid associated compounds such as phytosterols, ergothionine, lipoic acid and antioxidants including beta-carotene and tocopherol. Exemplary first lipids include, but are not limited to, C8, C10, C12, C14, C16 and C18 triacylglycerides, lipids containing omega-3 highly unsaturated fatty acids, such as docosahexaenoic acid (DHA), eicosapentaenoic acid (EPA), and/or docosapentaenoic acid (DPA). First lipids also include arachidonic acid, stearidonic acid, cholesterol, desmesterol, astaxanthin, canthaxanthin, and n-6 and n-3 highly unsaturated fatty acids such as eicosapentaenoic acid, docosapentaenoic acid and docosahexaenoic acid.

Other lipids and microorganisms which may be suitable for use in the instant invention will be readily apparent to those skilled in the art.

[0348] Preferred first lipids are lipids containing a relatively large amount of C18 and C16 fatty acids.

1. C18:1

[0349] In a preferred embodiment of the present invention, the first lipid of a composition comprises at least 50% of a C18:1 lipid, preferably at least 60%, and more preferably at least 80%.

2. C10, C12 and C14

[0350] In another preferred embodiment of the present invention, the first lipid of a composition comprises at least 10% of a C10:0, C12:0 and C14:0 lipid combined, preferably at least 15%, more preferably at least 20%, and most preferably at least 30%.

B. Second Lipids

[0351] The second lipid can be any oil, optionally selected from the group consisting of oil from soy, rapeseed, canola, palm, coconut, corn, waste vegetable, Chinese tallow, olive, sunflower, cotton seed, chicken fat, beef tallow, porcine tallow, microalgae, macroalgae, *Cuphea*, flax, peanut, choice white grease, lard, *Camelina sativa*, mustard seed, cashew nut, oats, lupine, kenaf, calendula, hemp, coffee, linseed, hazelnuts, euphorbia, pumpkin seed, coriander, camellia, sesame, safflower, rice, tung oil tree, cocoa, copra, pium poppy, castor beans, pecan nuts, jojoba, jatropha, macadamia nuts, Brazil nuts, and avocado. A preferred second lipid is coconut oil. Another preferred second lipid is palm oil or soy oil. A second lipid can also be a fossil oil such as crude oil or a distillate fraction of crude oil.

C. Compositions Comprising A First Lipid and A Second Lipid

[0352] The first lipid and the second lipid can be provided at different ratios. In a preferred embodiment, the ratio of the first lipid to the second lipid is between 1 and 100. In another preferred embodiment, the ratio of the first lipid to the second lipid is between 1 and 10. In another preferred embodiment, the ratio of the first lipid to the second lipid is about 1. In another embodiment, the ratio of the second lipid to the first lipid is between 10 and 1.

[0353] Additional embodiments of the present invention include optional functional components that would allow one of ordinary skill in the art to perform any of the method variations described herein.

[0354] Although the forgoing invention has been described in some detail by way of illustration and example for clarity and understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings of this invention that certain variations, changes, modifications and substitution of equivalents may be made thereto without necessarily departing from the spirit and scope of this invention. As a result, the embodiments described herein are subject to various modifications, changes and the like, with the scope of this invention being determined solely by reference to the claims appended hereto. Those of skill in the art will readily recognize a variety of non-critical parameters that could be changed, altered or modified to yield essentially similar results.

[0355] While each of the elements of the present invention is described herein as containing multiple embodiments, it should be understood that, unless indicated otherwise, each of the embodiments of a given element of the present invention is capable of being used with each of the embodiments of the other elements of the present invention and each such use is intended to form a distinct embodiment of the present invention.

[0356] The referenced patents, patent applications, and scientific literature, including accession numbers to GenBank database sequences, referred to herein are hereby incorporated by reference in their entirety as if each individual publication, patent or patent application were specifically and individually indicated to be incorporated by reference. Any conflict between any reference cited herein and the specific teachings of this specification shall be resolved in favor of the latter. Likewise, any conflict between an art-understood definition of a word or phrase and a definition of the word or phrase as specifically taught in this specification shall be resolved in favor of the latter. The publications mentioned herein are cited for the purpose of describing and disclosing reagents, methodologies and concepts that may be used in connection with the present invention. Nothing herein is to be construed as an admission that these references are prior art in relation to the inventions described herein. In particular, the following patent applications are hereby incorporated by reference in their entireties for all purposes: US Provisional Application No. 61/028,493, filed February 13, 2008, entitled "Extraction of Lipids from Microorganisms"; US Provisional Application No. 61/036,918, filed March 14, 2008, entitled "Oil Separation Methods"; US Provisional

Application No. 61/043,318, filed April 8, 2008, entitled “Fractionation of Oil-Bearing Biomass”; PCT Patent Application No. PCT/US2009/066142, filed November 30, 2009, entitled “Production of Tailored Oils in Heterotrophic Microorganisms”; and PCT Patent Application No. PCT/US2009/066141, filed November 30, 2009, entitled “Manufacturing of Tailored Oils in Recombinant Heterotrophic Microorganisms”.

[0357] Although this invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications. The present invention also has a wide variety of applications. This application is intended to cover any variations, uses or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth.

IX. EXAMPLES

Example 1: Cultivation of Microalgae to Achieve High Oil Content

[0358] Microalgae strains were cultivated to achieve a high percentage of oil by dry cell weight. Cryopreserved cells were thawed at room temperature and 500 µl of cells were added to 4.5 ml of medium (4.2 g/L K₂HPO₄, 3.1 g/L NaH₂PO₄, 0.2 g/L MgSO₄·7H₂O, 0.25 g/L citric acid monohydrate, 0.025 g/L CaCl₂·2H₂O, 2 g/L yeast extract) plus 2% glucose and grown for 7 days at 28°C with agitation (200 rpm) in a 6-well plate. Dry cell weights were determined by centrifuging 1 ml of culture at 14,000 rpm for 5 min in a pre-weighed Eppendorf tube. The culture supernatant was discarded and the resulting cell pellet washed with 1 ml of deionized water. The culture was again centrifuged, the supernatant discarded, and the cell pellets placed at -80°C until frozen. Samples were then lyophilized for 24 hours and dry cell wights calculated. For determination of total lipid in cultures, 3 ml of culture was removed and subjected to anaylsis using an Ankom system (Ankom Inc., Macedon, NY) according to the manufacturer’s protocol. Samples were subjected to solvent extraction with an Ankom XT10 extractor according to manufacturer’s protocol. Total lipid was determined as the difference in mass between acid hydrolyzed dried samples and solvent extracted, dried samples. Percent oil dry cell weight measurements are shown in Table 12.

[0359] Table 12. Percent oil dry cell weight for microalgae.

Species	Strain	% Oil	Strain #
---------	--------	-------	----------

Chlorella kessleri	UTEX 387	39.42	4
Chlorella kessleri	UTEX 2229	54.07	5
Chlorella kessleri	UTEX 398	41.67	6
Parachlorella kessleri	SAG 11.80	37.78	7
Parachlorella kessleri	SAG 14.82	50.70	8
Parachlorella kessleri	SAG 21.11 H9	37.92	9
Prototheca stagnora	UTEX 327	13.14	10
Prototheca moriformis	UTEX 1441	18.02	11
Prototheca moriformis	UTEX 1435	27.17	12
Chlorella minutissima	UTEX 2341	31.39	13
Chlorella protothecoides	UTEX 250	34.24	1
Chlorella protothecoides	UTEX 25	40.00	2
Chlorella protothecoides	CCAP 211/8D	47.56	3
Chlorella sp.	UTEX 2068	45.32	14
Chlorella sp.	CCAP 211/92	46.51	15
Chlorella sorokiniana	SAG 211.40B	46.67	16
Parachlorella beijerinckii	SAG 2046	30.98	17
Chlorella luteoviridis	SAG 2203	37.88	18
Chlorella vulgaris	CCAP 211/11K	35.85	19
Chlorella reisiigii	CCAP 11/8	31.17	20
Chlorella eplipsoidea	CCAP 211/42	32.93	21
Chlorella saccharophila	CCAP 211/31	34.84	22
Chlorella saccharophila	CCAP 211/32	30.51	23

Example 2: Cultivation of *Chlorella protothecoides*

[0360] Three fermentation processes were performed with three different media formulations with the goal of generating algal biomass with high oil content. The first formulation (Media 1) was based on medium described in Wu et al. (1994 *Science in China*, vol. 37, No. 3, pp. 326-335) and consisted of per liter: KH_2PO_4 , 0.7g; K_2HPO_4 , 0.3g; $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.3g; $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 3mg; thiamine hydrochloride, 10 μg ; glucose, 20g; glycine, 0.1g; H_3BO_3 , 2.9mg; $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 1.8mg; $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 220 μg ; $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 80 μg ; and $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$, 22.9mg. The second medium (Media 2) was derived from the flask media described in Example 1 and consisted of per liter: K_2HPO_4 , 4.2g; NaH_2PO_4 , 3.1g; $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.24g; citric acid monohydrate, 0.25g; calcium chloride dehydrate, 25mg; glucose, 20g; yeast extract, 2g. The third medium (Media 3) was a hybrid and consisted of per liter: K_2HPO_4 , 4.2g; NaH_2PO_4 , 3.1g; $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.24g; citric acid monohydrate, 0.25g; calcium chloride dehydrate, 25mg; glucose, 20g; yeast extract, 2g; H_3BO_3 , 2.9mg; $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 1.8 mg; $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 220 μg ; $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 80 μg ; and $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$, 22.9mg. All three media formulations were prepared and autoclave sterilized in lab scale fermentor vessels for 30 minutes at 121°C. Sterile glucose was added to each vessel following cool down post autoclave sterilization.

[0361] Inoculum for each fermentor was *Chlorella protothecoides* (UTEX 250), prepared in two flask stages using the medium and temperature conditions of the fermentor inoculated. Each fermentor was inoculated with 10% (v/v) mid-log culture. The three lab scale fermentors were held at 28°C for the duration of the experiment. The microalgal cell growth in Media 1 was also evaluated at a temperature of 23°C. For all fermentor evaluations, pH was maintained at 6.6-6.8, agitations at 500rpm, and airflow at 1 vvm. Fermentation cultures were cultivated for 11 days. Biomass accumulation was measured by optical density at 750 nm and dry cell weight.

[0362] Lipid/oil concentration was determined using direct transesterification with standard gas chromatography methods. Briefly, samples of fermentation broth with biomass was blotted onto blotting paper and transferred to centrifuge tubes and dried in a vacuum oven at 65-70°C for 1 hour. When the samples were dried, 2mL of 5% H₂SO₄ in methanol was added to the tubes. The tubes were then heated on a heat block at 65-70°C for 3.5hours, while being vortexed and sonicated intermittently. 2ml of heptane was then added and the tubes were shaken vigorously. 2Ml of 6% K₂CO₃ was added and the tubes were shaken vigorously to mix and then centrifuged at 800rpm for 2 minutes. The supernatant was then transferred to GC vials containing Na₂SO₄ drying agent and ran using standard gas chromatography methods. Percent oil/lipid was based on a dry cell weight basis. The dry cell weights for cells grown using: Media 1 at 23°C was 9.4g/L; Media 1 at 28°C was 1.0g/L, Media 2 at 28°C was 21.2g/L; and Media 3 at 28°C was 21.5g/L. The lipid/oil concentration for cells grown using: Media 1 at 23°C was 3g/L; Media 1 at 28°C was 0.4g/L; Media 2 at 28°C was 18 g/L; and Media 3 at 28°C was 19g/L. The percent oil based on dry cell weight for cells grown using: Media 1 at 23°C was 32%; Media 1 at 28°C was 40%; Media 2 at 28°C was 85%; and Media 3 at 28°C was 88%. The lipid profiles (in area %, after normalizing to the internal standard) for algal biomass generated using the three different media formulations at 28°C are summarized below in Table 13.

[0363] Table 13. Lipid profiles for *Chlorella protothecoides* grown under different media conditions.

	Media 1 28°C (in Area %)	Media 2 28°C (in Area %)	Media 3 28°C (in Area %)
C14:0	1.40	0.85	0.72
C16:0	8.71	7.75	7.43
C16:1	--	0.18	0.17
C17:0	--	0.16	0.15
C17:1	--	0.15	0.15

C18:0	3.77	3.66	4.25
C18:1	73.39	72.72	73.83
C18:2	11.23	12.82	11.41
C18:3 alpha	1.50	0.90	1.02
C20:0	--	0.33	0.37
C20:1	--	0.10	0.39
C20:1	--	0.25	--
C22:0	--	0.13	0.11

Example 3: Genotyping of Microalgae with High Oil Content

[0364] Microalgae samples from the 23 strains listed in Table 12 above were genotyped. Genomic DNA was isolated from algal biomass as follows. Cells (approximately 200 mg) were centrifuged from liquid cultures for 5 minutes at 14,000 x g. Cells were then resuspended in sterile distilled water, centrifuged for 5 minutes at 14,000 x g and the supernatant discarded. A single glass bead ~2mm in diameter was added to the biomass and tubes were placed at -80°C for at least 15 minutes. Samples were removed and 150 µl of grinding buffer (1% Sarkosyl, 0.25 M sucrose, 50 mM NaCl, 20 mM EDTA, 100 mM Tris-HCl, pH 8.0, RNase A 0.5 µg/µl) was added. Pellets were resuspended by vortexing briefly, followed by the addition of 40 µl of 5M NaCl. Samples were vortexed briefly, followed by the addition of 66 µl of 5% CTAB (Cetyl trimethylammonium bromide) and a final brief vortex. Samples were next incubated at 65°C for 10 minutes after which they were centrifuged at 14,000 x g for 10 minutes. The supernatant was transferred to a fresh tube and extracted once with 300 µl Phenol:Chloroform:Isoamyl alcohol 12:12:1, followed by centrifugation for 5 minutes at 14,000 x g. The resulting aqueous phase was transferred to a fresh tube containing 0.7 vol of isopropanol (~190 µl), mixed by inversion and incubated at room temperature for 30 minutes or overnight at 4°C. DNA was recovered via centrifugation at 14,000 x g for 10 minutes. The resulting pellet was then washed twice with 70% ethanol, followed by a final wash with 100% ethanol. Pellets were air dried for 20-30 minutes at room temperature followed by resuspension in 50 µl of 10 mM TrisCl, 1mM EDTA (pH 8.0).

[0365] Five µl of total algal DNA, prepared as described above, was diluted 1:50 in 10 mM Tris, pH 8.0. PCR reactions, final volume 20 µl, were set up as follows. Ten µl of 2 x iProof HF master mix (BIO-RAD) was added to 0.4 µl primer SZ02613 (5'-TGTTGAAGAATGAGCCGGCGAC-3' (SEQ ID NO:5) at 10mM stock concentration). This primer sequence runs from position 567-588 in Gen Bank accession no. L43357 and is highly conserved in higher plants and algal plastid genomes. This was followed by the addition of 0.4 µl primer SZ02615 (5'-CAGTGAGCTATTACGCACTC-3' (SEQ ID NO:6)

at 10 mM stock concentration). This primer sequence is complementary to position 1112-1093 in Gen Bank accession no. L43357 and is highly conserved in higher plants and algal plastid genomes. Next, 5 µl of diluted total DNA and 3.2 µl dH₂O were added. PCR reactions were run as follows: 98°C, 45''; 98°C, 8''; 53°C, 12''; 72°C, 20'' for 35 cycles followed by 72°C for 1 min and holding at 25°C. For purification of PCR products, 20 µl of 10 mM Tris, pH 8.0, was added to each reaction, followed by extraction with 40 µl of Phenol:Chloroform:isoamyl alcohol 12:12:1, vortexing and centrifuging at 14,000 x g for 5 minutes. PCR reactions were applied to S-400 columns (GE Healthcare) and centrifuged for 2 minutes at 3,000 x g. Purified PCR products were subsequently TOPO cloned into PCR8/GW/TOPO and positive clones selected for on LB/Spec plates. Purified plasmid DNA was sequenced in both directions using M13 forward and reverse primers. Sequences from strains 1-23 (designated in Example 1, Table 12) are listed as SEQ ID NOs: 7-29 in the attached Sequence Listing.

Example 4: Genomic DNA analysis of 23S rRNA from *Chlorella protothecoides* and *Prototheca*

[0366] Genomic DNA from 8 strains of *Chlorella protothecoides* (UTEX 25, UTEX 249, UTEX 250, UTEX 256, UTEX 264, UTEX 411, CCAP 211/17 and CCAP 211/8d) was isolated and genomic DNA analysis of 23S rRNA was performed according to the methods described above in Example 3. All strains of *Chlorella protothecoides* tested were identical in sequence except for UTEX 25. Sequences for all eight strains are listed as SEQ ID NOs: 30 and 31 in the attached Sequence Listing.

[0367] The 23s rRNA genomic sequence for *Prototheca moriformis* UTEX 1436 (SEQ ID NO: 32) was also compared to other *Prototheca* species (UTEX 1435, UTEX 1437, and UTEX 1439) and the above described *Chlorella protothecoides* strains. The comparison showed that the 23s rRNA genomic sequence for *Prototheca moriformis* UTEX 1436 was dissimilar to the other *Prototheca* genotypes (SEQ ID NO: 33).

Example 5: Diversity of Lipid Chains in Microalgal Species

[0368] Lipid samples from a subset of strains grown in Example 1, Table 12 were analyzed for lipid profile using HPLC. Results are shown below in Table 14.

[0369] Table 14. Diversity of lipid chains in microalgal species.

Strain #	C:14:0	C:16:0	C:16:1	C:18:0	C:18:1	C:18:2	C:18:3	C:20:0	C:20:1
from Table 12									
1	0.57	10.30	0	3.77	70.52	14.24	1.45	0.27	0
2	0.61	8.70	0.30	2.42	71.98	14.21	1.15	0.20	0.24
4	0.68	9.82	0	2.83	65.78	12.94	1.46	0	0
5	1.47	21.96	0	4.35	22.64	9.58	5.2	3.88	3.3
10	0	12.01	0	0	50.33	17.14	0	0	0
11	1.41	29.44	0.70	3.05	57.72	12.37	0.97	0.33	0
12	1.09	25.77	0	2.75	54.01	11.90	2.44	0	0

Example 6: Extraction Of Lipids From Microalgae Using Coconut Oil

1. Cell Production

[0370] An F-Tank batch of *Chlorella protothecoides* (about 1,200 gallons) was used to generate biomass for extraction processes. The batch (#ZA07126) was allowed to run for 100 hours, while controlling the glucose levels at 16 g/L, after which time the corn syrup feed was terminated. Residual glucose levels dropped to <0 g/L two hours later. This resulted in a final age of 102 hours. The final broth volume was 1,120 gallons. Both in-process contamination checks and a thorough analysis of a final broth sample failed to show any signs of contamination.

2. Cell Disruption

[0371] Lyophilized *Chlorella protothecoides* cells were resuspended to 200 g/L with DI water containing 20 g/L KOH. This cell suspension was autoclaved at 130°C for 30 minutes and cooled to room temperature. A 10 ml sample of this material was sonicated using a Misonix 3000 sonicator equipped with a micro-tip on level 7. The suspension was sonicated for 6 minutes on a 30 sec on/off cycle for a total of 3 cycles. More than 70% breakage was observed by microscope analysis. An aliquot of this suspension (1 ml sample) was centrifuged at 14 K rpm for 15 min. Three layers were observed: a pellet of heavy solids, an aqueous phase and an emulsion phase. No oil layer was observed.

[0372] A second set of experiments were performed looking at cell breakage efficiency at a variety of times of sonication to achieve maximal cell breakage. 6 ml samples of *Chlorella protothecoides* cells prepared as described above were sonicated using a Misonix 3000 sonicator equipped with a micro-tip on level 8 with interval times of 20 seconds on and 20 seconds off. The sonication times ranged from 10 minutes, 20 minutes and 30 minutes and were all performed with the samples on ice. The samples were then centrifuged at 4300 rpm for 30 minutes. Three layers were observed: a pellet of heavy solids, an aqueous phase and

an emulsion phase. No oil layer was observed. Each time point sample was also analyzed qualitatively for total cell breakage by looking for whole cells in the lysate under 100x magnification. No whole cells were observed in any of the time point samples, indicating that 10 minutes of sonication was sufficient to achieve maximal cell breakage. Micrographs of the lysate under magnification showed oil droplets and cell debris consisting of organelles and disrupted cell membranes. To further inspection of the cell pellets from each of the samples showed no whole cells were observed, only cell debris.

3. Algal Oil Purification For Control

[0373] *Chlorella protothecoides* oil was obtained from dried biomass via hexane extraction using standard methods known in the art (see for example US Patents 5,567,732; 6,255,505; and Yamada *et al.*, Industrial applications of single cell oils, Eds. Kyle and Ratledge, 118-138 (1992)).

4. HPLC Analysis Of Fatty Acid Profiles

[0374] The fatty acid concentrations and profiles of (i) *Chlorella* cells, (ii) pure algal oil, (iii) pure coconut oil, and (iv) the extracted oil samples were determined by HPLC.

[0375] A 50 µl sample of the (200 g/L) washed *Chlorella* cell suspension (before caustic and heat treatment; see above) was hydrolyzed by incubation at 80°C in an isopropanol/KOH saturated solution for 4 hours. The cells were then centrifuged at 14 K rpm for 10 min and the hydrolysate was loaded onto an HPLC column. Samples were analyzed using an Agilent 1100 HPLC using the following method. The samples were derivatized with bromophenacyl bromide (60 mg/ml) and loaded onto a Luna 5u C8(2) 100A 150x2 mm column (Phenomenex). The samples were eluted from the column using a gradient of water to 100% Acetonitrile:tetrahydrofuran (95:5). Signals were detected using DAD array detector at a wavelength of 254nm. The result of such an analysis detailing the lipid profiles for pure oils (algal and coconut oil), established separately (i.e., not mixed with each other), is shown in Figure 1. This result demonstrated that the lipid profiles of pure algal oil and coconut oil are different. As such, coconut oil can be used as an organic solvent to extract lipids from microalgae and the percent of lipids extracted can be determined (see below).

5. Construction And Validation Of Theoretical Curves

[0376] The fatty acids from mixtures of pure algal oil extracted from *Chlorella* and coconut oil were measured as percentages of the total lipid and are shown in Table 15.

[0377] Table 15. Fatty acid content as percentage of total lipid.

% Coconut Oil	% Algal Oil	C18:1	C12	C14	C18:1/C12	C18:1/C14
0	100	59	0.01	1.00	5,900	59.00
2	98	57.9	1.09	8.14	53.13	7.12
5	95	56.3	2.71	8.39	20.76	6.71
10	90	53.5	5.41	8.81	9.89	6.08
20	80	48.0	10.81	9.64	4.44	4.98
25	75	45.3	13.51	10.06	3.35	4.5
40	60	37.0	21.61	11.31	1.71	3.27
50	50	31.5	27.01	12.15	1.17	2.59
60	40	26.0	32.40	12.98	0.80	2.00
27	25	17.8	40.5	14.23	0.44	1.25
80	20	15.0	43.2	14.65	0.35	1.02
89.75	10.25	9.6	48.47	15.46	0.20	0.62
94.75	5.25	6.9	51.17	15.88	0.13	0.43
100	0	4	54	17.00	0.07	0.24

[0378] In pure algal oil, 59% of the total fatty acids are of the C18:1 type, 0.01% are of the C12 type, and 1% are of the C14 type. The remaining 39.99% of algal fatty acids are C10:1, C10, C18:3, C18:2, C16 or C18 fatty acids, the majority being C18:2 fatty acids (see Figure 1). In pure coconut oil, 4% of the fatty acids are of the C18:1 type, 54% are of the C12 type, and 17% are of the C14 type. The remaining 25% are C10:1, C10, C18:3, C18:2, C16 or C18 fatty acids (see Figure 1). Upon mixing algal oil with coconut oil, the ratio of the respective fatty acids changes. For example, in a 1:1 mixture of pure algal oil and pure coconut oil, 31.5% of the total fatty acids are of the C18:1 type (where the majority is from algal oil), 27.01% are of the C12 type (which are almost exclusively from coconut oil), and 12.15% are of the C14 type (where the majority is from coconut oil). This reference table is useful for determining the percent of algal fatty acids and the percent of coconut fatty acids upon HPLC analysis of a lipid:organic solvent composition obtained from the lightest layer of a microalgal biomass lysate that has been subjected to centrifugation.

[0379] A curve was constructed based on the expected total C18:1 vs C12 or C14 in a mixture of the two pure oils, i.e., algal oil and coconut oil. The theoretical measurements for both curves were validated by HPLC analysis of a 1:1 mixture of pure algal to pure coconut oil which should result in 50% of each ratio. The measured value from the C18:1 to C12 and C14 was 51% and 65% respectively when fitted to the corresponding curve (Figure 2).

6. Separating Microbial Oil From An Emulsion Using An Organic Solvent (Coconut Oil Example)

[0380] In order to free the oil from the *Chlorella* lysate, coconut oil was added to 1ml of disrupted *Chlorella* cell material at ratios of 3:1, 1:1, 1:3 and 1:10 (algal oil to coconut oil), mixed well and heated for 15 min at 70°C and then centrifuged for 15 min at 14 K rpm in a microcentrifuge. Upon centrifugation 4 layers were observed: a pellet of heavy solids, an aqueous phase, an emulsion phase an oil phase. Then the oil phase was separated from the other layers by pipetting. The oil phase included a mixture of algal oil and coconut oil of which an aliquot as subjected to HPLC analysis. The measurement of algal oil in the suspension was based on oil content of cells. Controls included (i) 3 mg of pure algal oil, (ii) 3 mg of pure coconut oil; and (iii) a 1:1 mixture of pure algal oil and pure coconut oil. A 5 µl sample from each was analyzed by HPLC.

7. Results

[0381] The oil content of the starting cell mass (CM) was approximately 30% as determined by HPLC. Coconut oil was used to extract lipids from a disrupted cell suspension of *Chlorella* (200 g/L was suspension) in mixtures of 3:1, 1:1, 1:3 and 1:10 (algal to coconut oil). The samples in this experiment included: (1) 3mg pure algal oil (control); (2) 3 mg pure coconut oil (control); (3) 1:1 mix of 3 mg of each pure oil (control); (4) CM (1ml) + coconut oil (3:1); (5) CM (1ml) + coconut oil (1:1); (6) CM (1ml) + coconut oil (1:3); (7) CM (1ml) + coconut oil (1:10); and (8) CM (1ml).

[0382] No oil layer was present in the sample without coconut oil (8). The fatty acid composition (expressed as % fatty acid of total lipid) of the pure oil controls and each of the extraction mixtures (experiments (1) - (7)) is shown in Table 16.

[0383] Table 16. Fatty acid composition of pure oil controls and extraction mixtures.

% Fatty Acid of Total	Algal Oil (1)	Coconut Oil (2)	1:1 Mix of Algal Oil and Coconut Oil (3)	C.p./C.oil(3:1) (4)	C.p./C.oil (1:1) (5)	C.p./C.oil(1:3) (6)	C.p./C.oil (1:10) (7)
C10:1	0%	1%	0%	0%	0%	0%	0%
C10	0%	10%	5%	9%	10%	9%	10%
C12	0%	56%	27%	50%	53%	53%	54%
C18:3	4%	0%	1%	0%	0%	0%	0%
C14	1%	18%	9%	18%	17%	18%	18%
C18:2	25%	1%	11%	1%	1%	1%	1%
C16	11%	8%	9%	8%	8%	8%	8%
C18:1	59%	4%	33%	11%	8%	8%	6%
C18:0	0%	2%	4%	3%	3%	3%	2%

C.p., *Chlorella protothecoides*; C.oil, coconut oil

[0384] The ratios of C18:1 to C12 and to C14 were calculated and fitted into the theoretical graph (see above and Figure 2). The extraction efficiency of the coconut oil was determined by dividing the calculated percent of algal oil (from theoretical curve) by the expected percent (oil content in starting material). The extraction efficiency ranged from 44% in the 1:10 (algal oil to coconut oil) mixture (see experiment (7) above) to 16% in the 3:1 mixture (see experiment (4) above) as determined from the C18:1 to C12 data (Table 17). Thus, the higher the percent of coconut oil in the extraction procedure, the higher the yield of extracting lipids from *Chorella*.

[0385] Table 17. Extraction efficiency as a function of coconut oil fraction.

% Algal Oil in Extraction	Measured C18:1 to C12	% Measured	% Extraction (C18:1/C12)
9 (7)	0.12	4.0	44%
25 (6)	0.14	6.0	24%
50 (5)	0.16	7.0	14%
75 (4)	0.22	12.0	16%

[0386] The range of the extraction efficiency was 33% (see experiment (7) above) to 14% (see experiment (4) above) as determined from the C18:1 to C14 data (Table 18). Thus, the higher the percent of coconut oil in the extraction procedure, the higher the yield of extracting lipids from *Chorella*.

[0387] Table 18. Extraction efficiency as a function of coconut oil fraction.

% Algal Oil in Extraction	Measured C18:1 to C14	% Measured	% Extraction (C18:1/C14)
9 (7)	0.35	3.0	33%
25 (6)	0.43	5.3	21%
50 (5)	0.47	6.3	13%
75 (4)	0.62	10.3	14%

[0388] The above experiments demonstrated that an organic solvent (in this case coconut oil) efficiently extracted lipids from an algal cell mass to yield a lipid:organic solvent composition.

[0389] An additional set of experiments were performed to look at the efficiency of extracting lipids from a disrupted cell suspension of *Chlorella protothecoides* (sonication) using four emulsion breaking surfactants. The surfactants that were used in this study were:

oleamide DEA (NINOL 201), laurelamide DEA (NINOL 55LL), laurelamide DEA (NINOL 96L) and cocoamide DEA (NINOL 40CO). Each surfactant was added in the concentrations of 0.2%, 0.5% and 1.0% to the cell lysate suspensions and were placed in a oven at 55°C for 15 hours. Each tube was then centrifuged at 4200 rpm for 20 minutes. In each sample, four phases could be seen: a solid phase, an aqueous phase, and emulsion layer and an oil layer on the top. Negative control samples with no added surfactant produced no visible oil phase, but produced only the solid, aqueous and emulsion layers. The oil layers from each of the surfactant treated samples were carefully pipetted off and were weighed to determine overall oil yields. The results are summarized below in Table 19. These results show that an organic solvent, such as a surfactant, efficiently extracted lipids from lipid-containing algae biomass to yield a lipid:organic solvent composition.

[0390] Table 19. Use of surfactants to extract lipids from algae biomass.

Surfactant Name	NINOL 201	NINOL 201	NINOL 201	NINOL 55LL	NINOL 55LL	NINOL 55LL	NINOL 40CO	NINOL 40CO	NINOL 40CO	NINOL 96SL	NINOL 96SL	NINOL 96SL
Surfactant	Oleamide DEA	Oleamide DEA	Oleamide DEA	Laurelamid DEA	Laurelamid DEA	Laurelamid DEA	Cocoamide DEA	Cocoamide DEA	Cocoamide DEA	Laurelamid DEA	Laurelamid DEA	Laurelamid DEA
Percent v/v	1.0%	0.5%	0.2%	1.0%	0.5%	0.2%	1.0%	0.5%	0.2%	1.0%	0.5%	0.2%
Oil in Form g/L	62	62	62	62	62	62	62	62	62	62	62	62
(g) Oil in 6 mL sample (theoretical)	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
Temperature (C)	55	55	55	55	55	55	55	55	55	55	55	55
Time (hrs)	15	15	15	15	15	15	15	15	15	15	15	15
TOTAL												
Oil in sample (g)	0.183	0.237	0.041	0.262	0.012	0.026	0.276	0.082	0.000	0.286	0.035	0.000
% of theoretical	58.07%	76.55%	13.11%	84.44%	4.03%	8.40%	88.01%	26.45%	0.00%	92.30%	11.41%	0.00%

Example 7: Lysing of Chlorella Protothecoides

1. Biomass Production

[0391] An F-Tank batch of *Chlorella protothecoides* (about 1,200 gallons) was used to generate biomass for extraction processes. The batch (#ZA07126) was allowed to run for 100 hours, while controlling the glucose levels at 16 g/L, after which time the corn syrup feed was terminated. Residual glucose levels dropped to <0 g/L two hours later. This resulted in a final age of 102 hours. The final broth volume was 1,120 gallons. Both in-process contamination checks and a thorough analysis of a final broth sample failed to show any signs of contamination.

2. Heat and Chemical Treatment

[0392] Cells were resuspended in water to a biomass concentration of 150 g/L and aliquoted into 27 x 5 mL tubes. Each tube was conditioned per the matrix below in Table 20.

[0393] Table 20. Chemical/heat treatment matrix.

Tube	Condition	[KOH] mN	[H ₂ SO ₄] mN	Temp, 30 min
1	Control	0	0	25 °C
2	40 mN KOH	40	0	25 °C
3	80 mN KOH	80	0	25 °C
4	120 mN KOH	120	0	25 °C
5	160 mN KOH	160	0	25 °C
6	40 mN H ₂ SO ₄	0	40	25 °C
7	80 mN H ₂ SO ₄	0	80	25 °C
8	120 mN H ₂ SO ₄	0	120	25 °C
9	160 mN H ₂ SO ₄	0	160	25 °C
10	Control	0	0	65 °C
11	40 mN KOH	40	0	65 °C
12	80 mN KOH	80	0	65 °C
13	120 mN KOH	120	0	65 °C
14	160 mN KOH	160	0	65 °C
15	40 mN H ₂ SO ₄	0	40	65 °C
16	80 mN H ₂ SO ₄	0	80	65 °C
17	120 mN H ₂ SO ₄	0	120	65 °C
18	160 mN H ₂ SO ₄	0	160	65 °C
19	Control	0	0	130 °C
20	40 mN KOH	40	0	130 °C
21	80 mN KOH	80	0	130 °C
22	120 mN KOH	120	0	130 °C
23	160 mN KOH	160	0	130 °C
24	40 mN H ₂ SO ₄	0	40	130 °C
25	80 mN H ₂ SO ₄	0	80	130 °C
26	120 mN H ₂ SO ₄	0	120	130 °C
27	160 mN H ₂ SO ₄	0	160	130 °C

[0394] After treatment, a 1.5 ml sample was centrifuged at 14 K rpm for 15 min, and the size of each of the resulting layers was measured.

3. Emulsion Water washing

[0395] A fresh set of samples were generated under conditions for samples 21-27 using a total of 10 mL cell suspension, and aliquoted into 3 x 3 mL samples. The resulting emulsions were isolated in new 15 mL tubes, and subjected to the following water wash scheme according to Table 21.

[0396] Table 21. Emulsion water washing matrix.

	Wash 1	Wash 2
Sample	Ratio of water wash volume:emulsion	Ratio of water wash volume:emulsion
21	1 to 1	1 to 1

22	1 to 1	1 to 1
23	1 to 1	1 to 1
24	1 to 1	1 to 1
25	1 to 1	1 to 1
26	1 to 1	1 to 1
27	1 to 1	1 to 1
Sample	Ratio of water wash volume:emulsion	Ratio of water wash volume:emulsion
21	5 to 1	5 to 1
22	5 to 1	5 to 1
23	5 to 1	5 to 1
24	5 to 1	5 to 1
25	5 to 1	5 to 1
26	5 to 1	5 to 1
27	5 to 1	5 to 1
Sample	Ratio of water wash volume:emulsion	Ratio of water wash volume:emulsion
21	10 to 1	10 to 1
22	10 to 1	10 to 1
23	10 to 1	10 to 1
24	10 to 1	10 to 1
25	10 to 1	10 to 1
26	10 to 1	10 to 1
27	10 to 1	10 to 1

[0397] The water was added to the emulsion, vortexed, and centrifuged at 4,400 rpm for 10 minutes. Following the 1st wash, the aqueous phase was removed, and a 2nd water wash performed at the same ratio, vortexed, and centrifuged at 4,400 rpm for 10 minutes. Following treatment, the emulsions were observed for the appearance of a visible oil layer.

4. Sonication

[0398] A sampling of conditions (samples #1, 9, 10, 14, 18, 19, 23, 27) from the chemical/heat treatment matrix were sonicated using a Misonix 3000 sonicator equipped with a micro-tip on level 7. The suspension was sonicated for 3 minutes on a 30 sec on/20 sec off cycle. After treatment, a 1.5 ml sample was centrifuged at 14 K rpm for 15 min and the size of each of the resulting layers was measured.

5. Emulsion Freezing

[0399] Samples which generated emulsions were subjected to freezing at -20°C for >24 hours. Following freezing, the samples were thawed at room temperature, vortexed, and centrifuged at 4,400 rpm for 10 minutes. The samples studied are listed in Table 22.

[0400] Table 22. Emulsions used in freezing cycle.

Sample	Sonicated
160 mN HCl, 130 °C for 30 min	Yes, 30 min cycle
120 mN HCl, 130 °C for 30 min	No
160 mN KOH 130 °C for 30 min	Yes, 30 min cycle
No chemical, 130 °C for 30 min	Yes, 30 min cycle
No chemical, no heat	Yes, 30 min cycle

[0401] Following treatment, the emulsions were observed for the appearance of a visible oil layer.

6. Enzyme Treatment Experiments #1

[0402] Several enzymes were evaluated as an alternative method of cell breakage. These enzymes, all obtained from Sigma Aldrich, were Hemicellulase (P/N H2125), Pectinase (P/N P2401), Cellulase (C9422), and Driselase (D9515). 5 mL of cell material was prepared to various biomass concentrations, and exchanged into buffer. The enzyme concentrations were then prepared as described in Table 23.

[0403] Table 23. Enzyme treatment matrix.

Enzyme	Enzyme conc. (%)	Incubation temp, °C	Incubation time, hrs	Incubation buffer	[BM] g/L
Hemicellulase	1%	25 °C	12	100 mM Acetate, pH 4.6	150
Pectinase	1%	25 °C	12	100 mM Acetate, pH 4.6	150
Cellulase	1%	25 °C	12	100 mM Acetate, pH 4.6	150
Driselase	1%	25 °C	12	100 mM Acetate, pH 4.6	150
Cocktail of all 4 enzymes	1% of each	25 °C	12	100 mM Acetate, pH 4.6	150
Hemicellulase	1%	37 °C	1	100 mM Acetate, pH 4.6	150
Pectinase	1%	37 °C	1	100 mM Acetate, pH 4.6	150
Cellulase	1%	37 °C	1	100 mM Acetate, pH 4.6	150
Driselase	1%	37 °C	1	100 mM Acetate, pH 4.6	150
Cocktail of all 4 enzymes	1% of each	37 °C	1	100 mM Acetate, pH 4.6	150
Hemicellulase/Driselase	4%/2%	37 °C	12	50 mM Citrate, pH 4.5	150
Hemicellulase/Driselase	4%/2%	50 °C	12	50 mM Citrate, pH 4.5	150
Hemicellulase/Driselase	4%/2%	37 °C	12	50 mM Citrate, pH 4.5	100
Hemicellulase/Driselase	4%/2%	50 °C	12	50 mM Citrate, pH 4.5	100
Hemicellulase/Driselase	4%/2%	37 °C	12	50 mM Citrate, pH 4.5	50
Hemicellulase/Driselase	4%/2%	50 °C	12	50 mM Citrate, pH 4.5	50
Hemicellulase/Driselase	4%/2%	37 °C	12	50 mM Citrate, pH 4.5	25
Hemicellulase/Driselase	4%/2%	50 °C	12	50 mM Citrate, pH 4.5	25

7. Enzyme Treatment Experiments #2

[0404] Additional enzymes were evaluated alone, and in combination with one another, and compared against disruption of cells by storing and aging the cells with or without enzyme treatment. A polysaccharide-degrading enzyme mannose and a protease were evaluated for their ability to disrupt cells of *Chlorella protothecoides*. Also, the effect of storing and aging a sample of cultured cells on the weakening of the cell structure to facilitate oil extraction was evaluated.

[0405] Three experiments were conducted. *Chlorella protothecoides* cells were generated via heterotrophic growth using glucose as the sole carbon source. In experiments in which dried cells were used, fermentation broth was centrifuged and the cell paste was subjected to drum drying. In experiments in which cells were stored for 7 days, storage was at 4°C. A polysaccharide-degrading enzyme (Mannaway 4.0L) and protease (Alcalase 2.4 FG) (both from Novozymes) were used in the experiments, as indicated. Enzymes were added at a concentration of 0.2% weight/volume.

[0406] In the first experiment, dried cells were reconstituted in deionized water to a dry cell weight of 155 grams/liter in flasks, and maintained at pH 7.5. Flasks were placed in an incubating shaker for 22 hours at 50°C. Enzymatic cell lysis for sample 1 in Experiment 1 was very effective when combining enzymes, as evidenced by 90.4% of total oil being found in the emulsion and not in the pellet. Neither enzyme individually was as effective as the combination of both, nor was the control without enzyme treatment. The results are shown in Table 27 below.

[0407] In the second experiment, fresh fermentation broth was separated into two aliquots for fresh and 7 day broth experiments. The fresh cells were concentrated by centrifugation and reconstituted in deionized water to 170 g/l in flasks. Resuspended cell samples were then subdivided into two sets of 4 flasks, half of which were stored for 7 days at 4°C before further treatment. For experimental treatment, 3 flasks contained combinations of a protease and a polysaccharide-degrading mannase while the 4th contained no enzymes (see Table 27 for experimental conditions). Flasks were placed in an incubating shaker for 22 hours at 50°C. This second experiment indicated extensive cell lysis with dual enzyme treatment at day 0 and day 7, as well as lysis in the absence of enzymes after 7 days. The results are shown in Table 28 below.

[0408] In the third experiment, fermentation broth was collected from fresh fermentations, concentrated by centrifugation and re-suspended in deionized water at two different concentrations of 71 and 115 g/l and agitated at 150rpm in Applikon fermentors for 22 hours at 50°C. pH was maintained at 7.5 by feedback acid/base control. The results of this third experiment indicated that cell disruption and/or emulsion formation is facilitated by addition of enzymes and aging of the cells. The results are shown in Table 29 below.

8. Oil Analysis and Characterization

[0409] TLC was performed on aluminum backed silica TLC plates (Sigma# 60805). The developing solvent (mobile phase) was hexane/diethyleter/acetic acid (80:20:1, by vol). Spots were visualized by spraying with 10% sulfuric acid and charring on a hotplate. Recovered oil samples from control, acidic, and basic conditions were analyzed to determine if any significant product degradation and/or modification occurred due to the processing conditions.

9. Results

(a) Heat and Chemical Treatment

[0410] The results from the chemical/heat treatment matrix are reported in Figure 3 and Table 24 below.

[0411] Table 24. Measurement of layers from chemical/heat treatment matrix.

Tube	Condition	[KOH] mN	[H ₂ SO ₄] mN	Temp, 30 min	Pellet, mm	Aqueous, mm	Emulsion, mm
1	Control	0	0	25 °C	15	15	0
2	40 mN KOH	40	0	25 °C	15	15	0
3	80 mN KOH	80	0	25 °C	15	15	0
4	120 mN KOH	120	0	25 °C	15	15	0
5	160 mN KOH	160	0	25 °C	15	15	0
6	40 mN H ₂ SO ₄	0	40	25 °C	15	15	0
7	80 mN H ₂ SO ₄	0	80	25 °C	15	15	0
8	120 mN H ₂ SO ₄	0	120	25 °C	15	15	0
9	160 mN H ₂ SO ₄	0	160	25 °C	15	15	0
10	Control	0	0	65 °C	10	19	1
11	40 mN KOH	40	0	65 °C	10	19	1
12	80 mN KOH	80	0	65 °C	10	19	1
13	120 mN KOH	120	0	65 °C	9	19	2
14	160 mN KOH	160	0	65 °C	8	19	3
15	40 mN H ₂ SO ₄	0	40	65 °C	10	19	1
16	80 mN H ₂ SO ₄	0	80	65 °C	10	19	1
17	120 mN H ₂ SO ₄	0	120	65 °C	10	19	1
18	160 mN H ₂ SO ₄	0	160	65 °C	10	19	1

	H ₂ SO ₄						
19	Control	0	0	65 °C	5	15	10
20	40 mN KOH	40	0	130 °C	10	18	2
21	80 mN KOH	80	0	130 °C	5	15	10
22	120 mN KOH	120	0	130 °C	5	15	10
23	160 mN KOH	160	0	130 °C	5	15	10
24	40 mN H ₂ SO ₄	0	40	130 °C	2	23	5
25	80 mN H ₂ SO ₄	0	80	130 °C	1	19	10
26	120 mN H ₂ SO ₄	0	120	130 °C	1	19	10
27	160 mN H ₂ SO ₄	0	160	130 °C	1	19	10

[0412] At 25°C, no difference was observed between the chemically treated samples and the control, and no emulsion formed.

[0413] At 65°C, an emulsion began to appear for the KOH treated samples, increasing in size with higher concentrations of caustic. The aqueous phase also became more turbid with increasing concentrations of KOH. There appeared to be a proportional decrease in the size of the pellet as the emulsion size increased. For the acidic treated samples at 65°C, a smaller emulsion formed relative to the caustic treated samples, and no increase in the size of the emulsion was observed with increasing level of acid, and the aqueous phase was clear and uncolored.

[0414] Samples heated to 130°C all showed significantly larger emulsions compared to 65°C, and comprised roughly 1/3 of the total sample volume. The aqueous phase of the caustic treated samples showed increasing dark color with increasing concentration of KOH.

[0415] In the acidic treated samples, the aqueous phase remained clear for all samples. In contrast to the base treated materials, the pellet virtually disappeared for the acid treated preparations, indicating that the cells and cell debris had partitioned to the emulsion and aqueous phase. In none of the 27 conditions tested was a visible oil layer observed.

(b) Emulsion Water Washing

[0416] The emulsion of the caustic treated samples was poorly defined and difficult to separate from the aqueous phase to prepare the emulsion for water washing. In contrast, the acidic emulsion was firm and partitioned easily from the aqueous phase. Under all washing conditions for the KOH treated samples, the emulsion became thinner and thinner until it was nearly gone. No visible oil layer was observed during either wash.

[0417] The acidic emulsions were easier to manipulate and separate from the aqueous phase. The emulsion size remained essentially unchanged during both washes. However, no visible oil layer was observed during either wash.

(c) **Sonication**

[0418] The results of the sonication are recorded in Table 25 below.

[0419] Table 25. Impact of sonication compared to corresponding pre-sonicated samples.

Tube	[KOH] mN	[H ₂ SO ₄] mN	Temp, 30 min	Pellet, mm	Emulsion, mm
1	0	0	25 °C	-5	+1
9	0	160	25 °C	-5	+1
10	0	0	65 °C	-5	+1
14	160	0	65 °C	-3	+2
18	0	160	65 °C	-5	+1
19	0	0	130 °C	-2	-7
23	160	0	130 °C	-3	-7
27	0	160	130 °C	+1	-7

[0420] The 25°C and 65°C sample emulsions increased in size while the pellet decreased relative to their pre-sonicated counterparts. For the 130°C samples, the emulsion decreased in size substantially, and the pellet of the control and base treated samples decreased, while the pellet increased for the acid treated preparation. There was some oil layer formation for the sonicated samples, particularly at the elevated temperatures.

(d) **Temperature Reduction of Emulsion**

[0421] The results of the emulsion freezing are recorded in Table 26 below.

[0422] Table 26. Results of emulsion freezing.

Sample	Sonicated	Visible oil layer
160 mN HCl, 130 °C for 30 min	Yes, 30 min cycle	Yes
120 mN HCl, 130 °C for 30 min	No	No
160 mN KOH 130 °C for 30 min	Yes, 30 min cycle	Yes
No chemical, 130 °C for 30 min	Yes, 30 min cycle	Yes
No chemical, no heat	Yes, 30 min cycle	Yes

[0423] Reducing the temperature of the emulsion to – 20°C proved to be an effective emulsion breaker. In all samples where the emulsion was sonicated, a large oil layer was observed. A representative result is shown in Figure 4. When compared to the lipid content

of the cell material by HPLC, the oil layer constituted the majority of the oil from the emulsion.

(e) Enzymatic Treatment #1

[0424] Samples incubated at 25°C and 37°C at 1% enzyme levels showed minimal emulsion formation. Considerably larger emulsions formed at the higher enzyme concentrations of 4%/2% Hemicellulase/Driselase combination, with the 50°C producing a larger emulsion compared to the 37°C sample. No oil layer was observed with enzyme treatment alone. Enzyme-generated emulsions subjected to sonication and then– 20°C, generated an oil layer.

(f) Enzymatic Treatment #2

[0425] Enzyme treatment combining a polysaccharide-degrading mannanase and a protease was found effective to disrupt cells from *Chlorella protothecoides*. Another technique found effective in weakening cell structure enough to facilitate oil extraction was storing/aging the cells. Tables 27, 28 and 29 show the results of the three experiments described above, and Figure 6 illustrates the resulting layers from enzyme-treated (left tube) vs. untreated (right tube) biomass. In the untreated tube, the cell pellet is considerably larger in size because it comprises a large lipid component. In neither case was lipid found in the aqueous phase.

[0426] Table 27. Enzymatic treatment #2, experiment 1 results.

Sample Number				
	1	2	3	4
ALCALASE 2.4 FG	+	+	-	-
MANNAWAY 4.0 L	+	-	+	-
% of oil in cell pellet	9.6%	30.7%	50.3%	97.8%
% of oil in emulsion	90.4%	69.3%	49.7%	2.2%

[0427] Table 28. Enzymatic treatment #2, experiment 2 results.

Sample Number					
	AGE	1	2	3	4
ALCALASE 2.4 FG		+	+	-	-
MANNAWAY 4.0 L		+	-	+	-
DCW G/L		170.0	170.0	170.0	170.0
	DAYS	W/W%	W/W%	W/W%	W/W%
% of oil in cell pellet	0	28.8%	48.2%	67.8%	100.0%
% of oil in emulsion	0	71.2%	51.8%	32.2%	0%
% of oil in cell pellet	7	47.8%	81.3%	80.7%	81.8%

% of oil in emulsion	7	52.2%	18.7%	19.3%	18.2%
----------------------	---	-------	-------	-------	-------

[0428] Table 29. Enzymatic treatment #2, experiment 3 results.

	Sample Number				
	1	2	3	4	5
ALCALASE 2.4 FG	-	-	+	-	-
MANNAWAY 4.0 L	-	-	+	-	-
DCW G/L	71.7	115.1	115.1	71.7	115.1
TEMPERATURE	50	50	50	50	50
PH	7.5	7.5	7.5	7.5	7.5
	W/W%	W/W%	W/W%	W/W%	W/W%
% of oil in cell pellet	81.8%	63.0%	12.0%	49.1%	42.0%
% of oil in emulsion	18.2%	37.0%	88.0%	50.9%	58.0%
age	0	0	0	7	7

(g) Oil Analysis and Characterization

[0429] An oil sample was isolated from each of the chemically treated conditions and analyzed via TLC to determine if product degradation was occurring at the extreme pHs compared the control condition. TLC analysis is shown in Figure 5. The pattern showed a minor generation of free fatty acids (FFA's) under acidic conditions compared to the base treated and control sample.

Example 8: Methods for Culturing *Prototheca*

[0430] *Prototheca* strains were cultivated to achieve a high percentage of oil by dry cell weight. Cryopreserved cells were thawed at room temperature and 500 ul of cells were added to 4.5 ml of medium (4.2 g/L K₂HPO₄, 3.1 g/L NaH₂PO₄, 0.24 g/L MgSO₄·7H₂O, 0.25 g/L Citric Acid monohydrate, 0.025 g/L CaCl₂ 2H₂O, 2g/L yeast extract) plus 2% glucose and grown for 7 days at 28°C with agitation (200 rpm) in a 6-well plate. Dry cell weights were determined by centrifuging 1 ml of culture at 14,000 rpm for 5 min in a pre-weighed Eppendorf tube. The culture supernatant was discarded and the resulting cell pellet washed with 1 ml of deionized water. The culture was again centrifuged, the supernatant discarded, and the cell pellets placed at -80°C until frozen. Samples were then lyophilized for 24 hrs and dry cell weights calculated. For determination of total lipid in cultures, 3 ml of culture was removed and subjected to analysis using an Ankom system (Ankom Inc., Macedon, NY) according to the manufacturer's protocol. Samples were subjected to solvent extraction with an Amkom XT10 extractor according to the manufacturer's protocol. Total lipid was

determined as the difference in mass between acid hydrolyzed dried samples and solvent extracted, dried samples. Percent oil dry cell weight measurements are shown in Table 30.

[0431] Table 30. Percent oil by dry cell weight.

Species	Strain	% Oil
Prototheca stagnora	UTEX 327	13.14
Prototheca moriformis	UTEX 1441	18.02
Prototheca moriformis	UTEX 1435	27.17

[0432] Microalgae samples from the strains listed in Table 30, above, were genotyped. Genomic DNA was isolated from algal biomass as follows. Cells (approximately 200 mg) were centrifuged from liquid cultures 5 minutes at 14,000 x g. Cells were then resuspended in sterile distilled water, centrifuged 5 minutes at 14,000 x g and the supernatant discarded. A single glass bead ~2mm in diameter was added to the biomass and tubes were placed at -80°C for at least 15 minutes. Samples were removed and 150 µl of grinding buffer (1% Sarkosyl, 0.25 M Sucrose, 50 mM NaCl, 20 mM EDTA, 100 mM Tris-HCl, pH 8.0, RNase A 0.5 ug/ul) was added. Pellets were resuspended by vortexing briefly, followed by the addition of 40 ul of 5M NaCl. Samples were vortexed briefly, followed by the addition of 66 µl of 5% CTAB (Cetyl trimethylammonium bromide) and a final brief vortex. Samples were next incubated at 65°C for 10 minutes after which they were centrifuged at 14,000 x g for 10 minutes. The supernatant was transferred to a fresh tube and extracted once with 300 µl of Phenol:Chloroform:Isoamyl alcohol 12:12:1, followed by centrifugation for 5 minutes at 14,000 x g. The resulting aqueous phase was transferred to a fresh tube containing 0.7 vol of isopropanol (~190 µl), mixed by inversion and incubated at room temperature for 30 minutes or overnight at 4°C. DNA was recovered via centrifugation at 14,000 x g for 10 minutes. The resulting pellet was then washed twice with 70% ethanol, followed by a final wash with 100% ethanol. Pellets were air dried for 20-30 minutes at room temperature followed by resuspension in 50 µl of 10mM TrisCl, 1mM EDTA (pH 8.0).

[0433] Five µl of total algal DNA, prepared as described above, was diluted 1:50 in 10 mM Tris, pH 8.0. PCR reactions, final volume 20 µl, were set up as follows. Ten µl of 2 x iProof HF master mix (BIO-RAD) was added to 0.4 µl primer SZ02613 (5'-TGTTGAAGAATGAGCCGGCGAC-3' (SEQ ID NO:5) at 10mM stock concentration). This primer sequence runs from position 567-588 in Gen Bank accession no. L43357 and is highly conserved in higher plants and algal plastid genomes. This was followed by the addition of 0.4 µl primer SZ02615 (5'-CAGTGAGCTATTACGCACTC-3' (SEQ ID NO:6) at 10 mM

stock concentration). This primer sequence is complementary to position 1112-1093 in Gen Bank accession no. L43357 and is highly conserved in higher plants and algal plastid genomes. Next, 5 µl of diluted total DNA and 3.2 µl dH₂O were added. PCR reactions were run as follows: 98°C, 45''; 98°C, 8''; 53°C, 12''; 72°C, 20'' for 35 cycles followed by 72°C for 1 min and holding at 25°C. For purification of PCR products, 20 µl of 10 mM Tris, pH 8.0, was added to each reaction, followed by extraction with 40 µl of Phenol:Chloroform:isoamyl alcohol 12:12:1, vortexing and centrifuging at 14,000 x g for 5 minutes. PCR reactions were applied to S-400 columns (GE Healthcare) and centrifuged for 2 minutes at 3,000 x g. Purified PCR products were subsequently TOPO cloned into PCR8/GW/TOPO and positive clones selected for on LB/Spec plates. Purified plasmid DNA was sequenced in both directions using M13 forward and reverse primers. In total, twelve *Prototheca* strains were selected to have their 23S rRNA DNA sequenced and the sequences are listed in the Sequence Listing. A summary of the strains and Sequence Listing Numbers is included below in Table 31. The sequences were analyzed for overall divergence from the UTEX 1435 (SEQ ID NO: 17) sequence. Two pairs emerged (UTEX 329/UTEX 1533 and UTEX 329/UTEX 1440) as the most divergent. In both cases, pairwise alignment resulted in 75.0% pairwise sequence identity. The percent sequence identity to UTEX 1435 is also included below in Table 31.

[0434] Table 31. Genotyped *Prototheca* strains.

Species	Strain	% nt identity	SEQ ID NO.
<i>Prototheca kruegani</i>	UTEX 329	75.2	SEQ ID NO: 34
<i>Prototheca wickerhamii</i>	UTEX 1440	99	SEQ ID NO: 35
<i>Prototheca stagnora</i>	UTEX 1442	75.7	SEQ ID NO: 16
<i>Prototheca moriformis</i>	UTEX 288	75.4	SEQ ID NO: 36
<i>Prototheca moriformis</i>	UTEX 1439; 1441; 1435; 1437	100	SEQ ID NO: 17
<i>Prototheca wickerhamii</i>	UTEX 1533	99.8	SEQ ID NO: 37
<i>Prototheca moriformis</i>	UTEX 1434	75.9	SEQ ID NO: 38
<i>Prototheca zopfii</i>	UTEX 1438	75.7	SEQ ID NO: 39
<i>Prototheca moriformis</i>	UTEX 1436	88.9	SEQ ID NO: 32

[0435] Lipid samples from a subset of the above-listed strains were analyzed for lipid profile using HPLC. Results are shown below in Table 32.

[0436] Table 32. Diversity of lipid chains in *Prototheca* species.

Strain	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1
UTEX 327	0	12.01	0	0	50.33	17.14	0	0	0
UTEX	1.41	29.44	0.70	3.05	57.72	12.37	0.97	0.33	0

1441									
UTEX	1.09	25.77	0	2.75	54.01	11.90	2.44	0	0
1435									

[0437] Algal plastid transit peptides were identified through the analysis of UTEX 1435 (*Prototheca moriformis*) or UTEX 250 (*Chlorella protothecoides*) cDNA libraries. cDNAs encoding potentially plastid targeted proteins based upon BLAST hit homology to other known plastid targeted proteins were subjected to further analysis by the software programs PSORT (<http://psort.ims.u-tokyo.ac.jp/form.html>), ChloroP (<http://www.cbs.dtu.dk/services/ChloroP/>) and TargetP (<http://www.cbs.dtu.dk/services/TargetP/>). Candidate plastid transit peptides identified through at least one of these three programs were then PCR amplified from the appropriate genomic DNA. Below, in Table 33, is a summary of the algal plastid targeting amino acid sequences (PTS) that were identified from this screen. Also included are the amino acid sequences of plant fatty acyl-ACP thioesterases that are used in the heterologous expression Examples below.

[0438] Table 33. Summary of algal plastic targeting amino acid sequences.

<u>cDNA</u>	<u>SEQ ID NO.</u>
<i>P.moriformis</i> isopentenyl diphosphate synthase PTS	SEQ ID NO: 40
<i>P.moriformis</i> delta 12 fatty acid desaturase PTS	SEQ ID NO: 41
<i>P.moriformis</i> stearoyl ACP desaturase PTS	SEQ ID NO: 42
<i>C.protothecoides</i> stearoyl ACP desaturase PTS	SEQ ID NO: 43
<i>Cuphea hookeriana</i> fatty acyl-ACP thioesterase (C8-10)	SEQ ID NO: 44
<i>Umbellularia californica</i> fatty acyl-ACP thioesterase (C12)	SEQ ID NO: 45
<i>Cinnamomum camphora</i> fatty acyl-ACP thioesterase (C14)	SEQ ID NO: 46

Example 9: Methods for Transforming *Prototheca*

1. General Method for Biolistic transformation of *Prototheca*

[0439] S550d gold carriers from Seashell Technology were prepared according to the protocol from manufacturer. Linearized plasmid (20 µg) was mixed with 50 µl of binding buffer and 60 µl (30 mg) of S550d gold carriers and incubated in ice for 1 min. Precipitation buffer (100 µl) was added, and the mixture was incubated in ice for another 1 min. After vortexing, DNA-coated particles were pelleted by spinning at 10,000 rpm in an Eppendorf 5415C microfuge for 10 seconds. The gold pellet was washed once with 500 µl of cold 100% ethanol, pelleted by brief spinning in the microfuge, and resuspended with 50 µl of ice-cold

ethanol. After a brief (1-2 sec) sonication, 10 μ l of DNA-coated particles were immediately transferred to the carrier membrane.

[0440] *Prototheca* strains were grown in proteose medium (2g/L yeast extract, 2.94mM NaNO₃, 0.17mM CaCl₂•2H₂O, 0.3mM MgSO₄•7H₂O, 0.4mM K₂HPO₄, 1.28mM KH₂PO₄, 0.43mM NaCl) on a gyratory shaker until it reaches a cell density of 2×10^6 cells/ml. The cells were harvested, washed once with sterile distilled water, and resuspended in 50 μ l of medium. 1×10^7 cells were spread in the center third of a non-selective proteose media plate. The cells were bombarded with the PDS-1000/He Biolistic Particle Delivery system (Bio-Rad). Rupture disks (1100 and 1350 psi) were used, and the plates are placed 9 and 12 cm below the screen/macrocarrrier assembly. The cells were allowed to recover at 25°C for 12–24 h. Upon recovery, the cells were scraped from the plates with a rubber spatula, mixed with 100 μ l of medium and spread on plates containing the appropriate antibiotic selection. After 7-10 days of incubation at 25°C, colonies representing transformed cells were visible on the plates from 1100 and 1350 psi rupture discs and from 9 and 12 cm distances. Colonies were picked and spotted on selective agar plates for a second round of selection.

2. Transformation of *Prototheca* with G418 Resistance Gene

[0441] *Prototheca moriformis* and other *Prototheca* strains sensitive to G418 can be transformed using the methods described below. G418 is an aminoglycoside antibiotic that inhibits the function of 80S ribosomes and thereby inhibits protein synthesis. The corresponding resistance gene functions through phosphorylation, resulting in inactivation of G418. *Prototheca* strains UTEX 1435, UTEX 1439 and UTEX 1437 were selected for transformation. All three *Prototheca* strains were genotyped using the methods described above. All three *Prototheca* strains had identical 23s rRNA genomic sequences (SEQ ID NO:17).

[0442] All transformation cassettes were cloned as EcoRI-SacI fragments into pUC19. Standard molecular biology techniques were used in the construction of all vectors according to Sambrook and Russell, 2001. The *C. reinhardtii* beta-tubulin promoter/5'UTR was obtained from plasmid pHyg3 (Berthold *et al.*, (2002) Protist: 153(4), pp 401-412) by PCR as an EcoRI-AscI fragment. The *Chlorella vulgaris* nitrate reductase 3'UTR was obtained from genomic DNA isolated from UTEX strain 1803 via PCR using the following primer pairs:

Forward:

5' TGACCTAGGTGATTAATTAAGTTCGAGGCAGCAGCAGCTCGGATAGTATCG 3'
(SEQ ID NO:47)

Reverse:

5' CTACGAGCTCAAGCTTTCCATTTGTGTTC CCATCCCCTACTTCC 3'
(SEQ ID NO:48)

[0443] The *Chlorella sorokiniana* glutamate dehydrogenase promoter/UTR was obtained via PCR of genomic DNA isolated from UTEX strain 1230 via PCR using the following primer pairs:

Forward: 5' GATCAGAATTCCGCCTGCAACGCAAGG GCAGC 3' (SEQ ID NO:34)

Reverse: 5' GCATACTAGTGGCGGGACGGAGAGA GGGCG 3' (SEQ ID NO:35)

[0444] Codon optimization was based on the codons for *Prototheca moriformis*. The sequence of the non-codon optimized neomycin phosphotransferase (nptII) cassette was synthesized as an AscI-XhoI fragment and was based on upon the sequence of Genbank Accession No. YP_788126. The codon optimized nptII cassette was also based on this Genbank Accession number.

[0445] The three *Prototheca* strains were transformed using biolistic methods described above. Briefly, the *Prototheca* strains were grown heterophically in liquid medium containing 2% glucose until they reached the desired cell density (1×10^7 cells/mL to 5×10^7 cells/mL). The cells were harvested, washed once with sterile distilled water and resuspended at 1×10^8 cells/mL. 0.5mL of cells were then spread out on a non-selective solid media plate and allowed to dry in a sterile hood. The cells were bombarded with the PDS-1000/He Biolistic Particle Delivery System (BioRad). The cells were allowed to recover at 25° C for 24 hours. Upon recovery, the cells were removed by washing plates with 1 mL of sterile media and transferring to fresh plates containing 100 µg/mL G418. Cells were allowed to dry in a sterile hood and colonies were allowed to form on the plate at room temperature for up to three weeks. Colonies of UTEX 1435, UTEX 1439 and UTEX 1437 were picked and spotted on selective agar plates for a second round of selection.

[0446] A subset of colonies that survived a second round of selection described above, were cultured in small volume and genomic DNA and RNA were extracted using standard molecular biology methods. Southern blots were done on genomic DNA extracted from untransformed (WT), the transformants and plasmid DNA. DNA from each sample was run on 0.8% agarose gels after the following treatments: undigested (U), digested with AvrII (A),

digested with NcoI (N), digested with SacI (S). DNA from these gels was blotted on Nylon+ membranes (Amersham). These membranes were probed with a fragment corresponding to the entire coding region of the nptII gene (NeoR probe). Figure 7 shows maps of the cassettes used in the transformations. Figure 8 shows the results of Southern blot analysis on three transformants (all generated in UTEX strain 1435) (1, 2, and 3) transformed with either the beta-tubulin::neo::nit (SEQ ID NO: 49) (transformants 1 and 2) or glutamate dehydrogenase:neo::nit (SEQ ID NO: 50) (transformant 3). The glutamate dehydrogenase:neo::nit transforming plasmid was run as a control and cut with both NcoI and SacI. AvrII does not cut in this plasmid. Genomic DNA isolated from untransformed UTEX strain 1435 shows no hybridization to the NeoR probe.

[0447] Additional transformants containing the codon-optimized glutamate dehydrogenase:neo::nit (SEQ ID NO: 51) and codon-optimized β -tubulin::neo::nit (SEQ ID NO:52) constructs were picked and analyzed by Southern blot analysis. As expected, only digests with SacI show linearization of the transforming DNA. These transformation events are consistent with integration events that occur in the form of oligomers of the transforming plasmid. Only upon digestion with restriction enzymes that cut within the transforming plasmid DNA do these molecules collapse down the size of the transforming plasmid.

[0448] Southern blot analysis was also performed on transformants generated upon transformation of *Prototheca* strains UTEX 1437 and UTEX 1439 with the glutamate dehydrogenase::neo::nit cassette. The blot was probed with the NeoR probe and the results are similar to the UTEX 1435 transformants. The results are indicative of integration events characterized by oligomerization and integration of the transforming plasmid. This type of integration event is known to occur quite commonly in *Dictyostelium discoideum* (see, for example, Kuspa, A. and Loomis, W. (1992) *PNAS*, 89:8803-8807 and Morio *et al.*, (1995) *J. Plant Res.* 108:111-114).

[0449] To further confirm expression of the transforming plasmid, Northern blot analysis and RT-PCR analysis were performed on selected transformants. RNA extraction was performed using Trizol Reagent according to manufacturer's instructions. Northern blot analysis were run according to methods published in Sambrook and Russel, 2001. Total RNA (15 μ g) isolated from five UTEX 1435 transformants and untransformed UTEX 1435 (control lanes) was separated on 1% agarose-formaldehyde gel and blotted on nylon membrane. The blot was hybridized to the neo-non-optimized probe specific for transgene sequences in

transformants 1 and 3. The two other transformants RNAs express the codon-optimized version of the neo-transgene and, as expected, based on the sequence homology between the optimized and non-optimized neo genes, showed significantly lower hybridization signal.

[0450] RNA (1 µg) was extracted from untransformed *Prototheca* strain UTEX 1435 and two representative UTEX 1435 transformants and reverse transcribed using an oligo dT primer or a gene specific primer. Subsequently these cDNAs (in duplicate) were subjected to qPCR analysis on ABI Veriti Thermocycler using SYBR-Green qPCR chemistry using the following primers (nptII):

Forward: 5' GCCGCGACTGGCTGCTGCTGG 3' (SEQ ID NO:53)

Reverse: 5' AGGTCCTCGCCGTCGGGCATG 3' (SEQ ID NO:54)

[0451] Possible genomic DNA contamination was ruled out by a no reverse transcriptase negative control sample. The results indicated that the NeoR genes used to transform these strains is actively transcribed in the transformants.

3. Transformation of *Prototheca* with Secreted Heterologous Sucrose Invertase

[0452] All of the following experiments were performed using liquid medium/agar plates based on the basal medium described in Ueno et al., (2002) *J Bioscience and Bioengineering* 94(2):160-65, with the addition of trace minerals described in US Patent No. 5,900,370, and 1x DAS Vitamin Cocktail (1000x solution): tricine: 9g, thiamine HCL: 0.67g, biotin: 0.01g, cyannocobalamin (vitamin B12): 0.008g, calcium pantothenate: 0.02g and p-aminobenzoic acid: 0.04g).

[0453] Two plasmid constructs were assembled using standard recombinant DNA techniques. The yeast sucrose invertase genes (one codon optimized and one non-codon optimized), *suc2*, were under the control of the *Chlorella reinhardtii* beta-tubulin promoter/5'UTR and had the *Chlorella vulgaris* nitrate reductase 3'UTR. The sequences (including the 5'UTR and 3'UTR sequences) for the non-codon optimized (Crβ-tub::NCO-suc2::CvNitRed) construct, SEQ ID NO: 55, and codon optimized (Crβ-tub::CO-suc2::CvNitRed) construct, SEQ ID NO: 56, are listed in the Sequence Listing. Codon optimization was based on *Prototheca* sp (see Table 4). Figure 9 shows a schematic of the two constructs with the relevant restriction cloning sites and arrows indicating the direction of transcription. Selection was provided by Neo R. Codon optimization was based on preferred codon usage in *Prototheca* strains in Table 4.

[0454] Preparation of the DNA/gold microcarrier: DNA/gold microcarriers were prepared immediately before use and stored on ice until applied to macrocarriers. The plasmid DNA (in TE buffer) was added to 50 μ l of binding buffer. Saturation of the gold beads was achieved at 15 μ g plasmid DNA for 3 mg gold carrier. The binding buffer and DNA were mixed well via vortexing. The DNA and binding buffer should be pre-mix prior to gold addition to ensure uniformed plasmid binding to gold carrier particles. 60 μ l of S550d (Seashell Technologies, San Diego, CA) gold carrier was added to the DNA/binding buffer mixture. For a gold stock at 50 mg/ml, addition of 60 μ l results in an optimal ratio of 15 μ g DNA/3 mg gold carrier. The gold carrier/DNA mixture was allowed to incubate on ice for 1 minute and then 100 μ l of precipitation buffer was added. The mixture was allowed to incubate again on ice for 1 minute and then briefly vortexed and centrifuged at 10,000 rpm at room temperature for 10 seconds to pellet the gold carrier. The supernatant was carefully removed with a pipette and the pellet was washed with 500 μ l of ice cold 100% ethanol. The gold particles were re-pelleted by centrifuging again at 10,000 rpm for 10 seconds. The ethanol was removed and 50 μ l of ice cold ethanol was added to the gold mixture. Immediately prior to applying the gold to macrocarriers, the gold/ethanol was resuspended with a brief 1-2 second pulse at level 2 on a MISONIX sonicator using the micro tip. Immediately after resuspension, 10 μ l of the dispersed gold particles was transferred to the macrocarrier and allowed to dry in a sterile hood.

[0455] The two *Prototheca moriformis* strains (UTEX 1435 and 1441) were grown heterotrophically in liquid medium containing 2% glucose from cryopreserved vials. Each strain was grown to a density of 10^7 cells/ml. This seed culture was then diluted with fresh media to a density of 10^5 cells/ml and allowed to grow for 12-15 hours to achieve a final cell density of approximately 10^6 cells/ml. The microalgae were aliquoted into 50 ml conical tubes and centrifuged for 10 minutes at 3500 rpm. The cells were washed with fresh medium and centrifuged again for 10 minutes at 3500 rpm. The cells were then resuspended at a density of 1.25×10^8 cells/ml in fresh medium.

[0456] In a sterile hood, 0.4ml of the above-prepared cells were removed and placed directly in the center of an agar plate (without selection agent). The plate was gently swirled with a level circular motion to evenly distribute the cells to a diameter of no more than 3cm. The cells were allowed to dry onto the plates in the sterile hood for approximately 30-40 minutes and then were bombarded at a rupture disk pressure of 1350 psi and a plate to

macrocarrier distance of 6 cm. The plates were then covered and wrapped with parafilm and allowed to incubate under low light for 24 hours.

[0457] After the 24 hour recovery, 1ml of sterile medium (with no glucose) was added to the lawn of cells. The cells were resuspended using a sterile loop, applied in a circular motion to the lawn of cells and the resuspended cells were collected using a sterile pipette. The cells were then plated onto a fresh agar plate with 2% glucose and 100 µg/ml G418. The appearance of colonies occurred 7-12 days after plating. Individual colonies were picked and grown in selective medium with 2% glucose and 100 µg/ml G418. The wildtype (untransformed) and transgenic cells were then analyzed for successful introduction, integration and expression of the transgene.

[0458] Genomic DNA from transformed *Prototheca moriformis* UTEX 1435 and 1441 and their wildtype (untransformed) counterparts were isolated using standard methods. Briefly, the cells were centrifuged for 5 minutes at 14,000 rpm in a standard table top Eppendorf centrifuge (model 5418) and flash frozen prior to DNA extraction. Cell pellets were lysed by adding 200 µL of Lysis buffer (100 mM Tris HCl, pH 8.0, 1% Lauryl Sarcosine, 50 mM NaCl, 20 mM EDTA, 0.25 M sucrose, 0.5 mg/ml RNase A) for every 100-200 mg of cells (wet weight) and vortexing for 30-60 seconds. Cetyl trimethylammonium bromide (CTAB) and NaCl were brought to 1% and 1 M, respectively, and cell extracts were incubated at 60-65°C for 10 minutes. Subsequently, extracts were clarified via centrifugation at 14,000 rpm for 10 minutes and the resulting supernatant was extracted with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1). Samples were then centrifuged for 5 minutes at 14,000 rpm and the aqueous phase removed. DNA was precipitated with 0.7 volumes of isopropanol. DNA was pelleted via centrifugation at 14,000 rpm for 10 minutes and washed twice with 80% ethanol, and once with ethanol. After drying, DNA was resuspended in 10 mM Tris HCl, pH 8.0 and DNA concentrations were determined by using PicoGreen fluorescence quantification assay (Molecular Probes).

[0459] RNA from transformed *Prototheca moriformis* UTEX 1435 and 1441 and their wildtype (untransformed) counterparts were isolated using standard methods. Briefly, the cells were centrifuged for 5 minutes at 14,000 rpm in a standard table top Eppendorf centrifuge (model 5418) and flash frozen before RNA extraction. Cell pellets were lysed by addition of 1 mL of Trizol reagent (Sigma) for every 100 mg of cells (wet weight) and by vortexing for 1-2 minutes. Samples were incubated at room temperature for 5 minutes and

subsequently adjusted with 200 μ L of chloroform per 1 mL of Trizol reagent. After extensive shaking, cells were incubated at room temperature for 15 minutes and then subjected to centrifugation at 14000 rpm for 15 minutes in a refrigerated table top microcentrifuge. RNA partitioning to the upper aqueous phase was removed and precipitated by addition of isopropanol (500 μ L per 1ml of Trizol reagent). RNA was collected by centrifugation for 10 minutes and the resulting pellet washed twice with 1 mL of 80% ethanol, dried, and resuspended in RNase free water. RNA concentration was estimated by RiboGreen fluorescence quantification assay (Molecular Probes).

[0460] Expression of neomycin phosphotransferase gene conferring G418 antibiotic resistance and yeast invertase was assayed in non-transformed *Prototheca moriformis* UTEX 1435 and 1441 and transformants T98 (UTEX 1435 transformant) and T97 (UTEX 1441 transformant) using reverse transcription quantitative PCR analysis (RT-qPCR). 20 ng total RNA (isolated as described above) was subjected to one step RT-qPCR analysis using iScript SYBR Green RT-PCR kit (BioRad Laboratories) and primer pairs targeting the neomycin resistance gene (forward primer 5'CCGCCGTGCTGGACGTGGTG 3' and reverse primer 5'GGTGGCGGGGTCCAGGGTGT 3'; SEQ ID NOs: 57 and 58, respectively) and *suc2* invertase transcripts (forward primer 5' CGGCCGGCGGCTCCTTCAAC 3' and reverse primer 5' GGCGTCCCGTAGGTCGGGT 3'; SEQ ID NO: 59 and 60, respectively). Endogenous beta-tubulin transcripts served as an internal positive control for PCR amplification and as a normalization reference to estimate relative transcript levels.

[0461] Both codon optimized and non-codon optimized constructs were transformed into UTEX 1435 and 1441 *Prototheca moriformis* cells as described above. Initially, transformants were obtained with both constructs and the presence of the transgene was verified by Southern blot analysis followed by RTPCR to confirm the presence of the DNA and mRNA from the transgene. For the Southern blot analysis, genomic DNA isolated as described above was electrophoresed on 0.7% agarose gels in 1x TAE buffer. Gells were processed as described in Sambrook et al. (Molecular Cloning; A Laboratory Manual, 2nd Edition. Cold Spring Harbor Laboratory Press, 1989). Probes were prepared by random priming and hybridizations carried out as described in Sambrook et al. Transformants from both the codon optimized and the non-codon optimized constructs showed the presence of the invertase cassette, while the non-transformed control was negative. Invertase mRNA was also detected in transformants with both the codon optimized and non-codon optimized constructs.

[0462] To confirm that the transformants were expressing an active invertase protein, the transformants were plated on sucrose plates. The transformants containing the non-codon optimized cassette failed to grow on the sucrose containing plates, indicating that, while the gene and the mRNA encoding the SUC2 protein were present, the protein was either (1) not being translated, or (2) being translated, but not accumulating to levels sufficient to allow for growth on sucrose as the sole carbon source. The transformants with the codon optimized cassette grew on the sucrose containing plates. To assess the levels of invertase being expressed by these transformants, two clones (T98 and T97) were subjected to an invertase assay of whole cells scraped from solid medium and direct sampling and quantitation of sugars in the culture supernatants after 48 hours of growth in liquid medium containing 2% sucrose as the sole carbon source.

[0463] For the invertase assay, the cells (T98 and T97) were grown on plates containing 2% sucrose, scraped off and assayed for invertase activity. 10 μ l of the scraped cells was mixed with 40 μ l of 50mM NaOAc pH 5.1. 12.5 μ l of 0.5M sucrose was added to the cell mixture and incubated at 37°C for 10-30 minutes. To stop the reaction, 75 μ l of 0.2M K_2HPO_4 was added. To assay for glucose liberated, 500 μ l of reconstituted reagent (glucose oxidase/peroxidase + o-Dianisidine) from Sigma (GAGO-20 assay kit) was added to each tube and incubated at 37°C for 30 minutes. A glucose standard curve was also created at this time (range: 25 μ g to 0.3 μ g glucose). After incubation, 500 μ l of 6N HCl was added to stop the reaction and to develop the color. The samples were read at 540nm. The amount of glucose liberated was calculated from the glucose standard curve using the formula $y=mx+c$, where y is the 540nm reading, and x is μ g of glucose. Weight of glucose was converted to moles of glucose, and given the equimolar relationship between moles of sucrose hydrolyzed to moles of glucose generated, the data was expressed as nmoles of sucrose hydrolyzed per unit time. The assay showed that both T98 and T97 clones were able to hydrolyze sucrose, indicating that a functional sucrose invertase was being produced and secreted by the cells.

[0464] For the sugar analysis on liquid culture media after 48 hours of algal growth, T97 and T98 cells were grown in 2% sucrose containing medium for 48 hours and the culture media were processed for sugar analysis. Culture broths from each transformant (and negative non-transformed cell control) were centrifuged at 14,000 rpm for 5 minutes. The resulting supernatant was removed and subjected to HPLC/ELSD (evaporative light scattering detection). The amount of sugar in each sample was determined using external standards and liner regression analysis. The sucrose levels in the culture media of the

transformants were very low (less than 1.2g/L, and in most cases 0g/L). In the negative controls, the sucrose levels remained high, at approximately 19g/L after 48 hours of growth.

[0465] These results were consistent with the invertase activity results, and taken together, indicated that the codon optimized transformants, T97 and T98, secreted an active sucrose invertase that allowed the microalgae to utilize sucrose as the sole carbon source in contrast to (1) the non-codon optimized transformants and (2) the non-transformed wildtype microalgae, both of which could not utilize sucrose as the sole carbon source in the culture medium.

[0466] *Prototheca moriformis* strains, T98 and T97, expressing a functional, secreted sucrose invertase (SUC2) transgene were assayed for growth and lipid production using sucrose as the sole carbon source.

[0467] Wild type (untransformed), T98 and T97 strains were grown in growth media (as described above) containing either 4% glucose or 4% sucrose as the sole carbon source under heterotrophic conditions for approximately 6 days. Growth, as determined by A750 optical density readings were taken of all four samples every 24 hours and the dry cell weight of the cultures and lipid profiles were determined after the 6 days of growth. The optical density readings of the transgenic strains grown in both the glucose and sucrose conditions were comparable to the wildtype strains grown in the glucose conditions. These results indicate that the transgenic strains were able to grow on either glucose or sucrose as the sole carbon source at a rate equal to wildtype strains in glucose conditions. The non-transformed, wildtype strains did not grow in the sucrose-only condition.

[0468] The biomass for the wildtype strain grown on glucose and T98 strain grown on sucrose was analyzed for lipid profile. Lipid samples were prepared from dried biomass (lyophilized) using an Acid Hydrolysis System (Ankom Technology, NY) according to manufacturer's instructions. Lipid profile determinations were carried as described in Example 6. The lipid profile for the non-transformed *Prototheca moriformis* UTEX 1435 strain, grown on glucose as the sole carbon source and two clonal T98 strains (UTEX 1435 transformed with a sucrose invertase transgene), grown on sucrose as the sole carbon source, are disclosed in Table 34 (wildtype UTEX 1435 and T98 clone 8 and clone 11 below. C:19:0 lipid was used as an internal calibration control).

[0469] Table 34. Lipid profile of wildtype UTEX 1435 and UTEX 1435 clones with *suc2* transgene.

Name	wildtype (Area % - ISTD)	T98 clone 11 (Area % - ISTD)	T98 clone 8 (Area % - ISTD)
C 12:0	0.05	0.05	0.05
C 14:0	1.66	1.51	1.48
C 14:1	0.04	nd	nd
C 15:0	0.05	0.05	0.04
C 16:0	27.27	26.39	26.50
C 16:1	0.86	0.80	0.84
C 17:0	0.15	0.18	0.14
C 17:1	0.05	0.07	0.05
C 18:0	3.35	4.37	4.50
C 18:1	53.05	54.48	54.50
C 18:2	11.79	10.33	10.24
C 19:0 (ISTD)	--	--	--
C 18:3 alpha	0.90	0.84	0.81
C 20:0	0.32	0.40	0.38
C 20:1	0.10	0.13	0.12
C 20:1	0.04	0.05	0.04
C 22:0	0.12	0.16	0.12
C 20:3	0.07	0.08	0.07
C 24:0	0.12	0.11	0.10
nd - denotes none detected			

[0470] Oil extracted from wildtype *Prototheca moriformis* UTEX 1435 (via solvent extraction or using an expeller press) was analyzed for carotenoids, chlorophyll, tocopherols, other sterols and tocotrienols. The results are summarized below in Table 35.

[0471] Table 35. Carotenoid, chlorophyll, tocopherol/sterols and tocotrienol analysis in oil extracted from *Prototheca moriformis* (UTEX 1435).

	Pressed oil (mcg/ml)	Solvent extracted oil (mcg/ml)
cis-Lutein	0.041	0.042
trans-Lutein	0.140	0.112
trans-Zeaxanthin	0.045	0.039
cis-Zeaxanthin	0.007	0.013
t-alpha-Cryptoxanthin	0.007	0.010
t-beta-Cryptoxanthin	0.009	0.010
t-alpha-Carotene	0.003	0.001
c-alpha-Carotene	none detected	none detected
t-beta-Carotene	0.010	0.009
9-cis-beta-Carotene	0.004	0.002
Lycopene	none detected	none detected
Total Carotenoids	0.267	0.238
Chlorophyll	<0.01 mg/kg	<0.01 mg/kg
Tocopherols and Sterols		
	Pressed oil (mg/100g)	Solvent extracted oil (mg/100g)
gamma Tocopherol	0.49	0.49
Campesterol	6.09	6.05
Stigmasterol	47.6	47.8
Beta-sitosterol	11.6	11.5
Other sterols	445	446

Tocotrienols		
	Pressed oil (mg/g)	Solvent extracted oil (mg/g)
alpha Tocotrienol	0.26	0.26
beta Tocotrienol	<0.01	<0.01
gamma Tocotrienol	0.10	0.10
delta Tocotrienol	<0.01	<0.01
Total Tocotrienols	0.36	0.36

[0472] The ability of using sucrose as the sole carbon source as the selection factor for clones containing the *suc2* transgene construct instead of G418 (or another antibiotic) was assessed using the positive *suc2* gene transformants. A subset of the positive transformants was grown on plates containing sucrose as the sole carbon source and without antibiotic selection for 24 doublings. The clones were then challenged with plates containing glucose as the sole carbon source and G418. There was a subset of clones that did not grow on the glucose + G418 condition, indicating a loss of expression of the transgene. An additional experiment was performed using a plate containing sucrose as the sole carbon source and no G418 and streaking out a *suc2* transgene expressing clone on one half of the plate and wild-type *Prototheca moriformis* on the other half of the plate. Growth was seen with both the wild-type and transgene-containing *Prototheca moriformis* cells. Wild-type *Prototheca moriformis* has not demonstrated the ability to grow on sucrose, therefore, this result shows that unlike antibiotic resistance, the use of sucrose/invertase selection is not cell-autonomous. It is very likely that the transformants were secreting enough sucrose invertase into the plate/media to support wildtype growth as the sucrose was hydrolyzed into fructose and glucose.

Example 10: Recombinant *Prototheca* with Exogenous TE Gene

[0473] As described above, *Prototheca* strains can be transformed with exogenous genes. *Prototheca moriformis* (UTEX 1435) was transformed, using methods described above, with either *Umbellularia californica* C12 thioesterase gene or *Cinnamomum camphora* C14 thioesterase gene (both codon optimized according to Table 4). Each of the transformation constructs contained a *Chlorella sorokiniana* glutamate dehydrogenase promoter/5'UTR region (SEQ ID NO: 61) to drive expression of the thioesterase transgene. The thioesterase transgenes coding regions of *Umbellularia californica* C12 thioesterase (SEQ ID NO: 62) or *Cinnamomum camphora* C14 thioesterase (SEQ ID NO: 63), each with the native putative plastid targeting sequence. Immediately following the thioesterase coding sequence is the

coding sequence for a c-terminal 3x-FLAG tag (SEQ ID NO: 64), followed by the *Chlorella vulgaris* nitrate reductase 3'UTR (SEQ ID NO: 65).

[0474] Preparation of the DNA, gold microcarrier and *Prototheca moriformis* (UTEX 1435) cells were performed using the methods described above in Example 9. The microalgae were bombarded using the gold microcarrier—DNA mixture and plated on selection plates containing 2% glucose and 100µg/ml G418. The colonies were allowed to develop for 7 to 12 days and colonies were picked from each transformation plate and screened for DNA construct incorporation using Southern blots assays and expression of the thioesterase constructs were screened using RT-PCR.

[0475] Positive clones were picked from both the C12 and C14 thioesterase transformation plates and screened for construct incorporation using Southern blot assays. Southern blot assays were carried out using standard methods (and described above in Example 9) using an optimized c probes, based on the sequence in SEQ ID NO: 62 and SEQ ID NO: 63.

Transforming plasmid DNA was run as a positive control. Out of the clones that were positive for construct incorporation, a subset was selected for reverse transcription quantitative PCR (RT-qPCR) analysis for C12 thioesterase and C14 thioesterase expression.

[0476] RNA isolation was performed using methods described in Example 9 above and RT-qPCR of the positive clones were performed using 20 ng of total RNA from each clone using the below-described primer pair and iScript SYBR Green RT-PCR kit (Bio-Rad Laboratories) according to manufacturer's protocol. Wildtype (non-transformed) *Prototheca moriformis* total RNA was included as a negative control. mRNA expression was expressed as relative fold expression (RFE) as compared to negative control. . The primers that were used in the C12 thioesterase transformation RT-qPCR screening were:

U. californica C12 thioesterase PCR primers:

Forward: 5' CTGGGCGACGGCTTCGGCAC 3' (SEQ ID NO: 66)

Reverse: 5' AAGTCGCGGCGCATGCCGTT 3' (SEQ ID NO: 67)

[0477] The primers that were used in the C14 thioesterase transformation RT-qPCR screening were:

Cinnamomum camphora C14 thioesterase PCR primers:

Forward: 5' TACCCCGCCTGGGGCGACAC 3' (SEQ ID NO: 68)

Reverse: 5' CTTGCTCAGGCGGCGGGTGC 3' (SEQ ID NO: 69)

[0478] RT-qPCR results for C12 thioesterase expression in the positive clones showed an increased RFE of about 40 fold to over 2000 fold increased expression as compared to negative control. Similar results were seen with C14 thioesterase expression in the positive clones with an increase RFE of about 60-fold to over 1200 fold increased expression as compared to negative control.

[0479] A subset of the positive clones from each transformation (as screened by Southern blotting and RT-qPCR assays) were selected and grown under nitrogen-replete conditions and analyzed for total lipid production and profile. Lipid samples were prepared from dried biomass from each clone. 20-40 mg of dried biomass from each transgenic clone was resuspended in 2 mL of 3% H₂SO₄ in MeOH, and 200ul of toluene containing an appropriate amount of a suitable internal standard (C19:0) was added. The mixture was sonicated briefly to disperse the biomass, then heated at 65-70°C for two hours. 2 mL of heptane was added to extract the fatty acid methyl esters, followed by addition of 2 mL of 6% K₂CO₃ (aq) to neutralize the acid. The mixture was agitated vigorously, and a portion of the upper layer was transferred to a vial containing Na₂SO₄ (anhydrous) for gas chromatography analysis using standard FAME GC/FID (fatty acid methyl ester gas chromatography flame ionization detection) methods. Lipid profile (expressed as Area %) of the positive clones as compared to wildtype negative control are summarized in Tables 36 and 37 below. As shown in Table 36, the fold increase of C12 production in the C12 transformants ranged from about a 5-fold increase (clone C12-5) to over 11-fold increase (clone C12-1). Fold increase of C14 production in the C14 transformants ranged from about a 1.5 fold increase to about a 2.5 fold increase.

[0480] Table 36. Summary of total lipid profile of the *Prototheca moriformis* C12 thioesterase transformants.

	Wildtype	C12-1	C12-2	C12-3	C12-4	C12-5	C12-6	C12-7	C12-8
C6:0	0.03	nd	nd	nd	nd	nd	nd	nd	nd
C8:0	0.11	0.09	nd	0.11	nd	nd	nd	nd	nd
C10:0	nd	nd	nd	0.01	0.01	nd	nd	0.01	nd
C12:0	0.09	1.04	0.27	0.72	0.71	0.50	0.67	0.61	0.92
C14:0	2.77	2.68	2.84	2.68	2.65	2.79	2.73	2.56	2.69
C14:1	0.01	nd	nd	0.02	nd	nd	nd	0.01	nd
C15:0	0.30	0.09	0.10	0.54	0.19	0.09	0.13	0.97	0.09
C15:1	0.05	nd	nd	0.02	nd	nd	nd	nd	nd
C16:0	24.13	23.12	24.06	22.91	22.85	23.61	23.14	21.90	23.18
C16:1	0.57	0.62	0.10	0.52	0.69	0.63	0.69	0.49	0.63
C17:0	0.47	0.24	0.27	1.02	0.36	0.17	0.26	2.21	0.19
C17:1	0.08	nd	0.09	0.27	0.10	0.05	0.09	0.80	0.05
C18:0	nd	nd	2.14	1.75	2.23	2.16	2.38	1.62	2.47

C18:1	22.10	23.15	24.61	21.90	23.52	19.30	22.95	20.22	22.85
C18:1	nd	0.33	0.24	nd	nd	0.09	0.09	nd	0.11
C18:2	37.16	34.71	35.29	35.44	35.24	36.29	35.54	36.01	35.31
C18:3 alpha	11.68	11.29	9.26	11.62	10.76	13.61	10.64	11.97	10.81
C20:0	0.15	0.16	0.19	0.16	0.16	0.14	0.18	0.14	0.18
C20:1	0.22	0.17	0.19	0.20	0.21	0.19	0.21	0.20	0.21
C20:2	0.05	nd	0.04	0.05	0.05	0.05	0.04	0.05	0.04
C22:0	nd	nd	nd	0.01	nd	nd	nd	0.02	nd
C22:1	nd	nd	nd	nd	nd	0.01	nd	0.01	nd
C20:3	0.05	nd	0.07	0.06	0.06	0.10	0.07	0.05	0.06
C20:4	nd	nd	nd	nd	nd	0.02	nd	nd	nd
C24:0	nd	nd	0.24	0.01	0.20	0.19	0.19	0.14	0.20

[0481] Table 37. Summary of total lipid profile of the *Prototheca moriformis* C14 thioesterase transformants.

	Wildtype	C14-1	C14-2	C14-3	C14-4	C14-5	C14-6	C14-7
C6:0	0.03	nd	nd	nd	nd	nd	nd	nd
C8:0	0.11	nd	nd	nd	nd	nd	nd	nd
C10:0	nd	0.01	nd	0.01	nd	0.01	nd	nd
C12:0	0.09	0.20	0.16	0.25	0.21	0.19	0.40	0.17
C14:0	2.77	4.31	4.76	4.94	4.66	4.30	6.75	4.02
C14:1	0.01	nd	0.01	nd	nd	0.01	nd	nd
C15:0	0.30	0.43	0.45	0.12	0.09	0.67	0.10	0.33
C15:1	0.05	nd	nd	nd	nd	nd	nd	nd
C16:0	24.13	22.85	23.20	23.83	23.84	23.48	24.04	23.34
C16:1	0.57	0.65	0.61	0.60	0.60	0.47	0.56	0.67
C17:0	0.47	0.77	0.76	0.21	0.19	1.11	0.18	0.54
C17:1	0.08	0.23	0.15	0.06	0.05	0.24	0.05	0.12
C18:0	nd	1.96	1.46	2.48	2.34	1.84	2.50	2.06
C18:1	22.10	22.25	19.92	22.36	20.57	19.50	20.63	22.03
C18:1	nd	nd	nd	nd	nd	nd	0.10	nd
C18:2	37.16	34.97	36.11	34.35	35.70	35.49	34.03	35.60
C18:3 alpha	11.68	10.71	12.00	10.15	11.03	12.08	9.98	10.47
C20:0	0.15	0.16	0.19	0.17	0.17	0.14	0.18	0.16
C20:1	0.22	0.20	0.12	0.19	0.19	0.19	0.17	0.20
C20:2	0.05	0.04	0.02	0.03	0.04	0.05	0.03	0.04
C22:0	nd	nd	nd	nd	0.02	0.01	nd	nd
C22:1	nd	0.01	nd	nd	nd	nd	nd	0.01
C20:3	0.05	0.08	0.03	0.06	0.09	0.05	0.05	0.07
C20:4	nd	0.01	nd	nd	nd	nd	0.02	nd
C24:0	nd	0.17	0.14	0.19	0.20	0.16	0.22	0.17

[0482] The above-described experiments indicate the successful transformation of *Prototheca moriformis* (UTEX 1435) with transgene constructs of two different thioesterases (C12 and C14), which involved not only the successful expression of the transgene, but also the correct targeting of the expressed protein to the plastid and a functional effect (the expected change in lipid profile) as a result of the transformation. The same transformation experiment was performed using an expression construct containing a codon-optimized

(according to Table 4) *Cuphea hookeriana* C8-10 thioesterase coding region with the native plastid targeting sequence (SEQ ID NO: 70) yielded no change in lipid profile. While the introduction of the *Cuphea hookeriana* C8-10 transgene into *Prototheca moriformis* (UTEX 1435) was successful and confirmed by Southern blot analysis, no change in C8 or C10 fatty acid production was detected in the transformants compared to the wildtype strain.

Example 11: Generation of *Prototheca moriformis* strain with exogenous plant TE with algal plastid targeting sequence

[0483] In order to investigate whether the use of algal chloroplast/plastid targeting sequences would improve medium chain (C8-C14) thioesterase expression and subsequent medium chain lipid production in *Prototheca moriformis* (UTEX 1435), several putative algal plastid targeting sequences were cloned from *Chlorella protothecoides* and *Prototheca moriformis*. Thioesterase constructs based on *Cuphea hookeriana* C8-10 thioesterase, *Umbellularia californica* C12 thioesterase, and *Cinnamomum camphora* C14 thioesterase were made using made with a *Chlorella sorokiniana* glutamate dehydrogenase promoter/5'UTR and a *Chlorella vulgaris* nitrate reductase 3'UTR. The thioesterase coding sequences were modified by removing the native plastid targeting sequences and replacing them with plastid targeting sequences from the *Chlorella protothecoides* and the *Prototheca moriformis* genomes. The thioesterase expression constructs and their corresponding sequence identification numbers are listed below. Each transformation plasmid also contained a Neo resistance construct that was identical to the ones described in Example 9 above. Additionally, another algal-derived promoter, the *Chlamydomonas reinhardtii* β -tubulin promoter, was also tested in conjunction with the thioesterase constructs. "Native" plastid targeting sequence refers to the higher plant thioesterase plastid targeting sequence. A summary of the constructs used in these experiments is provided below:

Construct Name	Promoter/ 5'UTR	Plastid targeting seq	Gene	3'UTR	SEQ ID NO.
Construct 1	<i>C. sorokiniana</i> glutamate dehydrogenase	<i>C. protothecoides</i> stearyl ACP desaturase	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 71
Construct 2	<i>C. sorokiniana</i> glutamate dehydrogenase	<i>P. moriformis</i> delta 12 fatty acid desaturase	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 72
Construct 3	<i>C. sorokiniana</i> glutamate dehydrogenase	<i>P. moriformis</i> isopentenyl diphosphate synthase	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 73
Construct 4	<i>C. sorokiniana</i>	<i>P. moriformis</i>	<i>Umbellularia</i>	<i>C. vulgaris</i>	SEQ ID NO: 74

	glutamate dehydrogenase	isopentenyl diphosphate synthase	<i>californica</i> C12 TE	nitrate reductase	
Construct 5	<i>C. sorokiniana</i> glutamate dehydrogenase	<i>P. moriformis</i> stearoyl ACP desaturase	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 75
Construct 6	<i>C. sorokiniana</i> glutamate dehydrogenase	<i>C. protothecoides</i> stearoyl ACP desaturase	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 76
Construct 7	<i>C. sorokiniana</i> glutamate dehydrogenase	<i>P. moriformis</i> delta 12 fatty acid desaturase	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 77
Construct 8	<i>C. sorokiniana</i> glutamate dehydrogenase	<i>C. protothecoides</i> stearoyl ACP desaturase	<i>Cinnamomum</i> <i>camphora</i> C14 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 78
Construct 9	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	Native	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 79
Construct 10	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>P. moriformis</i> isopentenyl diphosphate synthase	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 80
Construct 11	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>P. moriformis</i> delta 12 fatty acid desaturase	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 81
Construct 12	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>C. protothecoides</i> stearoyl ACP desaturase	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 82
Construct 13	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>P. moriformis</i> stearoyl ACP desaturase	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 83
Construct 14	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	Native	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 84
Construct 15	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>P. moriformis</i> isopentenyl diphosphate	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 85
Construct 16	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>P. moriformis</i> delta 12 fatty acid desaturase	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 86
Construct 17	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>C. protothecoides</i> stearoyl ACP desaturase	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 87
Construct 18	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulin	<i>P. moriformis</i> stearoyl ACP desaturase	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 88

Construct 19	<i>Chlamydomonas reinhardtii</i> β -tubulin	Native	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 89
Construct 20	<i>Chlamydomonas reinhardtii</i> β -tubulin	<i>P. moriformis</i> isopentenyl diphosphate synthase	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 90
Construct 21	<i>Chlamydomonas reinhardtii</i> β -tubulin	<i>P. moriformis</i> delta 12 fatty acid desaturase	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 91
Construct 22	<i>Chlamydomonas reinhardtii</i> β -tubulin	<i>C. protothecoides</i> stearoyl ACP desaturase	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 92
Construct 23	<i>Chlamydomonas reinhardtii</i> β -tubulin	<i>P. moriformis</i> stearoyl ACP desaturase	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrate reductase	SEQ ID NO: 93

[0484] Each construct was transformed into *Prototheca moriformis* (UTEX 1435) and selection was performed using G418 using the methods described in Example 9 above. Several positive clones from each transformation were picked and screened for the presence thioesterase transgene using Southern blotting analysis. Expression of the thioesterase transgene was confirmed using RT-PCR. A subset of the positive clones (as confirmed by Southern blotting analysis and RT-PCR) from each transformation was selected and grown for lipid profile analysis. Lipid samples were prepared from dried biomass samples of each clone and lipid profile analysis was performed using acid hydrolysis methods described in Example 9. Changes in area percent of the fatty acid corresponding to the thioesterase transgene were compared to wildtype levels, and clones transformed with a thioesterase with the native plastid targeting sequence.

[0485] The clones transformed with *Cuphea hookeriana* C8-10 thioesterase constructs with the native plastid targeting sequence had the same level of C8 and C10 fatty acids as wildtype. The clones transformed with *Cuphea hookeriana* C8-10 thioesterase constructs (Constructs 1-3) with algal plastid targeting sequences had over a 10-fold increase in C10 fatty acids for Construct 3 and over 40-fold increase in C10 fatty acids for Constructs 1 and 2 (as compared to wildtype). The clones transformed with *Umbellularia californica* C12 thioesterase constructs with the native plastid targeting sequence had a modest 6-8 fold increase in C12 fatty acid levels as compared to wildtype. The clones transformed with the *Umbellularia californica* C12 thioesterase constructs with the algal plasid targeting

constructs (Constructs 4-7) had over an 80-fold increase in C12 fatty acid level for Construct 4, about an 20-fold increase in C12 fatty acid level for Construct 6, about a 10-fold increase in C12 fatty acid level for Construct 7 and about a 3-fold increase in C12 fatty acid level for Construct 5 (all compared to wildtype). The clones transformed with *Cinnamomum camphora* C14 thioesterase with either the native plastid targeting sequence or the construct 8 (with the *Chlorella protothecoides* stearoyl ACP desaturase plastid targeting sequence) had about a 2-3 fold increase in C14 fatty acid levels as compared to wildtype. In general clones transformed with an algal plastid targeting sequence thioesterase constructs had a higher fold increase in the corresponding chain-length fatty acid levels than when using the native higher plant targeting sequence.

1. *Chlamydomonas reinhardtii* β -tubulin promoter

[0486] Additional heterologous thioesterase expression constructs were prepared using the *Chlamydomonas reinhardtii* β - tubulin promoter instead of the *C. sorokinana* glutamate dehydrogenase promoter. The construct elements and sequence of the expression constructs are listed above. Each construct was transformed into *Prototheca moriformis* UTEX 1435 host cells using the methods described above. Lipid profiles were generated from a subset of positive clones for each construct in order to assess the success and productivity of each construct. The lipid profiles compare the fatty acid levels (expressed in area %) to wildtype host cells. The “Mean” column represents the numerical average of the subset of positive clones. The “Sample” column represents the best positive clone that was screened (best being defined as the sample that produced the greatest change in area % of the corresponding chain-length fatty acid production). The “low-high” column represents the lowest area % and the highest area % of the fatty acid from the clones that were screened. The lipid profile results of Constructs 9-12 and 14-23 are summarized below.

Construct 9. <i>Cuphea hookeriana</i> C8-10 TE				
Fatty Acid	wildtype	Mean	Sample	low/high
C 8:0	0	0.05	0.30	0-0.29
C 10:0	0.01	0.63	2.19	0-2.19
C 12:0	0.03	0.06	0.10	0-0.10
C 14:0	1.40	1.50	1.41	1.36-3.59
C 16:0	24.01	24.96	24.20	
C 16:1	0.67	0.80	0.85	
C 17:0	0	0.16	0.16	
C 17:1	0	0.91	0	
C 18:0	4.15	17.52	3.19	
C 18:1	55.83	44.81	57.54	
C 18:2	10.14	7.58	8.83	
C 18:3 α	0.93	0.68	0.76	
C 20:0	0.33	0.21	0.29	

C 24:0 0 0.05 0.11

Construct 10. *Cuphea hookeriana* C8-10 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 8:0	0	0.01	0.02	0-0.03
C 10:0	0	0.16	0.35	0-0.35
C 12:0	0.04	0.05	0.07	0-0.07
C 14:0	1.13	1.62	1.81	0-0.05
C 14:1	0	0.04	0.04	
C 15:0	0.06	0.05	0.05	
C 16:0	19.94	26.42	28.08	
C 16:1	0.84	0.96	0.96	
C 17:0	0.19	0.14	0.13	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	3.62	3.43	
C 18:1	63.96	54.90	53.91	
C 18:2	9.62	9.83	9.11	
C18:3 γ	0	0.01	0	
C 18:3 α	0.63	0.79	0.73	
C 20:0	0.26	0.35	0.33	
C 20:1	0.06	0.08	0.09	
C 20:1	0.08	0.06	0.07	
C 22:0	0	0.08	0.09	
C 24:0	0.13	0.13	0.11	

Construct 11. *Cuphea hookeriana* C8-10 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 8:0	0	0.82	1.57	0-1.87
C 10:0	0	3.86	6.76	0-6.76
C 12:0	0.04	0.13	0.20	0.03- 0.20
C 14:0	1.13	1.80	1.98	1.64- 2.05
C 14:1	0	0.04	0.04	
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	25.60	25.44	
C 16:1	0.84	1.01	1.02	
C 17:0	0.19	0.13	0.11	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	2.98	2.38	
C 18:1	63.96	51.59	48.85	
C 18:2	9.62	9.85	9.62	
C18:3 γ	0	0.01	0	
C 18:3 α	0.63	0.91	0.92	
C 20:0	0.26	0.29	0.26	
C 20:1	0.06	0.06	0	
C 20:1	0.08	0.06	0.03	
C 22:0	0	0.08	0.08	
C 24:0	0.13	0.06	0	

Construct 12. *Cuphea hookeriana* C8-10 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 8:0	0	0.31	0.85	0-0.85
C 10:0	0	2.16	4.35	0.20- 4.35
C 12:0	0.04	0.10	0.15	0-0.18
C 14:0	1.13	1.96	1.82	1.66- 2.97

C 14:1	0	0.03	0.04
C 15:0	0.06	0.07	0.07
C 16:0	19.94	26.08	25.00
C 16:1	0.84	1.04	0.88
C 17:0	0.19	0.16	0.16
C 17:1	0.10	0.05	0.07
C 18:0	2.68	3.02	3.19
C 18:1	63.96	51.08	52.15
C 18:2	9.62	11.44	9.47
C18:3 γ	0	0.01	0
C 18:3 α	0.63	0.98	0.90
C 20:0	0.26	0.30	0.28
C 20:1	0.06	0.06	0.05
C 20:1	0.08	0.04	0
C 22:0	0	0.07	0
C 24:0	0.13	0.05	0

Construct 14. *Umbellularia californica* C12 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0.01	0.02	0.03	0.02-0.03
C 12:0	0.03	2.62	3.91	0.04-3.91
C 14:0	1.40	1.99	2.11	1.83-2.19
C 16:0	24.01	27.64	27.01	
C 16:1	0.67	0.92	0.92	
C 18:0	4.15	2.99	2.87	
C 18:1	55.83	53.22	52.89	
C 18:2	10.14	8.68	8.41	
C 18:3 α	0.93	0.78	0.74	
C 20:0	0.33	0.29	0.27	

Construct 15. *Umbellularia californica* C12 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0	0.05	0.08	0-0.08
C 12:0	0.04	8.12	12.80	4.35-12.80
C 13:0	0	0.02	0.03	0-0.03
C 14:0	1.13	2.67	3.02	2.18-3.37
C 14:1	0	0.04	0.03	0.03-0.10
C 15:0	0.06	0.07	0.06	
C 16:0	19.94	25.26	23.15	
C 16:1	0.84	0.99	0.86	
C 17:0	0.19	0.14	0.14	
C 17:1	0.10	0.05	0.05	
C 18:0	2.68	2.59	2.84	
C 18:1	63.96	46.91	44.93	
C 18:2	9.62	10.59	10.01	
C 18:3 α	0.63	0.92	0.83	
C 20:0	0.26	0.27	0.24	
C 20:1	0.06	0.06	0.06	
C 20:1	0.08	0.05	0.04	
C 22:0	0	0.07	0.09	
C 24:0	0.13	0.13	0.12	

Construct 16. <i>Umbellularia californica</i> C12 TE				
Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0	0.03	0.04	0.02-0.04
C 12:0	0.04	2.43	5.32	0.98-5.32
C 13:0	0	0.01	0.02	0-0.02
C 14:0	1.13	1.77	1.93	1.62-1.93
C 14:1	0	0.03	0.02	0.02-0.04
C 15:0	0.06	0.06	0.05	
C 16:0	19.94	24.89	22.29	
C 16:1	0.84	0.91	0.82	
C 17:0	0.19	0.16	0.15	
C 17:1	0.10	0.06	0.06	
C 18:0	2.68	3.81	3.67	
C 18:1	63.96	53.19	52.82	
C 18:2	9.62	10.38	10.57	
C 18:3 α	0.63	0.80	0.77	
C 20:0	0.26	0.35	0.32	
C 20:1	0.06	0.06	0.07	
C 20:1	0.08	0.07	0.08	
C 22:0	0	0.08	0.07	
C 24:0	0.13	0.15	0.14	

Construct 17. <i>Umbellularia californica</i> C12 TE				
Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0	0.04	0.07	0.03-0.08
C 12:0	0.04	7.02	14.11	4.32-14.11
C 13:0	0	0.03	0.04	0.01-0.04
C 14:0	1.13	2.25	3.01	1.95-3.01
C 14:1	0	0.03	0.03	0.02-0.03
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	23.20	21.46	
C 16:1	0.84	0.82	0.77	
C 17:0	0.19	0.15	0.14	
C 17:1	0.10	0.06	0.06	
C 18:0	2.68	3.47	2.93	
C 18:1	63.96	50.30	45.17	
C 18:2	9.62	10.33	9.98	
C18:3 γ	0	0.01	0	
C 18:3 α	0.63	0.84	0.86	
C 20:0	0.26	0.32	0.27	
C 20:1	0.06	0.07	0.06	
C 20:1	0.08	0.06	0.06	
C 22:0	0	0.08	0.09	
C 24:0	0.13	0.14	0.13	

Construct 18. <i>Umbellularia californica</i> C12 TE				
Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0	0.03	0.05	0.01-0.05
C 12:0	0.04	5.06	7.77	0.37-

				7.77
C 13:0	0	0.02	0	0-0.03
C 14:0	1.13	2.11	2.39	1.82- 2.39
C 14:1	0	0.03	0.03	0.02- 0.05
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	24.60	23.95	
C 16:1	0.84	0.86	0.83	
C 17:0	0.19	0.15	0.14	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	3.31	2.96	
C 18:1	63.96	51.26	49.70	
C 18:2	9.62	10.18	10.02	
C 18:3 γ	0	0.01	0.02	
C 18:3 α	0.63	0.86	0.86	
C 20:0	0.26	0.32	0.29	
C 20:1	0.06	0.05	0.05	
C 20:1	0.08	0.07	0.04	
C 22:0	0	0.08	0.08	
C 24:0	0.13	0.13	0.13	

Construct 19. *Cinnamomum camphora* C14 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0.02	0.01	0.01	0.01- 0.02
C 12:0	0.05	0.27	0.40	0.08- 0.41
C 14:0	1.52	4.47	5.81	2.10- 5.81
C 16:0	25.16	28.14	28.55	
C 16:1	0.72	0.84	0.82	
C 18:0	3.70	3.17	2.87	
C 18:1	54.28	51.89	51.01	
C 18:2	12.24	9.36	8.62	
C 18:3 α	0.87	0.74	0.75	
C 20:0	0.33	0.33	0.31	

Construct 20. *Cinnamomum camphora* C14 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0.01	0.01	0.02	0.01- 0.02
C 12:0	0.03	0.39	0.65	0.08- 0.65
C 13:0	0	0.01	0.01	0.01- 0.02
C 14:0	1.40	5.61	8.4	2.1-8.4
C 14:1	0	0.03	0.03	0.02- 0.03
C 15:0	0	0.06	0.07	
C 16:0	24.01	25.93	25.57	
C 16:1	0.67	0.75	0.71	
C 17:0	0	0.13	0.12	
C 17:1	0	0.05	0.05	
C 18:0	4.15	3.30	3.23	
C 18:1	55.83	51.00	48.48	
C 18:2	10.14	10.38	10.35	
C 18:3 α	0.93	0.91	0.88	
C 20:0	0.33	0.35	0.32	

C 20:1	0	0.08	0.08
C 20:1	0	0.07	0.07
C 22:0	0	0.08	0.08
C 24:0	0	0.14	0.13

Construct 21. *Cinnamomum camphora* C14 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0.01	0.01	0.01	0-0.01
C 12:0	0.03	0.10	0.27	0.04- 0.27
C 14:0	1.40	2.28	4.40	1.47- 4.40
C 16:0	24.01	26.10	26.38	
C 16:1	0.67	0.79	0.73	
C 17:0	0	0.15	0.16	
C 17:1	0	0.06	0.06	
C 18:0	4.15	3.59	3.51	
C 18:1	55.83	53.53	50.86	
C 18:2	10.14	10.83	11.11	
C 18:3 α	0.93	0.97	0.87	
C 20:0	0.33	0.36	0.37	
C 20:1	0	0.09	0.08	
C 20:1	0	0.07	0.07	
C 22:0	0	0.09	0.09	

Construct 22. *Cinnamomum camphora* C14 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0.01	0.02	0.02	0.02- 0.02
C 12:0	0.03	1.22	1.83	0.59- 1.83
C 13:0	0	0.02	0.03	0.01- 0.03
C 14:0	1.40	12.77	17.33	7.97- 17.33
C 14:1	0	0.02	0.02	0.02- 0.04
C 15:0	0	0.07	0.08	
C 16:0	24.01	24.79	24.22	
C 16:1	0.67	0.64	0.58	
C 17:0	0	0.11	0.10	
C 17:1	0	0.04	0.04	
C 18:0	4.15	2.85	2.75	
C 18:1	55.83	45.16	41.23	
C 18:2	10.14	9.96	9.65	
C 18:3 α	0.93	0.91	0.85	
C 20:0	0.33	0.30	0.30	
C 20:1	0	0.07	0.06	
C 20:1	0	0.06	0.05	
C 22:0	0	0.08	0.08	

Construct 23. *Cinnamomum camphora* C14 TE

Fatty Acid	wildtype	Mean	Sample	low/high
C 10:0	0.01	0.01	0.02	0-0.02
C 12:0	0.05	0.57	1.08	0.16- 1.08
C 13:0	0	0.02	0.02	0-0.02
C 14:0	1.45	7.18	11.24	2.96-

C 14:1	0.02	0.03	0.03	11.24 0.02- 0.03
C 15:0	0.06	0.07	0.07	
C 16:0	24.13	25.78	25.21	
C 16:1	0.77	0.72	0.66	
C 17:0	0.19	0.13	0.11	
C 17:1	0.08	0.05	0.04	
C 18:0	3.53	3.35	3.12	
C 18:1	56.15	49.65	46.35	
C 18:2	11.26	10.17	9.72	
C 18:3 α	0.84	0.95	0.83	
C 20:0	0.32	0.34	0.32	
C 20:1	0.09	0.08	0.09	
C 20:1	0.07	0.05	0.06	
C 22:0	0.07	0.08	0.08	
C 24:0	0.13	0.13	0.12	

[0487] Constructs 9-12 were expression vectors containing the *Cuphea hookeriana* C8-10 thioesterase construct. As can be seen in the data summaries above, the best results were seen with Construct 11, with the Sample C8 fatty acid being 1.57 Area % (as compared to 0 in wildtype) and C10 fatty acid being 6.76 Area % (as compared to 0 in wildtype). There was also a modest increase in C12 fatty acids (approximately 2-5 fold increase). While the native plastid targeting sequence produced no change when under the control of the *C.sorokinana* glutamate dehydrogenase promoter, the same expression construct driven by the *C.reinhardtii* β -tubulin promoter produced significant changes in C8-10 fatty acids in the host cell. This is further evidence of the idiosyncrasies of heterologous expression of thioesterases in *Prototheca* species. All of the clones containing the *C.reinhardtii* β -tubulin promoter C8-10 thioesterase construct had greater increases in C8-10 fatty acids than the clones containing the *C.sorokinana* glutamate dehydrogenase promoter C8-10 thioesterase construct. Lipid profile data for Construct 13 was not obtained and therefore, not included above.

[0488] Constructs 14-18 were expression vectors containing the *Umbellularia californica* C12 thioesterase construct. As can be seen in the data summaries above, the best results were seen with Constructs 15 (*P.moriformis* isopentenyl diphosphate synthase plastid targeting sequence) and 17 (*C.protothecoides* stearoyl ACP desaturase plastid targeting sequence). The greatest change in C12 fatty acid production was seen with Construct 17, with C12 fatty acids levels of up to 14.11 area %, as compared to 0.04 area % in wildtype. Modest changes (about 2-fold) were also seen with C14 fatty acid levels. When compared to the same constructs with the *C.sorokinana* glutamate dehydrogenase promoter, the same trends were true with the *C.reinhardtii* β -tubulin promoter—the *C.protothecoides* stearoyl ACP desaturase and

P.moriformis isopentenyl diphosphate synthase plastid targeting sequences produced the greatest change in C12 fatty acid levels with both promoters.

[0489] Constructs 19-23 were expression vectors containing the *Cinnamomum camphora* C14 thioesterase construct. As can be seen in the data summaries above, the best results were seen with Constructs 22 and Construct 23. The greatest change in C14 fatty acid production was seen with Construct 22, with C14 fatty acid levels of up to 17.35 area % (when the values for C140 and C141 are combined), as compared to 1.40% in wildtype. Changes in C12 fatty acids were also seen (5-60 fold). When compared to the same constructs with the *C.sorokinana* glutamate dehydrogenase promoter, the same trends were true with the *C.reinhardtii* β -tubulin promoter—the *C.protothecoides* stearoyl ACP desaturase and *P.moriformis* stearoyl ACP desaturase plastid targeting sequences produced the greatest change in C14 fatty acid levels with both promoters. Consistently with all thioesterase expression constructs, the *C.reinhardtii* β -tubulin promoter constructs produced greater changes in C8-14 fatty acid levels than the *C.sorokiniana* glutamate dehydrogenase

[0490] Two positive clones from the Construct 22 were selected and grown under high selective pressure (50mg/L G418). After 6 days in culture, the clones were harvested and their lipid profile was determined using the methods described above. The lipid profile data is summarized below and is expressed in area %.

Construct 22 clones + 50mg/L G418		
Fatty Acid	Construct 22 A	Construct 22 B
C 12:0	3.21	3.37
C 14:0	27.55	26.99
C 16:0	25.68	24.37
C 16:1	0.99	0.92
C 18:0	1.37	1.23
C 18:1	28.35	31.07
C 18:2	11.73	11.05
C 18:3 α	0.92	0.81
C 20:0	0.16	0.17

[0491] Both clones, when grown under constant, high selective pressure, produced an increased amount of C14 and C12 fatty acids, about double the levels seen with Construct 22 above. These clones yielded over 30 area % of C12-14 fatty acids, as compared to 1.5 area % of C12-14 fatty acids seen in wildtype cells.

WHAT IS CLAIMED IS:

1. A method of extracting a lipid from a microorganism comprising the steps of:
 - (a) lysing a cultured microorganism to produce a lysate, wherein the microorganism:
 - (i) has not been subjected to a drying step between culturing and lysing; and
 - (ii) contains a lipid;
 - (b) treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent; and
 - (c) separating the lysate into layers comprising a lipid:organic solvent layer and an aqueous layer;
whereby the lipid is extracted from the microorganism.
2. The method of claim 1, wherein the treating step comprises treating with more than about 5% v/v of an organic solvent to the lysate.
3. The method of claim 1, wherein the layers further comprise a lipid:aqueous emulsion layer and/or a cell pellet.
4. The method of claim 1, wherein step (a) is performed at an acidic pH.
5. The method of claim 1, wherein step (a) comprises one or more methods of lysing the microorganism selected from the list consisting of heating, sonication, mechanical lysis, osmotic shock, expression of an autolysis gene, exposure to pH above 8, exposure to pH below 6, and digestion with an enzyme.
6. The method of claim 1, wherein the microorganism produces a lipid that comprises at least 25% w/v of total cellular lipid as C18 triacylglycerols.
7. The method of claim 1, wherein the microorganism produces a lipid that comprises at least 25% w/v of total cellular lipid as C16 triacylglycerols.
8. The method of claim 1, wherein the microorganism produces a lipid that comprises at least 10% w/v of total cellular lipid as C14 triacylglycerols.

9. The method of claim 1, wherein the microorganism produces a lipid that comprises at least 10% w/v of total cellular lipid as C12 triacylglycerols.
10. The method of claim 1, wherein the microorganism produces a lipid that comprises at least 10% w/v of total cellular lipid as C10 triacylglycerols.
11. The method of claim 1, wherein the microorganism produces a lipid with a lipid profile comprising at least 4% C8-C14.
12. The method of claim 1, wherein the microorganism produces a lipid with a lipid profile comprising between 4% and 40% C10-C14.
13. The method of claim 1, wherein step (a) comprises acidic lysis and heating.
14. The method of claim 1, wherein the organic solvent comprises an oil selected from the group consisting of oil from soy, rapeseed, canola, palm, palm kernel, coconut, corn, waste vegetable, Chinese tallow, olive, sunflower, cotton seed, chicken fat, beef tallow, porcine tallow, microalgae, macroalgae, *Cuphea*, flax, peanut, choice white grease, lard, *Camelina sativa*, mustard seed cashew nut, oats, lupine, kenaf, calendula, hemp, coffee, linseed, hazelnut, euphorbia, pumpkin seed, coriander, camellia, sesame, safflower, rice, tung oil tree, cocoa, copra, pium poppy, castor beans, pecan, jojoba, jatropha, macadamia, Brazil nuts, avocado, a fossil oil, or a distillate fraction thereof.
15. The method of claim 1, wherein the organic solvent is a surfactant.
16. The method of claim 1, wherein step (c) comprises reducing the temperature of the mixture below 25°C.
17. The method of claim 1, wherein the microorganism is selected from the group of microorganisms consisting of a bacterium, a cyanobacterium, a eukaryotic microalgae, an oleaginous yeast, and a fungus.
18. The method of claim 17, wherein the microorganism is of the genus *Chlorella* or *Prototheca*.
19. The method of claim 1, further comprising the step of:
(d) removing the lipid:organic solvent composition from the other layer(s).

20. The method of claim 19, wherein step (d) is performed without substantially separating the lipid from the organic solvent composition.
21. A method of producing fuel comprising the steps of:
- (a) lysing a lipid-containing microorganism to produce a lysate;
 - (b) treating the lysate with an organic solvent for a period of time sufficient to allow the lipid from the microorganism to become solubilized in the organic solvent;
 - (c) separating the lysate into a lipid:organic solvent composition and an aqueous composition and, optionally, an emulsified composition and/or cell pellet composition;
 - (d) removing the lipid:organic solvent composition from the aqueous composition, emulsion composition, or cell pellet composition; and
 - (e) subjecting the lipid:organic solvent composition to a chemical reaction, whereby a fuel is produced.
22. The method of claim 21, wherein the chemical reaction is transesterification and the fuel is biodiesel.
23. The method of claim 22, wherein the biodiesel meets or exceeds the ASTM D6751 biodiesel standard.
24. The method of claim 22, wherein the biodiesel meets or exceeds the EN 14214 biodiesel standard.
25. The method of claim 21, wherein the chemical reaction comprises one or more reactions selected from the group consisting of hydroprocessing, deoxygenation, and isomerization and the fuel is renewable diesel.
26. The method of claim 25, wherein the renewable diesel meets or exceeds the ASTM D 975 standard.
27. The method of claim 21, wherein the chemical reaction comprises hydrotreating and the fuel is jet fuel.
28. The method of claim 21, wherein the microorganism has not been subjected to a drying step.

152

29. The method of claim 21, wherein the organic solvent comprises an oil selected from the group consisting of oil from soy, rapeseed, canola, palm, palm kernel, coconut, corn, waste vegetable, Chinese tallow, olive, sunflower, cotton seed, chicken fat, beef tallow, porcine tallow, microalgae, macroalgae, *Cuphea*, flax, peanut, choice white grease, lard, *Camelina sativa*, mustard seed, cashew nut, oats, lupine, kenaf, calendula, hemp, coffee, linseed, hazelnut, euphorbia, pumpkin seed, coriander, camellia, sesame, safflower, rice, tung oil tree, cocoa, copra, pium poppy, castor beans, pecan, jojoba, jatropha, macadamia, Brazil nuts, avocado, a fossil oil or a distillate fraction thereof.

30. The method of claim 21, wherein the organic solvent is a surfactant.

31. The method of claim 21, wherein the microorganism is selected from the group of microorganisms consisting of a bacterium, a cyanobacterium a eukaryotic microalgae, an oleaginous yeast, and a fungus.

32. The method of claim 31, wherein the microorganism is of the genus *Chlorella* or *Prototheca*.

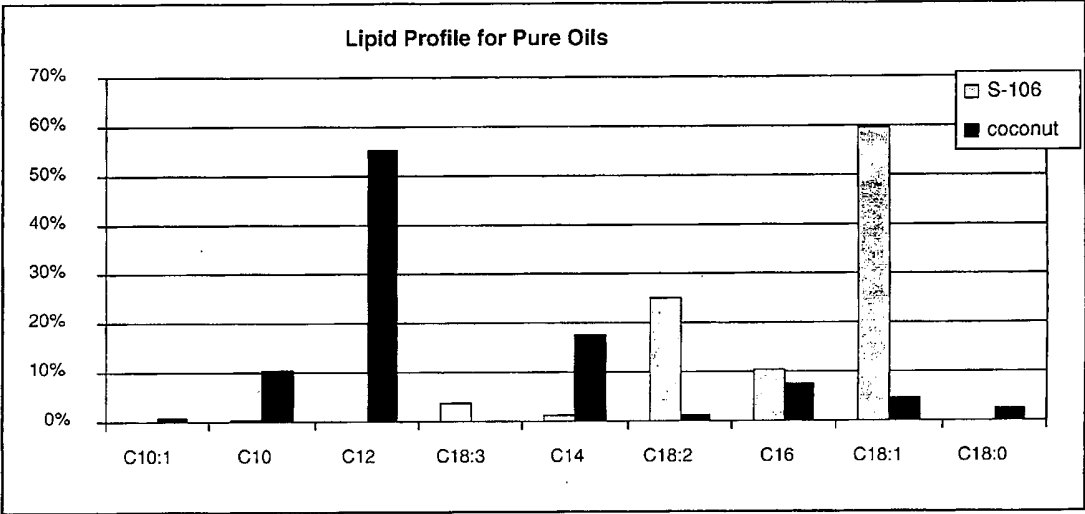
33. The method of claim 21, wherein the microorganism produces a lipid with a lipid profile comprising at least 4% C8-C14.

34. The method of claim 21, wherein the microorganism produces a lipid with a lipid profile comprising between 4% and 40% C10-C14.

35. The method of claim 21, wherein step (e) is performed without substantially separating the lipid from the organic solvent.

36. The method of claim 21, wherein the microorganism has at least 85% 23S rRNA genomic sequence identity to one or more sequences selected from the group consisting of SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, and SEQ ID NO:39.

Figure 1



1/8
SUBSTITUTE SHEET (RULE 26)

Figure 2

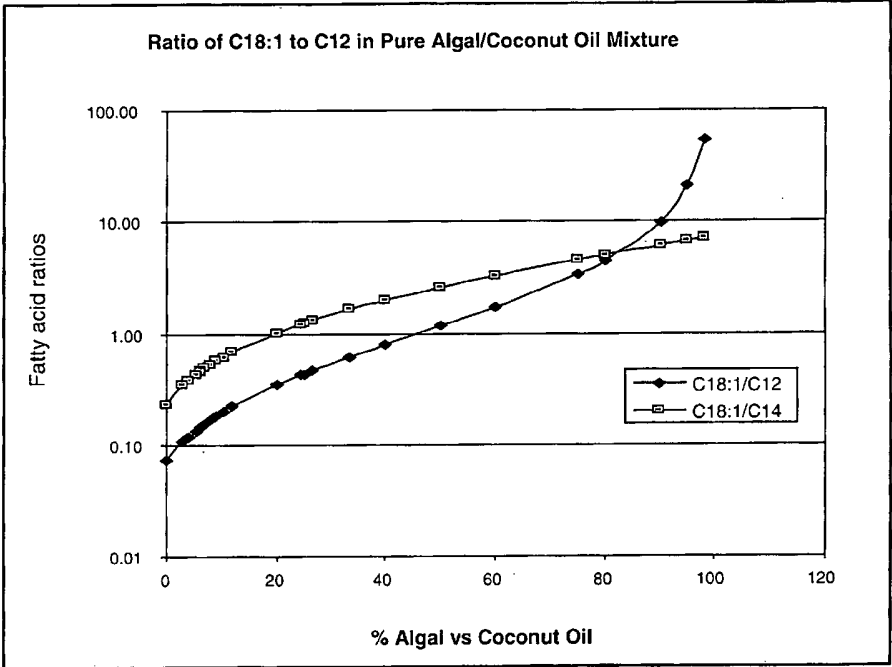


Figure 3

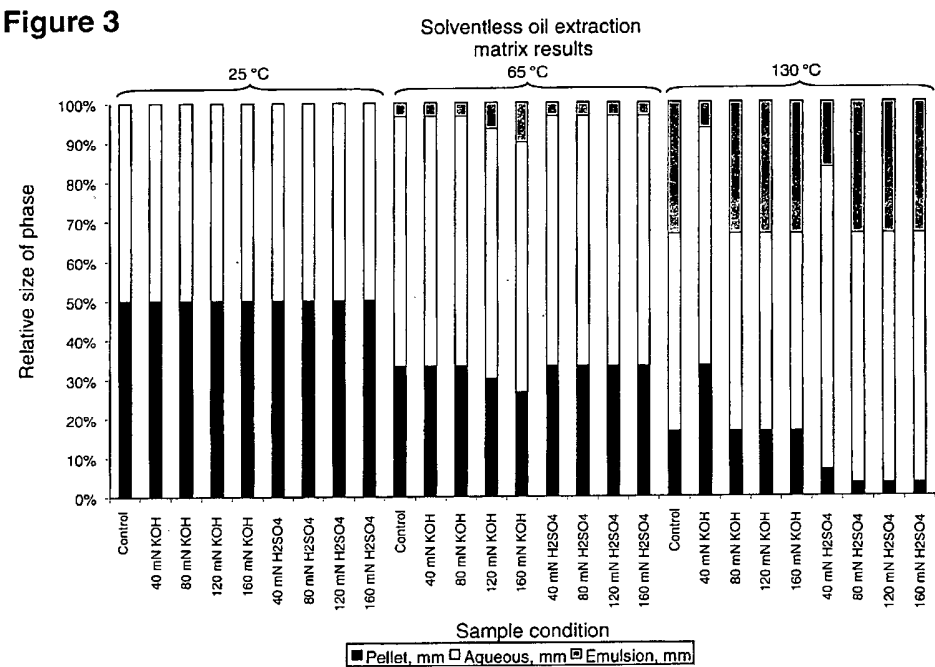


Figure 4

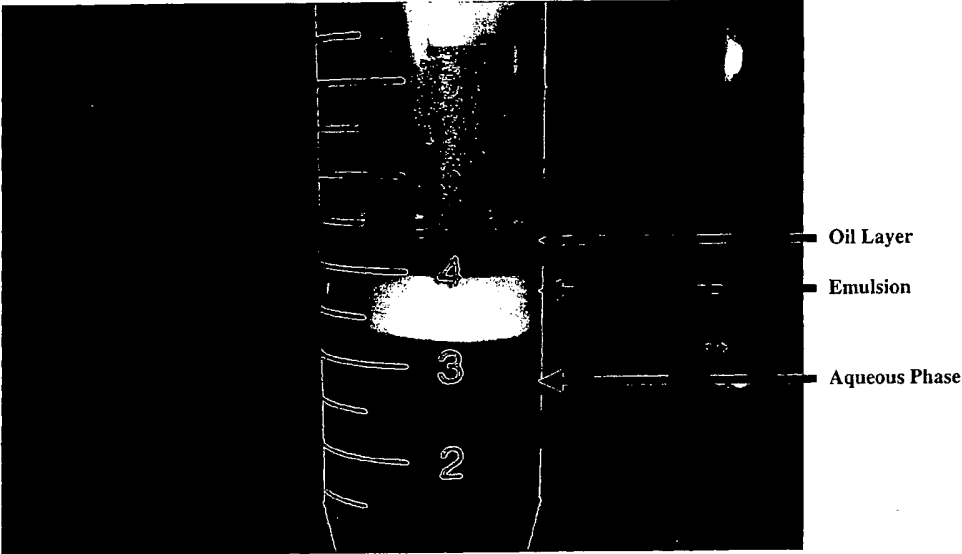


Figure 5

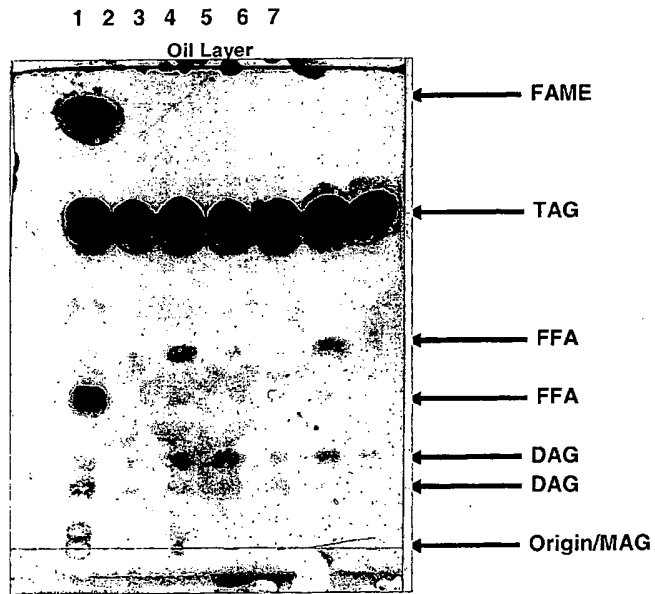


Figure 6

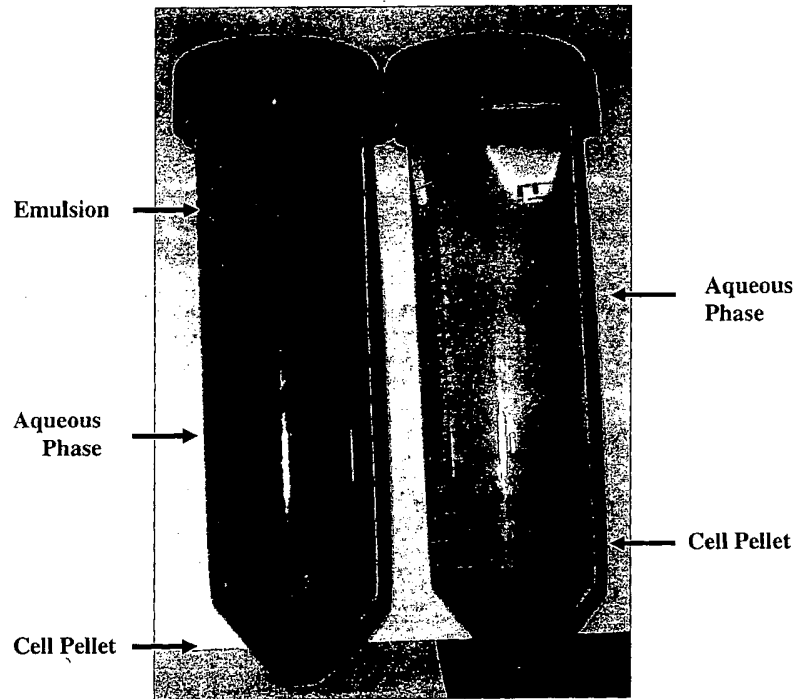


Figure 7

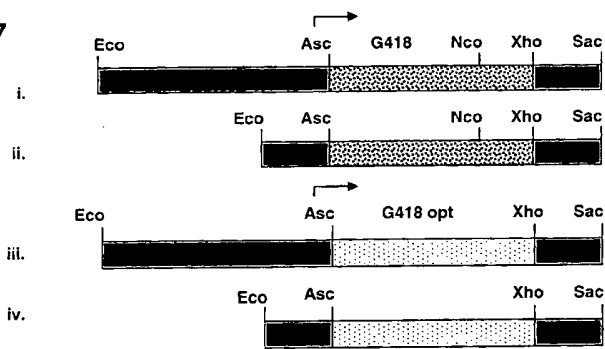
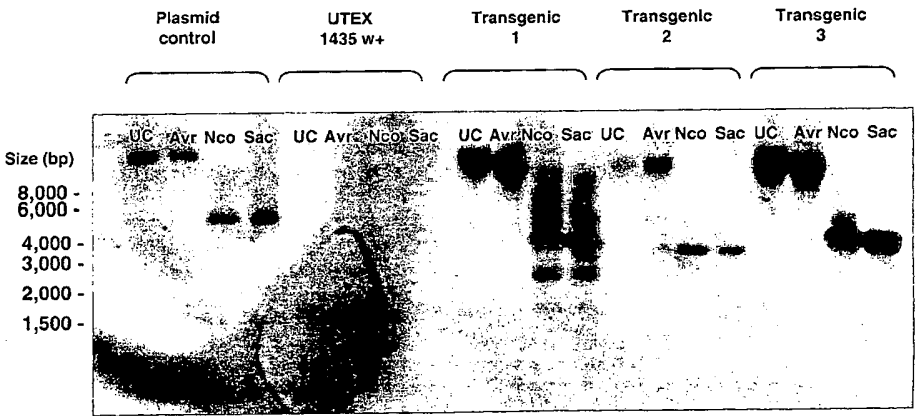


Figure 8



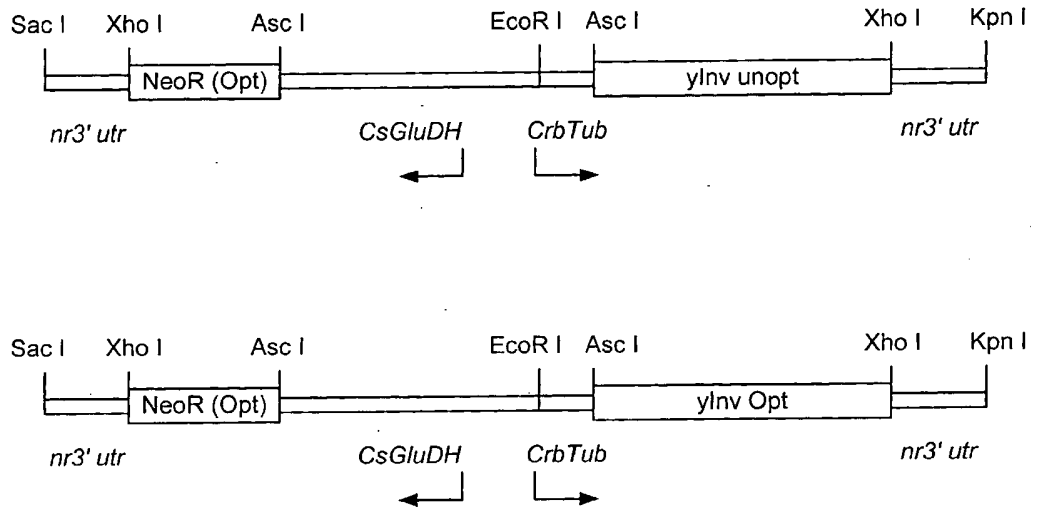


Figure 9

FRACCIONAMIENTO DE LA BIOMASA MICROBIANA QUE CONTIENE ACEITE

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

- 5 Esta solicitud reivindica el beneficio bajo 35 U.S.C. 119(e) de la solicitud provisional de Estados Unidos núm. 61/181,252, presentada el 26 de mayo de 2009, la que se incorporada en la presente como referencia en su totalidad para todos los propósitos.

REFERENCIA AL LISTADO DE SECUENCIA

- 10 Esta solicitud incluye un LISTADO DE SECUENCIAS, anexado como páginas 1-45.

CAMPO DE LA INVENCION

- 15 La invención se refiere, generalmente, a la producción de composiciones de aceite en microorganismos, tal como un lípido. En particular, la invención proporciona métodos para la extracción, recuperación, aislamiento y obtención de un lípido a partir de un microorganismo y composiciones que comprenden el lípido. La invención describe también los métodos para la producción de composiciones de hidrocarburo o lípido para la producción de biodiesel, diesel renovable, combustible de reactor, y surfactantes lipídicos que tienen diferentes longitudes de cadena de carbono, que incluyen C8, C10, C12 y C14.

ANTECEDENTES DE LA INVENCION

- 20 [0001] El combustible fósil es un término general para los depósitos geológicos de combustible enterrado de materiales orgánicos, formados a partir de plantas y animales descompuestos, que se transforman en aceite crudo, hulla, gas natural o aceites pesados por la exposición al calor y a la presión en la corteza terrestre durante cientos de millones de años.
- 25 [0002] Los combustibles fósiles son una fuente finita, no renovable. Con la modernización global en los siglos 20 y 21, las ansias de energía a partir de combustibles fósiles, especialmente la gasolina derivada del petróleo, es una de las causas de mayores conflictos regionales y globales. El incremento de la demanda energética por la globalización económica también ha ejercido presión en el aumento del costo de los hidrocarburos.

Además de la energía, muchas industrias, incluyendo los productores de plástico y sustancias químicas, dependen en gran medida de la disponibilidad de hidrocarburos como materia prima para sus procesos de fabricación. Las alternativas rentables a las fuentes actuales de suministro pueden ayudar a atenuar la presión ascendente sobre estas materias primas.

5 Los lípidos para uso en biocombustibles se pueden producir en microorganismos, tales como algas, hongos y bacterias. Típicamente, la fabricación de un lípido en un microorganismo implica cultivar los microorganismos, tales como algas, hongos o bacterias, que son capaces de producir un lípido deseado en un fermentador o biorreactor, aislar la biomasa microbiana, secarla y extraer los lípidos intracelulares.

10 Existe una necesidad de un proceso para la extracción de lípidos a partir del microorganismo que solucione los problemas anteriores de baja eficiencia y alto costo de extracción de lípido a partir del microorganismo. La presente invención proporciona una solución a estos problemas anteriores de la técnica.

BREVE RESUMEN DE LA INVENCION

15 En un aspecto, la presente invención se dirige a un método de extracción de un lípido a partir de un microorganismo. En una modalidad, el método comprende lisar un microorganismo cultivado que no se ha sometido a una etapa de secado entre el cultivo y la lisis, y que contiene un lípido, para producir un lisado, tratar el lisado con un solvente orgánico durante un período de tiempo suficiente para permitir que el lípido del
20 microorganismo se solubilice en el solvente orgánico, y separar el lisado en capas que comprenden una capa lípido:solvente orgánico, y una capa acuosa, a través de lo cual el lípido se extrae del microorganismo.

En otro aspecto, la presente invención se dirige a una composición que comprende un lípido aislado de un microorganismo, y un aceite obtenido de una fuente que no es el
25 microorganismo. En algunos casos, la proporción de lípido al aceite está entre 1 y 100. En otros casos, la proporción del lípido al aceite está entre 1 y 10. En algunas modalidades, el lípido comprende al menos 50% de un lípido C18:1, o al menos 10% de un lípido combinado C12 y C14.

En otro aspecto, la presente invención se dirige a un método para la producción de
30 biodiesel. En una modalidad, el método comprende lisar un microorganismo que contiene lípido para producir un lisado, tratar el lisado con un solvente orgánico durante un período de

tiempo suficiente para permitir que el lípido del microorganismo se solubilice en el solvente orgánico, separar el lisado en una composición lípido:solvente orgánico y una composición acuosa y, opcionalmente, una composición emulsionada y/o composición de sedimento celular, eliminar la composición lípido:solvente orgánico de la composición acuosa, composición de emulsión, o composición de sedimento celular, y transesterificar la composición lípido:solvente orgánico para producir biodiesel. En algunos casos, el biodiesel cumple o excede la norma de biodiesel ASTM D6751 y/o la norma de biodiesel EN 14214. En una modalidad, el microorganismo no se somete a una etapa de secado.

En algunas modalidades del método para la producción de biodiesel al menos 20% p/v del lípido total a partir del microorganismo es C18:1. En una modalidad, las capas comprenden además una capa lipídica:emulsión acuosa y/o un sedimento celular. La transesterificación del lípido: solvente orgánico se puede realizar sin separar sustancialmente el lípido del solvente orgánico.

Aún en otro aspecto, la presente invención se dirige a un método de producción de diesel renovable. En una modalidad, el método comprende lisar un microorganismo que contiene lípido para producir un lisado, tratar el lisado con un solvente orgánico durante un período de tiempo suficiente para permitir que el lípido del microorganismo se solubilice en el solvente orgánico, separar el lisado en una composición lípido:solvente orgánico y una composición acuosa y, opcionalmente, una composición emulsionada y/o composición de sedimento celular, eliminar la composición lípido:solvente orgánico de la composición acuosa, composición de la emulsión, o composición de sedimento celular, y tratar la composición lípido:solvente orgánico para producir un alcano de cadena lineal del producto diesel renovable. En algunos casos, el diesel renovable cumple o excede la norma ASTM D 975.

En diversas modalidades del método de producción de diesel renovable, el microorganismo no se ha sometido a una etapa de secado. En algunos casos, el tratamiento de la composición lípido:solvente orgánico se realiza sin la separación sustancial del lípido del solvente orgánico. En algunas modalidades, el tratamiento de la composición lípido:solvente orgánico comprende el hidrotratamiento, hidroprocesamiento, o la licuefacción indirecta.

En otro aspecto, la presente invención se dirige a un método para producir un combustible de reactor. En una modalidad, el método comprende lisar un microorganismo que contiene lípido para producir un lisado, tratar el lisado con un solvente orgánico durante

un período de tiempo suficiente para permitir que el lípido del microorganismo se solubilice en el solvente orgánico, separar el lisado en una composición lípido:solvente orgánico y una composición acuosa y, opcionalmente, una composición emulsionada o composición de sedimento celular, eliminar la composición lípido:solvente orgánico de la composición acuosa, composición de emulsión, o composición de sedimento celular, tratar la composición lípido:solvente orgánico para producir un alcano de cadena lineal y craquear el alcano de cadena lineal para producir el producto de combustible de reactor.

En diversas modalidades del método de producción de combustible de reactor, el microorganismo no se ha sometido a una etapa de secado. En algunos casos, el tratamiento de la composición lípido:solvente orgánico se realiza sin separar sustancialmente el lípido del solvente orgánico. En algunas modalidades, el tratamiento de la composición lípido: solvente orgánico se realiza haciendo fluir la composición lípido:solvente orgánico a una zona de craqueo catalítico fluido, y puede comprender, además, contactar la composición lípido:solvente orgánico con un catalizador en condiciones de craqueo. En una modalidad, el tratamiento de la composición lípido:solvente orgánico se realiza por hidrogenodesoxigenación de la composición lípido:solvente orgánico. En algunos casos, el método comprende, además, someter la composición lípido hidrogenodesoxigenada:solvente orgánico a isomerización.

En otro aspecto, la presente invención se dirige a un método de extracción de lípidos a partir de un microorganismo al contactar un microorganismo que contiene un lípido con un ácido para producir un lisado, separar el lisado en capas que comprende una capa acuosa y una capa de emulsión lipídica: acuosa, y extraer los lípidos de la capa de emulsión.

En diversas modalidades del método de extracción de lípido a partir de un microorganismo, el microorganismo se pone en contacto con el ácido para producir una concentración de ácido de 5-200 mN, y el contacto del microorganismo se puede realizar por encima de 25 °C. En diversas modalidades, un solvente orgánico se adiciona al microorganismo o lisado antes, simultáneamente con, o después de contactar el microorganismo con el ácido. En algunos casos, el microorganismo se pone en contacto con un ácido a un pH de no más de 4, a un pH de no más de 3, o a un pH de no más de 2. En algunas modalidades, además de contactar el microorganismo con el ácido, se utilizan también uno o más métodos adicionales de lisis del microorganismo. En algunos casos, el contacto del microorganismo con el ácido se realiza a una temperatura de 50-160°C. En otros casos, la etapa de contacto del microorganismo con el ácido se realiza a una temperatura de

20-65°C. En una modalidad, el lípido se extrae a partir de la emulsión mediante el contacto de la emulsión con un solvente orgánico, a través de lo cual el lípido se divide de la emulsión en el solvente orgánico. En algunos casos, la separación del lisado se realiza por el enfriamiento de la emulsión por debajo de 25°C, por debajo de 10°C, o hasta una temperatura
 5 igual o inferior a 0°C, a través de lo cual una capa lipídica se separa de la capa de emulsión. En una modalidad, el método comprende, además, centrifugar la emulsión después del enfriamiento para separar la capa de lípido. En algunas modalidades, el método comprende, además, separar la emulsión de la fase acuosa antes de separar los lípidos de la emulsión. En una modalidad, el lípido se extrae de la emulsión lipídica:acuosa sin el uso de un solvente
 10 orgánico.

En otro aspecto, la presente invención se dirige a un método para extraer lípido a partir de un microorganismo al lisar un microorganismo que contiene un lípido para producir un lisado que comprende una emulsión lipídica:acuosa, y enfriar la emulsión por debajo de 25 °C para separar el lípido de la emulsión.

15 En otro aspecto, la presente invención se dirige a un método para extraer lípido a partir de la biomasa microbiana generada mediante el cultivo de un microorganismo que produce un lípido. La extracción comprende lisar los microorganismos en la biomasa para producir un lisado que comprende una emulsión lipídica:acuosa. En algunos casos, la separación de una capa lipídica a partir de la emulsión comprende desestabilizar la emulsión
 20 con la adición de un surfactante a la emulsión. En algunos casos, el lisado se puede tratar con un solvente orgánico durante un período de tiempo suficiente para permitir a los lípidos de los microorganismos solubilizarse en el solvente orgánico, y el lisado se puede separar en dos o más capas, incluyendo una capa lipídica: solvente orgánico y al menos una capa acuosa. En algunos casos, el método puede incluir, además, eliminar la composición lípido:solvente
 25 orgánico de las otras capas. Opcionalmente, la composición lípido:solvente orgánico de las otras capas se puede realizar sin separar sustancialmente los lípidos del solvente orgánico. En algunos casos, el método puede incluir, además, la transesterificación de la composición lípido:solvente orgánico para producir un alquil éster de ácido graso, o tratar la composición lípido:solvente orgánico para producir un alcano de cadena lineal. En este último caso, el
 30 método puede incluir, además, craquear el alcano de cadena lineal.

En otro aspecto, la presente invención se dirige a una composición de ésteres de ácidos grasos que comprende ésteres derivados de un microorganismo, y ésteres derivados de

un aceite no microbiano. En algunas modalidades, no más de 20%, no más de 10%, o no más de 6%, de los ésteres en la composición se derivan de un aceite no microbiano. En algunos casos, los ésteres se seleccionan del grupo que consiste en ésteres de metilo, etilo y alquilo. En algunas modalidades, el microorganismo cultivado no se ha separado del medio líquido
 5 usado para cultivar el microorganismo, y/o está presente en una proporción de menos de 1:1 v/v de microorganismo cultivado al medio líquido extracelular.

En diversas modalidades del método de extracción de lípido a partir de un microorganismo, la emulsión se enfría por debajo de 10 °C, o de una temperatura igual o inferior a 0 °C. En una modalidad, el método comprende, además, centrifugar el lisado para
 10 producir capas que comprenden la emulsión y una capa acuosa. En algunos casos, el método comprende, además, separar la emulsión de la fase acuosa. En una modalidad, el lípido se extrae de la emulsión lipídica: acuosa sin el uso de un solvente orgánico. En algunos casos, el método comprende, además, tratar el lisado con un solvente orgánico durante un periodo de tiempo suficiente para permitir a los lípidos de los microorganismos solubilizarse en el
 15 solvente orgánico.

Los microorganismos útiles de acuerdo con la presente invención pueden seleccionarse del grupo de microorganismos que consisten en una bacteria, una canobacteria, una microalga eucariota, una levadura oleaginosa, y un hongo. En algunos casos, los microorganismos pueden seleccionarse de las Tablas 1, 2 ó 3. Tales microorganismos
 20 incluyen microorganismos del género *Chlorella*. En una modalidad, el microorganismo es *Chlorella protothecoides*. En algunos casos, los microorganismos de la presente invención producen un lípido que comprende al menos 10% p/v del lípido celular total como C18 triacilgliceroles, o al menos 10% p/v del lípido celular total como C16 triacilgliceroles. En algunos casos, los microorganismos producen un lípido que comprende al menos 10% p/v del
 25 lípido celular total como C14 triacilgliceroles, al menos 10% p/v del lípido celular total como C12 triacilgliceroles, o al menos 10% p/v del lípido celular total como C10 triacilgliceroles. En algunos casos, los microorganismos de la presente invención tienen al menos 85% de identidad de la secuencia genómica del ARNr 23s con una o más secuencias seleccionadas del grupo que consiste en sec. con núms. de ident.:7-33.

[0003] En algunos casos, el microorganismo es del género *Prototheca*. En una modalidad, el microorganismo es *Prototheca moriformis*. En algunos casos, los microorganismos de la presente invención tienen al menos 75% de identidad de secuencia

con una o más secuencias seleccionadas del grupo que consiste en sec. con núms. de ident.: 33, 34, 35, 16, 36, 17, 37, 38, 39, y 32.

En varias modalidades de acuerdo con la presente invención, el lisado de los microorganismos cultivados se realiza a un pH ácido, un pH ácido no mayor que 5, un pH ácido no mayor que 4, un pH ácido no mayor que 3, o un pH ácido no mayor que 2. En algunas modalidades, el lisado de los microorganismos cultivados se realiza a un pH de al menos 9. En algunos casos, lisar el microorganismo cultivado comprende uno o más métodos de lisis seleccionados de la lista que consiste en calentamiento, sonicación, lisis mecánica, choque osmótico, oscilación de presión, expresión de un gen de autólisis, exposición a un pH por encima de 8, exposición a pH por debajo de 6, y digestión con una enzima. En algunos casos, la lisis del microorganismo comprende lisis ácida y calentamiento. En una modalidad, el microorganismo se lisa por la digestión con una enzima de degradación de polisacáridos, que puede ser una enzima de degradación de polisacáridos a partir de *Chlorella* o un virus *Clorella*. En algunos casos, el microorganismo se lisa por la digestión con una proteasa. En otros casos, el microorganismo se digiere con una combinación de al menos una proteasa y al menos una enzima de degradación de polisacáridos. En algunos casos, la proteasa es alcalasa y/o la enzima de degradación de polisacáridos es Mannaway. En algunos casos, se usan otras combinaciones de lo anterior, por ejemplo, el contacto de la biomasa con una proteasa y unas enzimas que degradan polisacáridos en combinación con el calentamiento de la biomasa a una temperatura de al menos 30 grados Celsius.

En diversas modalidades de acuerdo con la presente invención, el tratamiento del lisado comprende tratar con más de aproximadamente 5% v/v de un solvente orgánico al lisado. En otros casos, la etapa de tratamiento comprende tratar con más de un 6% v/v de un solvente orgánico al lisado, o con más de un 7% v/v de un solvente orgánico al lisado. Aún en otros casos, el v/v del solvente orgánico al lisado es mayor que entre aproximadamente 5% y mayor que aproximadamente 25%. En una modalidad, el tratamiento del lisado se puede facilitar mediante la agitación del lisado.

En modalidades de la presente invención el solvente orgánico puede comprender un aceite. El aceite puede seleccionarse del grupo que consiste en aceite de soja, colza, canola, palma, almendra de palma, coco, maíz, residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de buey, sebo porcino, microalgas, macroalgas, *Cuphea*, lino, cacahuete, grasa blanca selecta, manteca de cerdo, *Camelina sativa*, semilla de

mostaza, nuez de marañón, avena, lupino, kenaf, caléndula, cáñamo, café, semilla de lino, avellana, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez, jojoba, jatrofa, macadamia, nueces del Brasil, aguacate, petróleo, o una fracción destilada de cualquiera de los aceites anteriores. En algunos casos el aceite es aceite de soja. En algunos casos el aceite es aceite de palma. En otros casos el aceite es aceite de coco. En aun otros casos el aceite es aceite de canola. En aun otros casos el aceite es aceite de jatrofa. En una modalidad, el tiempo suficiente para permitir que el lípido del microorganismo se solubilice en el solvente orgánico es entre 0.1 y 30 minutos.

10 En diversas modalidades de acuerdo con la presente invención, la separación del lisado comprende la centrifugación del lisado tratado, a través de lo cual se separa el lisado en una capa ligera que comprende la composición lípido:solvente orgánico y una capa pesada que comprende la composición acuosa y, opcionalmente, una composición emulsionada y/o composición de sedimento celular. En algunos casos, separar el lisado comprende el
15 asentamiento del lisado tratado, a través de lo cual se separa el lisado en una capa ligera que comprende la composición lípido:solvente orgánico y una capa pesada que comprende la composición acuosa y, opcionalmente, una composición emulsionada y/o composición de sedimento celular. La separación del lisado puede incluir la reducción de la temperatura de la mezcla por debajo de 25 °C, por debajo de 10°C, o una temperatura igual o inferior a 0°C. En
20 algunos aspectos de la invención, las capas comprenden, además, una capa de emulsión lipídica: acuosa y/o un sedimento celular.

En algunos casos, los métodos de la invención comprenden, además, eliminar la composición lípido:solvente orgánico de la(s) otra(s) capa(s). En una modalidad, la eliminación se realiza sin separar sustancialmente los lípidos del solvente orgánico.

25 En algunos métodos de la invención, el microorganismo se produce en un proceso de cultivo y, opcionalmente después se almacena durante un período de tiempo entre la terminación del proceso de cultivo y el cometido de etapas adicionales para la lisis del microorganismo. En algunos casos, lisar los microorganismos en la biomasa produce un lisado que comprende una emulsión lipídica: acuosa. En algunos casos, el microorganismo se
30 almacena durante al menos una hora entre la terminación del proceso de cultivo y el cometido de etapas adicionales para la lisis del microorganismo cultivado. En algunos casos, el microorganismo se almacena durante al menos veinticuatro horas entre la terminación del

proceso de cultivo y el cometido de etapas adicionales para la lisis del microorganismo cultivado. En algunos casos, el microorganismo se almacena durante al menos cuarenta y ocho horas entre la terminación del proceso de cultivo y el cometido de etapas adicionales para la lisis del microorganismo cultivado. En otros casos, se almacena el caldo de fermentación. En algunos casos, el caldo de fermentación se concentra, por ejemplo, por centrifugación o filtración, y las células se resuspenden en un medio acuoso antes de su almacenamiento. En algunos casos el medio acuoso es agua desionizada o destilada.

En algunas modalidades de la presente invención, los microorganismos (biomasa) preparados en un proceso de cultivo se almacenan opcionalmente a una temperatura por debajo de 15 grados Celsius entre la terminación del proceso de cultivo y el cometido de etapas adicionales acometidas para la lisis del microorganismo cultivado. En algunos casos, la biomasa se almacena a una temperatura por debajo de 5 grados Celsius entre la terminación del proceso de cultivo y el cometido de las etapas adicionales para lizar los microorganismos cultivados. En algunos casos, el microorganismo se almacena a una temperatura por encima de 30 grados Celsius entre la terminación del proceso de cultivo y el cometido de etapas adicionales para la lisis del microorganismo cultivado. En algunos casos, el microorganismo se almacena a una temperatura por encima de 40 grados Celsius entre la terminación del proceso de cultivo y el cometido de etapas adicionales para la lisis del microorganismo cultivado. En algunos casos, el microorganismo se somete a agitación durante el almacenamiento.

BREVE DESCRIPCIÓN DE LAS FIGURAS

La Figura 1 representa dos perfiles de lípidos de los aceites puros preparados de acuerdo con el método de la presente invención. Se muestran los ácidos grasos C10:1, C10, C12, C18:3, C14, C18:2, C16:6, C18:1 y C18:0 de aceite puro de *Chlorella* (barras sombreadas ligero) y aceite de coco (barras sombreadas oscuro). La mayoría del aceite puro de *Chlorella* comprende ácidos grasos C18:1 (59%), mientras que la mayoría del aceite de coco puro comprende ácidos grasos C12 (54%). Los detalles se describen en el Ejemplo 6.

La Figura 2 representa la proporción de ácidos grasos C18:1 a C12 y ácidos grasos C18:1 a C14 en la mezcla de aceite puro de algas/coco. Los detalles se describen en el Ejemplo 6.

[0004] La Figura 3 representa los resultados del tratamiento químico/térmico de *Chlorella protothecoides*. Los detalles se describen en el Ejemplo 7.

La Figura 4 representa una imagen de la capa de aceite recuperada a partir de una emulsión congelada, tratada con calor. Los detalles se describen en el Ejemplo 7 (Tratamiento enzimático #1).

La Figura 5 representa un resultado de un análisis de control por cromatografía en capa delgada (TLC), muestras de aceite generadas básica y ácida. Línea 1, Referencia @ 100 µg/cada una (FAME/TAG/FFA); Línea 2, sin tratamiento químico; Línea 3, tratamiento con 120 mN H₂SO₄; Línea 4, tratamiento con 160 mN KOH; Línea 5, sin tratamiento químico; Línea 6, tratamiento con 120 mN H₂SO₄; Línea 7, tratamiento con 160mN de KOH. FAME, metil éster de ácido graso, TAG, triacilglicéridos, FFA, ácidos grasos libres, DAG, diacilglicerol (1,3-dioleína); MAG, monoacilglicerol (1-monoooleína) Los detalles se describen en el Ejemplo 7.

La Figura 6 representa una imagen de las capas recuperadas a partir de la enzima tratada (tubo a la izquierda) contra al material de cultivo celular sin tratar (tubo a la derecha). Los detalles se describen en el Ejemplo 7 (Tratamiento enzimático #2).

[0005] La Figura 7 muestra los mapas de los casetes usados en las transformaciones de *Prototheca*, como se describe en el Ejemplo 9.

[0006] La Figura 8 muestra los resultados del análisis de transferencia de tipo Southern de tres transformantes de la cepa UTEX 1435, como se describe en el Ejemplo 9.

[0007] La Figura 9 muestra un esquema del constructo transgénica del codón optimizado y del codón no optimizado *suc2* (sacarosa invertasa de levadura (yInv)). Se indican los sitios de restricción relevantes del clonaje y las flechas indican la dirección de la transcripción.

DESCRIPCIÓN DETALLADA DE LA INVENCION

I. Definiciones

[0008] A menos que se defina de otra manera, todos los términos técnicos y científicos usados en la presente descripción tienen el significado comúnmente entendido por una persona experta en la materia a la que pertenece esta invención. Las siguientes referencias proporcionan al experto una definición general de muchos de los términos usados en esta invención: Singleton y otros, *Dictionary of Microbiology and Molecular Biology* (2da edición, 1994); *The Cambridge Dictionary of Science and Technology* (Ediciones Walker,

1988); *The Glossary of Genetics*, 5ta Edición, R. Rieger y otros (editores), Springer Verlag (1991); y Hale & Marham, *The Harper Collins Dictionary of Biology* (1991). Como se usa en la presente descripción, los siguientes términos tienen el significado atribuido a ellos a menos que se especifique de cualquier otra forma.

5 Como se usa con referencia a un ácido nucleico, la frase “activo en las microalgas” se refiere a un ácido nucleico que es funcional en microalgas. Por ejemplo, un promotor que se ha usado para manejar un gen de resistencia a los antibióticos para impartir resistencia antibiótica a una microalga transgénica está activo en las microalgas. Los ejemplos de promotores activos en las microalgas son promotores endógenos para determinadas especies
10 de algas y promotores encontrados en los virus de plantas.

[0009] Como se usa en la presente descripción, una “proteína portadora de acilos” o “ACP”, es una proteína que se une a una cadena acilo en crecimiento durante la síntesis de ácidos grasos como un tiol éster al tiol distal de la porción de 4-fosfopantoteína y comprende un componente del complejo de la ácido graso sintasa.

● 15 [0010] Una “molécula de acil-CoA” o “acil-CoA”, es una molécula que comprende una porción acilo covalentemente unida a la coenzima A a través de un enlace tiol éster en el tiol distal de a porción de 4-fosfopantoteína de la coenzima A.

 Como se usa en la presente descripción, el término “alquilo” se refiere a un radical hidrocarburo lineal o ramificado, y puede incluir radicales di- y multivalentes, que tienen el
20 número de átomos de carbono designado (es decir C1-C10 significa de uno a diez carbonos). Los ejemplos de radicales hidrocarburo saturados incluyen los grupos tales como metilo, etilo, n-propilo, isopropilo, n-butilo, t-butilo, isobutilo, sec-butilo, homólogos e isómeros de, por ejemplo, n-pentilo, n-hexilo, n-heptilo, n-octilo, y similares.

● 25 Como se usa en la presente descripción, el término “alquenilo” se refiere a un grupo alquilo insaturado que tiene uno o más enlaces dobles. Los ejemplos de grupos alquenilo incluyen vinilo, 2-propenilo, crotilo, 2-isopentenilo, 2-(butadienilo), 2,4-pentadienilo y 3-(1,4-pentadienilo), y los homólogos e isómeros superiores.

 Como se usa en la presente descripción, el término “alquinilo” se refiere a un grupo alquilo insaturado que tiene uno o más enlaces triples. Los ejemplos de grupos alquinilo
30 incluyen etinil (acetilenilo), 1-propinilo, 1- y 2-butinilo, y los homólogos e isómeros superiores.

Como se usa en la presente descripción, la frase “composición acuosa o emulsionada”, se refiere a la biomasa microbiana que contiene lípidos.

[0011] Como se usa en la presente descripción, la frase “porcentaje del área”, se refiere al área de picos observados usando los métodos de detección FAME GC/FID, en los que cada ácido graso en la muestra se convierte en un metil éster de ácido graso (FAME) antes de la detección. Por ejemplo, un pico separado se observa para un ácido graso de 14 átomos de carbono sin insaturaciones (C14:0) comparado con cualquier otro ácido graso, por ejemplo un C14:1. El área del pico para cada clase de FAME es directamente proporcional a su porcentaje de la composición en la mezcla y se calcula en base a la suma de todos los picos presentes en la muestra (*es decir* [área del pico específica/área total de todos los picos medidos] x 100). Al referirse a los perfiles lipídicos de los aceites y las células de la invención, “al menos 4% de C8-C14”, por ejemplo, significa que al menos el 4% de los ácidos grasos totales en la célula o en la composición del glicerolípidos extraídos, tiene una longitud de cadena que incluye 8, 10, 12 ó 14 átomos de carbono.

Como se usa en la presente descripción, el término “arilo” se refiere a un sustituyente hidrocarburo aromático poliinsaturado que tiene 5-12 miembros del anillo, que puede ser un anillo sencillo o múltiples anillos (hasta tres anillos) los cuales están fusionados o unidos covalentemente. Los ejemplos no limitantes de grupos arilo incluyen fenilo, 1-naftilo, 2-naftilo, 4-bifenilo, y bencilo. Otros grupos arilo son útiles, además, en la presente invención.

Como se usa en la presente descripción, el término “axénico” se refiere a un cultivo de un microorganismo que está libre de contaminación por otros organismos vivos.

Como se usa en la presente descripción, el término “base” se refiere a cualquier compuesto cuya pKa es mayor que la del agua.

Como se usa en la presente descripción, el término “biodiesel” se refiere a un éster de ácido graso producido a partir de la transesterificación de un lípido.

Como se usa en la presente descripción, el término “biomasa” se refiere al material producido por el crecimiento y/o propagación de las células. La biomasa puede contener células y/o contenidos intracelulares, así como material extracelular. El material extracelular incluye, pero no se limitan a, los compuestos secretados por una célula.

Como se usa en la presente descripción, el término "biorreactor" se refiere a un recinto o recinto parcial en el cual las células, por ejemplo, los microorganismos, se cultivan, opcionalmente en suspensión.

5 Como se usa en la presente descripción, el término "catalizador" se refiere a un agente, tal como una molécula o complejo macromolecular, capaz de facilitar o promover una reacción química de un reactivo a un producto sin convertirse en una parte del producto. Un catalizador aumenta así la velocidad de una reacción, después de lo cual, el catalizador puede actuar sobre otro reactante para formar el producto. Un catalizador generalmente disminuye la energía de activación total requerida para la reacción, de modo que esta ocurre más
10 rápidamente o a una temperatura más baja. Así, un equilibrio de la reacción puede lograrse más rápidamente. Los ejemplos de catalizadores incluyen enzimas, las cuales son catalizadores biológicos, y calor, el cual es un catalizador no biológico.

15 Como se usa en la presente descripción, el término "material celulósico" significa los productos de la digestión de celulosa, tales como glucosa, xilosa, arabinosa, disacáridos, oligosacáridos, lignina, furfurales, y otras moléculas.

[0012] Como se usa en la presente descripción, el término "co-cultivo", y variantes de este como "co-cultivar", definen la presencia de dos o más tipos de células en el mismo biorreactor. Los dos o más tipos de células pueden ser ambos microorganismos, como las microalgas, o pueden ser una célula de microalgas cultivada con un tipo celular diferente.
20 Las condiciones del cultivo pueden ser aquellas que fomentan el crecimiento y/o la propagación de los dos o más tipos celulares o aquellas que facilitan el crecimiento y/o proliferación de una, o un subconjunto de las dos o más células, mientras se mantiene el crecimiento celular para el resto.

25 Como se usa en la presente descripción, el término "cofactor" se refiere a cualquier molécula, distinta del sustrato, que es requerido por una enzima para llevar a cabo su actividad enzimática.

Un promotor "constitutivo" es un promotor que se activa bajo la mayoría de las condiciones ambientales y de desarrollo.

30 Como se usa en la presente descripción, el término "cultivado" y variantes de éste, se refieren al fomento intencional del crecimiento (incremento del tamaño celular, el contenido celular, y/o la actividad celular) y/o propagación (incremento de los números de

células vía mitosis) de una o varias células mediante el empleo de condiciones de cultivo seleccionadas. La combinación de ambos, crecimiento y propagación, puede ser nombrada como proliferación. Una o más de las células pueden ser las de un microorganismo, tal como una microalga. Los ejemplos de condiciones seleccionadas incluyen el uso de un medio
 5 definido (con características conocidas como pH, fuerza iónica, y fuente de carbono), y condiciones específicas de temperatura, tensión de oxígeno, niveles de dióxido de carbono, y crecimiento en un biorreactor. El término no se refiere al crecimiento o propagación de los microorganismos en la naturaleza o en otra parte sin la intervención humana, tal como el crecimiento natural de un organismo que finalmente se convierte en un fósil para producir el
 10 aceite crudo geológico.

Como se usa en la presente descripción, el término "cicloalquilo" se refiere a un hidrocarburo cíclico saturado que tiene 3 a 15 carbonos, y 1 a 3 anillos que se pueden fusionar o unir covalentemente. Los grupos cicloalquilo útiles en la presente invención incluyen, pero sin limitarse a, ciclopentilo, ciclohexilo, cicloheptilo y ciclooctilo. Los grupos
 15 bicicloalquilo útiles en la presente invención incluyen, pero sin limitarse a, [3.3.0]bicyclooctanilo, [2.2.2]bicyclooctanilo, [4.3.0]bicyclononano, [4.4.0]bicyclodecano (decalin), espiro[3.4]octanilo, espiro[2.5]octanilo, y demás.

Como se usa en la presente descripción, el término "cicloalquenilo" se refiere a un hidrocarburo cíclico no saturado que tiene de 3 a 15 carbonos, y de 1 a 3 anillos que se
 20 pueden fusionar o unir covalentemente. Los grupos cicloalquenilo útiles en la presente invención incluyen, pero no se limitan a, ciclopentenilo, ciclohexenilo, cicloheptenilo y ciclooctenilo. Los grupos bicicloalquenilo son útiles también en la presente invención.

Como se usa en la presente descripción, el término "citólisis" se refiere a la lisis de células en un ambiente hipotónico. La citólisis es provocada por la ósmosis excesiva, o por el
 25 movimiento del agua, hacia el interior de una célula (hiperhidratación). La membrana celular no puede resistir la presión osmótica del agua en su interior, y explota.

Como se usa en la presente descripción, el término "gen exógeno" se refiere a un ácido nucleico transformado en una célula. Una célula transformada puede definirse como una célula recombinante, dentro de la cual se puede introducir un gen(es) exógeno(s)
 30 adicional(es). El gen exógeno puede ser de una especie diferente (heterólogo), o de la misma especie (homólogo) con respecto a la célula que se transforma. En el caso de un gen homólogo, este ocupa un lugar diferente en el genoma de la célula relativa a la copia

endógena del gen. El gen exógeno puede estar presente en más de una copia en la célula. El gen exógeno puede mantenerse en una célula como una inserción en el genoma o como una molécula episomal.

5 Como se usa en la presente descripción, el término "proporcionado exógenamente" en el contexto de cultivar una célula, se refiere a una molécula proporcionada a un medio de cultivo de un cultivo celular.

[0013] Como se usa en la presente descripción, los términos "vector de expresión" o "constructo de expresión" se refiere a un constructo de ácido nucleico, generado recombinante o sintéticamente, con una serie de elementos de ácido nucleico especificados
10 que permiten la transcripción de un ácido nucleico particular en una célula huésped. El vector de expresión puede ser parte de un plásmido, virus, o fragmento de ácido nucleico. Típicamente, el vector de expresión incluye un ácido nucleico para ser transcrito operativamente unido a un promotor.

15 [0014] Como se usa en la presente descripción, una "acil graso-ACP tioesterasa" define a una enzima que cataliza la ruptura de un ácido graso de una proteína portadora de acilos (ACP) durante la síntesis de los lípidos.

[0015] Como se usa en la presente descripción, una "acil-graso CoA/aldehído reductasa" define a una enzima que cataliza la reducción de una molécula de acil-CoA a un alcohol primario.

20 Como se usa en la presente descripción, una "acil graso-CoA reductasa" es una enzima que cataliza la reducción de una molécula de acil-CoA a un aldehído.

[0016] Como se usa en la presente descripción, una "aldehído graso decarbonilasa" es una enzima que cataliza la conversión de un aldehído graso a un alcano.

25 Como se usa en la presente descripción, una "aldehído graso reductasa" es una enzima que cataliza la conversión de un aldehído a un alcohol primario.

Como se usa en la presente descripción, el término "fuente fija de carbono" significa molécula(s) que contienen carbono, preferentemente orgánico, que están presentes a temperatura y presión ambiental en forma sólida o líquida.

30 Como se usa en la presente descripción, el término "hongo," significa organismos heterótrofos caracterizados por una pared celular from quitinosa del reino de los hongos.

Como se usa en la presente descripción, el término "heteroarilo" se refiere a un sustituyente hidrocarburo aromático poliinsaturado, que tiene 5-12 miembros del anillo, el cual puede ser un anillo sencillo o múltiples anillos (hasta tres anillos) los cuales están fusionados o unidos covalentemente, y el cual tiene al menos un heteroátomo en el anillo, tal como N, O, o S. Un grupo heteroarilo puede estar unido al resto de la molécula a través de un heteroátomo. Los ejemplos no limitantes de grupos heteroarilo incluyen 1-pirrolilo, 2-pirrolilo, 3-pirrolilo, 3-pirazolilo, 2-imidazolilo, 4-imidazolilo, pirazinilo, 2-oxazolilo, 4-oxazolilo, 2-phenyl-4-oxazolilo, 5-oxazolilo, 3-isoxazolilo, 4-isoxazolilo, 5-isoxazolilo, 2-tiazolilo, 4-tiazolilo, 5-tiazolilo, 2-furilo, 3-furilo, 2-tienilo, 3-tienilo, 2-piridilo, 3-piridilo, 4-piridilo, 2-pirimidilo, 4-pirimidilo, 5-benzotiazolilo, purinilo, 2-benzimidazolilo, 5-indolilo, 1-isoquinolilo, 5-isoquinolilo, 2-quinoxalinilo, 5-quinoxalinilo, 3-quinolilo, y 6-quinolilo. Los grupos heteroarilo adicionales útiles en la presente invención incluyen piridilo N-óxido, tetrazolilo, benzofuranilo, benzotienilo, indazolilo, o cualquiera de los radicales sustituidos, especialmente mono- o di-sustituido.

Como se usa en la presente descripción, el término "heteroátomo" significa cualquier átomo que no es carbono o hidrógeno. Los ejemplos de heteroátomos incluyen magnesio, calcio, potasio, sodio, azufre, fósforo, hierro y cobre.

Como se usa en la presente descripción, el término "heterocicloalquilo" se refiere a un hidrocarburo cíclico saturado que tiene 3 a 15 miembros del anillo, y 1 a 3 anillos que pueden estar fusionados o unidos covalentemente, y el cual tiene al menos un heteroátomo en el anillo, tal como N, O, o S. Además, un heteroátomo puede ocupar la posición en la que el heterociclo se une al resto de la molécula. Los ejemplos de heterocicloalquilo incluyen 1 (1,2,5,6-tetrahidropiridilo), 1-piperidinilo, 2-piperidinilo, 3-piperidinilo, 4-morfolinilo, 3-morfolinilo, tetrahydrofuran-2-ilo, tetrahydrofuran-3-ilo, tetrahidrotien-2-ilo, tetrahidrotien-3-ilo, 1-piperazinilo, 2-piperazinilo, y similares.

Como se usa en la presente descripción, en el contexto de la biomasa, el término "homogenato" significa la biomasa que ha sido físicamente desintegrada.

Como se usa en la presente descripción, "hidrocarburo" se refiere a: (a) una molécula que contiene solamente átomos de hidrógeno y carbono donde los átomos de carbono se unen covalentemente para formar la cadena principal lineal, ramificada, cíclica o parcialmente cíclica a los que los átomos de hidrógeno se unen, o (b) una molécula que sólo contiene principalmente hidrógeno y átomos de carbono y que se puede convertir para

contener sólo átomos de hidrógeno y carbono por una a cuatro reacciones químicas. Los ejemplos no limitativos de esto último incluyen hidrocarburos que contienen un átomo de oxígeno entre un átomo de carbono y un hidrógeno para formar una molécula de alcohol, así como aldehídos que contienen un solo átomo de oxígeno. Los métodos para la reducción de

5 alcoholes a hidrocarburos que contienen sólo átomos de carbono e hidrógeno son bien conocidos. Otro ejemplo de un hidrocarburo es un éster, en el que un grupo orgánico reemplaza a un átomo de hidrógeno (o más de uno) en un ácido de oxígeno. La estructura molecular de los compuestos hidrocarburos varía de la más simple, en forma de metano (CH_4), que es un constituyente del gas natural, a una muy pesada y compleja, tales como

10 algunas moléculas como los asfaltenos encontrados en el aceite crudo, petróleo, y betún. Los hidrocarburos pueden encontrarse en forma gaseosa, líquida, o sólida, o en cualquier combinación de estas formas, y pueden tener uno o más dobles o triples enlaces entre los átomos del carbono adyacentes a la cadena principal. Por lo tanto, el término incluye alcanos, alquenos, lípidos y parafinas lineales, ramificadas, cíclicas o parcialmente cíclicas.

15 Los ejemplos incluyen propano, butano, pentano, hexano, octano, escualeno y carotenoides.

Como se usa en la presente descripción, el término "enzimas modificadoras de hidrocarburos" se refiere a una enzima que altera la estructura covalente de un hidrocarburo. Los ejemplos de enzimas modificadoras de hidrocarburos incluyen una lipasa, una acil graso ACP tioesterasa, una acil graso-CoA/aldehído reductasa, una acil graso-CoA reductasa, una

20 aldehído graso reductasa, y una aldehído graso decarbonilasa.

Como se usa en la presente descripción, el término "proporción hidrógeno:carbono" se refiere a la proporción de átomos de hidrógeno a átomos de carbono en una molécula en una base átomo-a-átomo. La relación puede usarse para referirse al número de átomos de carbono y de hidrógeno en una molécula de hidrocarburo. Por ejemplo, el hidrocarburo con

25 la relación más alta es el metano CH_4 (4:1).

Como se usa en la presente descripción, el término "fracción hidrófoba" se refiere a una porción, o fracción, de un material que es más soluble en una fase hidrófoba en comparación con una fase acuosa. Un fragmento hidrófobo es substancialmente insoluble en agua y normalmente es no polar.

30 [0017] Como se usa en la presente descripción, un "incremento del rendimiento lipídico" se refiere al aumento en la productividad de un cultivo microbiano, por ejemplo, incrementando el peso seco de células por litro de cultivo, aumentando el porcentaje

de células que constituyen los lípidos, o incrementando la cantidad total de lípidos por litro de volumen de cultivo por unidad de tiempo.

[0018] Un “promotor inducible”, es un promotor que media la transcripción de los genes operativamente vinculados en respuesta a un estímulo particular.

5 Como se usa en la presente descripción, el término “in situ” significa “en el lugar” o “en su posición original.” Por ejemplo, un cultivo puede contener una primera microalga que secreta un catalizador y un segundo microorganismo que secreta un sustrato, donde los tipos de células primera y segunda producen los componentes necesarios para que ocurra una reacción química particular in situ en el co-cultivo sin requerir la separación adicional o el
10 procesamiento de los materiales.

Como se usa en la presente descripción, el término “isómeros” se refiere a compuestos de la presente invención que poseen átomos de carbonos asimétricos (centros ópticos) o dobles enlaces. Se pretende que todos los racematos, diasterómeros, isómeros geométricos y los isómeros individuales estén incluidos dentro del alcance de la presente
15 invención.

Como se usa en la presente descripción, la frase “concentración limitante de un nutriente” se refiere a una concentración en un cultivo que limita la propagación de un organismo cultivado. Una concentración “no limitante de un nutriente” es una concentración que ayuda a la propagación máxima durante un período dado del cultivo. Así, el número de
20 células producidas durante un período dado del cultivo es más bajo en presencia de una concentración limitante de un nutriente que cuando el nutriente es no limitante. Se dice que un nutriente está “en exceso” en un cultivo, cuando el nutriente está presente en una concentración mayor que la necesaria para la propagación máxima.

Como se usa en la presente descripción, una “lipasa” es una enzima que cataliza la
25 hidrólisis de los enlaces éster en sustratos lípidos insolubles en agua. Las lipasas catalizan la hidrólisis de lípidos en glicerol y ácidos grasos.

[0019] Los “lípidos”, son una clase de moléculas que son solubles en solventes no polares (como el éter y el cloroformo) y son relativamente o completamente insolubles en agua. Las moléculas de lípidos tienen estas propiedades porque están
30 constituidas por largas colas de hidrocarburos que son de naturaleza hidrófoba. Los ejemplos de lípidos incluyen los ácidos grasos (saturados e insaturados); glicéridos o glicerolípidos

(como monoglicéridos, diglicéridos, triglicéridos o grasas neutras y fosfoglicéridos o glicerofosfolípidos); no glicéridos (esfingolípidos, esterol lípidos incluyendo colesterol y hormonas esteroideas, lípidos de prenol incluyendo terpenoides, alcoholes grasos, ceras, y poliquétidos); y derivados de lípidos complejos (lípidos unidos a azúcares, o glicolípidos, y lípidos unidos a proteínas). Las “grasas” son un subgrupo de lípidos denominados “triacylglicéridos”.

Como se usa en la presente descripción, la frase “composición lípido:solvente orgánico” se refiere a una mezcla de lípido y solvente orgánico.

Como se usa en la presente descripción, una “enzima de la vía lipídica” es cualquier enzima que desempeña un papel en el metabolismo de lípido, es decir, ya sea síntesis de lípido modificación o degradación. Este término incluye las proteínas que modifican químicamente los lípidos, así como las proteínas portadoras.

Como se usa en la presente descripción, el término “lisado” se refiere a una solución que contiene el contenido de las células lisadas.

[0020] Como se usa en la presente descripción, el término “lisis” se refiere a la rotura de la membrana plasmática y opcionalmente de la pared celular de un organismo biológico, que es suficiente para liberar al menos algún contenido intracelular, a menudo por mecanismos mecánicos, virales u osmóticos que comprometen la integridad de la célula.

Como se usa en la presente descripción, el término “lisis” se refiere a la ruptura de la membrana celular y, opcionalmente, de la pared celular de un organismo o célula biológica lo suficiente para liberar al menos algún contenido intracelular.

Como se usa en la presente descripción, el término “microalgas” significa un organismo microbiano que es ya sea (a) eucariota y contiene un cloroplasto o remanente de cloroplasto, o (b) es una cianobacteria. Las microalgas incluyen los fotoautótrofos obligados, que no pueden metabolizar una fuente fija de carbono en energía, como lo hacen los heterótrofos, los cuales pueden vivir únicamente a costa de una fuente fija de carbono. Las microalgas se pueden referir a los organismos unicelulares que se separan de las células hermanas poco después de la división celular, tales como *Chlamydomonas*, y se puede referir también a los microbios, tales como, por ejemplo, *Volvox*, que es un microbio fotosintético multicelular simple de dos tipos distintos de células. Las “Microalgas” se pueden referir también a las células tales como *Chlorella* y *Dunaliella*. Las “Microalgas” también incluyen

otros organismos fotosintéticos microbianos que presentan adhesión célula-célula, tales como *Agmenellum*, *Anabaena*, y *Pyrobotrys*, así como los organismos que contienen estructuras similares al cloroplasto que ya no son capaces de realizar fotosíntesis, tales como las microalgas del género *Prototheca* y algunos dinoflagelados. Las “Microalgas” incluyen microorganismos heterótrofos obligados que han perdido la capacidad de realizar la fotosíntesis, tales como determinadas especies de dinoflagelados.

Los términos “microorganismo” y “microbio” se usan indistintamente en la presente descripción para referirse a los organismos unicelulares microscópicos.

Como se usa en la presente descripción, el término “aceite” significa una sustancia que contiene carbono no polar, hidrófoba, lipofílica, incluyendo pero no limitado a, aceite crudo derivado geológicamente, fracciones destiladas de aceite crudo derivado geológicamente, aceite vegetal, aceite de alga, y lípidos microbianos.

Como se usa en la presente descripción, el término “levadura oleaginosa,” significa la levadura que puede acumular más de 10% de su peso seco celular como lípido. La levadura oleaginosa incluye organismos tales como *Yarrowia lipolytica*, así como las cepas modificadas de levadura tales como *Saccharomyces cerevisiae* que se han modificado para acumular más del 10% del peso seco celular como lípido.

Como se usa en la presente descripción, los términos “unida operativamente”, “en unión operativa,” o equivalentes gramaticales de éstos se refieren a una unión funcional entre dos secuencias, tal como una secuencia control (típicamente un promotor) y la secuencia unida. Un promotor está en una unión operativa con un gen exógeno si puede mediar la transcripción del gen.

Como se usa en la presente descripción, el término “solvente orgánico” se refiere a un material que contiene carbono que disuelve un soluto sólido, líquido, o gaseoso, resultando en una solución.

Como se usa en la presente descripción, el término “choque osmótico” se refiere a la ruptura de las células bacterianas, algas u otras en una solución a continuación de una reducción súbita en la presión osmótica. El choque osmótico se induce a veces para liberar los componentes celulares de estas células a una solución.

Como se usa en la presente descripción, el término “fotobiorreactor” se refiere a un recipiente, al menos parte del que es al menos parcialmente transparente o parcialmente

abierto, permitiendo de ese modo que la luz pase a través, en el que, por ejemplo, se cultivan una o más células de microalgas. Los fotobiorreactores pueden ser cerrados, como en el caso de una bolsa de polietileno o frasco Erlenmeyer, o pueden ser abiertos al medio ambiente, como en el caso de un estanque al aire libre.

5 Como se usa en la presente descripción, el término “polisacárido” (también llamado “glicano”) se refiere al carbohidrato compuesto de monosacáridos unidos a la vez por uniones glicosídicas. La celulosa es un ejemplo de un polisacárido que forma las paredes celulares de ciertas plantas. La celulosa puede ser despolimerizada por enzimas para producir monosacáridos como la glucosa y la xilosa, así como disacáridos más grandes y
10 oligosacáridos.

Como se usa en la presente descripción, el término “enzima de degradación de polisacáridos” se refiere a cualquier enzima capaz de catalizar la hidrólisis, o despolimerización, de cualquier polisacárido. Por ejemplo, la celulasa cataliza la hidrólisis de la celulosa.

● 15 Como se usa en la presente descripción, el término “puerto”, en el contexto de un biorreactor, se refiere a una abertura en el biorreactor que permite la entrada o salida de materiales tales como gases, líquidos y células. Los puertos se conectan normalmente a la tubería principal del fotobiorreactor.

Un “promotor” se define como una disposición de secuencias control de ácido nucleico que dirigen la transcripción de un ácido nucleico. Como se usa en la presente
20 descripción, un promotor incluye necesariamente secuencias de ácidos nucleicos cercanas al sitio de inicio de la transcripción, como, por ejemplo, en el caso del promotor de la polimerasa de tipo II, un elemento TATA. El promotor también incluye, opcionalmente, potenciadores distales o elementos represores, que se pueden encontrar a varios miles de
● 25 pares de bases del sitio de inicio de la transcripción.

Como se usa en la presente descripción, el término “recombinante” cuando se usa con referencia a, por ejemplo, una célula, o ácido nucleico, proteína, o vector, indica que la célula, ácido nucleico, proteína o vector, se ha modificado por la introducción de un ácido nucleico heterólogo o proteína o la alteración de un ácido nucleico o proteína nativa, o que la
30 célula se deriva de una célula siempre modificada. De este modo, por ejemplo, las células recombinantes expresan genes que no se encuentran dentro de la forma nativa de la célula (no recombinante) o expresan genes nativos que son, de cualquier otra forma, expresados

anormalmente, expresados por debajo o no expresados en absoluto. Por el término “ácido nucleico recombinante” se entiende en la presente un ácido nucleico formado originalmente *in vitro*, en general, por la manipulación de ácidos nucleicos, por ejemplo, usando las polimerasas y endonucleasas, en una forma que normalmente no se encuentra en la naturaleza. De esta manera, se logra la unión operativa de las diferentes secuencias. Así, un ácido nucleico aislado, en una forma lineal, o un vector de expresión formado *in vitro* por la ligadura de moléculas de ADN que normalmente no se unen en la naturaleza, se consideran recombinantes para los propósitos de la presente invención. Se entiende que una vez que un ácido nucleico recombinante se prepara y reintroduce en una célula hospedera u organismo, se replicará de forma no recombinante, es decir, usando la maquinaria celular *in vivo* de la célula huésped en lugar de manipulaciones *in vivo*, sin embargo, dichos ácidos nucleicos, una vez producidos de forma recombinante, aunque posteriormente replicados de forma no recombinante, todavía se consideran recombinantes para los propósitos de la invención. Del mismo modo, una “proteína recombinante” es una proteína producida usando técnicas de recombinación, es decir, a través de la expresión de un ácido nucleico recombinante como se representó anteriormente.

Como se usa en la presente descripción, el término “diesel renovable” se refiere a una mezcla de alcanos (tales como, C10:0, C12:0, C14:0, C16:0 y C18:0) producidos a través de la hidrogenación y desoxigenación de lípidos. El diesel renovable incluye también el combustible diesel derivado de la biomasa como se define en la Sección 45K (c) (3), usando el proceso de despolimerización térmica (EPAAct 2005).

Como se usa en la presente descripción, el término “sonicación” se refiere a un proceso de ruptura de materiales biológicos, tal como una célula, mediante el uso de energía de onda de sonido.

Como se usa en la presente descripción, el término “aguas residuales” se refiere a los residuos acuosos que típicamente contienen el agua de lavado, desechos de lavandería, heces, orina y otros desechos líquidos o semi-líquidos. Se incluyen algunas formas de residuales urbanos, así como en segundo lugar aguas residuales tratadas.

II. GENERAL

Las solicitudes de patente de Estados Unidos núms.: 60/941,581; 60/959,174; 60/968,291; y 61/024,069 se incorporan de este modo como referencia en su totalidad para todos los propósitos.

La invención se refiere generalmente a la producción de composiciones de hidrocarburos, tales como un lípido, en los microorganismos. En particular, la invención proporciona los métodos para la extracción, recuperación, aislamiento y obtención de un lípido a partir de un microorganismo y las composiciones que comprenden el lípido. La invención también describe los métodos para la producción de composiciones de hidrocarburos para uso como biodiesel, diesel renovable, combustible de reactor, y para la producción de surfactante lipídico que tiene una longitud de cadena de carbono de C12 y C14.

La invención se basa, en parte, en la idea de que determinados microorganismos se pueden usar para producir económicamente composiciones de hidrocarburos y en grandes cantidades para uso en el transporte de combustible y la industria petroquímica, entre otras aplicaciones. Los microorganismos adecuados incluyen microalgas, levadura oleaginosa, y hongos. Un género preferido de microalgas para uso en la invención es la microalga *Chlorella* que producen lípido. La transesterificación ácida de los lípidos produce ésteres de ácidos grasos de cadena larga que pueden ser útiles como biodiesel. Otros procesos enzimáticos se pueden adaptar para producir ácidos grasos, aldehídos, alcoholes y alcanos. La presente solicitud describe los métodos de modificación genética de múltiples especies y cepas de microorganismos, incluyendo *Chlorella* y microbios similares para proporcionar a los organismos que tienen características que facilitan la producción de lípidos adecuados para su conversión en biodiesel u otros compuestos de hidrocarburos. La presente solicitud describe también los métodos de cultivo de microalgas para la productividad aumentada y rendimiento lipídico aumentado.

Los microorganismos útiles en la invención producen lípidos o hidrocarburos adecuados para la producción de biodiesel o como materia prima para las aplicaciones industriales. Los hidrocarburos adecuados para la producción de biodiesel incluyen triglicéridos (TAG), que contienen moléculas de ácidos grasos de cadena larga. Los hidrocarburos adecuados para las aplicaciones industriales, tales como la fabricación de surfactantes, incluyen ácidos grasos, aldehídos, alcoholes y alcanos. En algunas modalidades, los ácidos grasos adecuados, o los alcoholes primarios, aldehídos, o alcanos correspondientes, generados por los métodos descritos en la presente descripción, contienen al menos 8, al menos 9, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 21, al menos 22, al menos 23, al menos 24, al menos 25, al menos 26, al menos 27, al menos 28, al

menos 29, al menos 30, al menos 31, al menos 32, al menos 33, al menos 34 , o al menos 35 átomos de carbono o más. Los ácidos grasos de cadena larga preferidos para el biodiesel generalmente contienen al menos 14 átomos de carbono o más.

Los ácidos grasos preferidos, o los alcoholes primarios, aldehídos, y alcanos correspondientes, para las aplicaciones industriales contienen al menos 8 o más átomos de carbono. En ciertas modalidades, los ácidos grasos anteriores, así como las otras moléculas de hidrocarburos correspondientes, son saturados (sin enlaces carbono-carbono dobles o triples), mono insaturados (enlace doble único); poli insaturados (dos o más enlaces dobles); son lineal (no cíclico), y/o tienen poca o ninguna ramificación en sus estructuras.

Los triacilgliceroles que contienen longitudes de cadena de carbono en el intervalo de C8 a C22 se prefieren. Preferidas para los surfactantes son C10-C14. Preferidas para el biodiesel o diesel renovable son C16 a C18. Preferidas para el combustible de reactor son C8-C10. Preferidos para la nutrición son los ácidos grasos poliinsaturados C22 (tal como DHA) y carotenoides (tal como la astaxantina).

III. MICROORGANISMOS ÚTILES PARA LA PRODUCCIÓN DE LÍPIDOS

Cualquier especie de organismo que produce el lípido o hidrocarburo adecuado se puede usar, aunque se prefieren los microorganismos que producen naturalmente altos niveles de lípido o hidrocarburo adecuado. La producción de hidrocarburos por microorganismos es revisada por Metzger *y otros*, *Appl Microbiol Biotechnol* (2005) 66: 486-496 y A Look Back en el U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae, NREL/TP-580-24190, John Sheehan, Terri Dunahay, John Benemann y Paul Roessler (1998).

Las consideraciones que afectan la selección de un microorganismo para su uso en la invención incluyen, además de la producción de hidrocarburos adecuados para el biodiesel o para aplicaciones industriales: (1) alto contenido de lípidos como porcentaje del peso celular; (2) facilidad de crecimiento; (3) facilidad de manipulación por ingeniería genética; y (3) facilidad de procesamiento. En modalidades particulares, el microorganismo de tipo salvaje o genéticamente modificado produce células que son al menos: aproximadamente 40%, aproximadamente 45%, aproximadamente 50%, aproximadamente 55%, aproximadamente 60%, aproximadamente 65%, o aproximadamente 70% o más de lípido. Los organismos preferidos crecen de manera heterótrofa (en azúcar en ausencia de luz) o pueden diseñarse para hacerlo usando, por ejemplo, los métodos descritos en las solicitudes

de patente de Estados Unidos núms. 60/837,839 y 60/968,291 de propiedad mancomunada, las que se incorporan en la presente como referencia en su totalidad. La facilidad de la transformación y disponibilidad de marcadores de selección y promotores, constitutivo y/ o inducible, que son funcionales en el microorganismo afecta la facilidad de la ingeniería genética. Las consideraciones de procesamiento se pueden incluir, por ejemplo, la disponibilidad de medios eficaces para la lisis de las células.

A. Algas

En una modalidad preferida de la presente invención, un microorganismo que produce un lípido o un microorganismo a partir del cual se puede extraer, recuperar, u obtener un lípido es un alga.

Los ejemplos de algas que pueden usarse para practicar la presente invención incluyen, pero sin limitarse a, las siguientes algas enumeradas en la Tabla 1.

Tabla 1. Ejemplos de algas.

Achmanthes orientalis, *Agmenellum*, *Amphiprora hyaline*, *Amphora coffeiformis*, *Amphora coffeiformis lineae*, *Amphora coffeiformis punctata*, *Amphora coffeiformis taylori*, *Amphora coffeiformis tenuis*, *Amphora delicatissima*, *Amphora delicatissima capitata*, *Amphora sp.*, *Anabaena*, *Ankistrodesmus*, *Ankistrodesmus falcatus*, *Boekelovia hooglandii*, *Borodinella sp.*, *Botryococcus braunii*, *Botryococcus sudeticus*, *Carteria*, *Chaetoceros gracilis*, *Chaetoceros muelleri*, *Chaetoceros muelleri subsalsum*, *Chaetoceros sp.*, *Chlorella anitrata*, *Chlorella Antarctica*, *Chlorella aureoviridis*, *Chlorella candida*, *Chlorella capsulate*, *Chlorella desiccata*, *Chlorella ellipsoidea*, *Chlorella emersonii*, *Chlorella fusca*, *Chlorella fusca var. vacuolata*, *Chlorella glucotropha*, *Chlorella infusionum*, *Chlorella infusionum var. actophila*, *Chlorella infusionum var. auxenophila*, *Chlorella kessleri*, *Chlorella lobophora (cepa SAG 37.88)*, *Chlorella luteoviridis*, *Chlorella luteoviridis var. aureoviridis*, *Chlorella luteoviridis var. lutescens*, *Chlorella miniata*, *Chlorella minutissima*, *Chlorella mutabilis*, *Chlorella nocturna*, *Chlorella parva*, *Chlorella photophila*, *Chlorella pringsheimii*, *Chlorella protothecoides (incluyendo cualquiera de las cepas UTEX 1806, 411, 264, 256, 255, 250, 249, 31, 29, 25, y las cepas CCAP 211/17 y 211/8d)*, *Chlorella protothecoides var. acidicola*, *Chlorella regularis*, *Chlorella regularis var. minima*, *Chlorella regularis var. umbricata*, *Chlorella reisiigii*, *Chlorella saccharophila*, *Chlorella saccharophila var. ellipsoidea*, *Chlorella salina*, *Chlorella simplex*, *Chlorella sorokiniana*, *Chlorella sp.*, *Chlorella*

sphaerica, *Chlorella stigmatophora*, *Chlorella vanniellii*, *Chlorella vulgaris*, *Chlorella vulgaris*, *Chlorella vulgaris f. tertia*, *Chlorella vulgaris var. autotrophica*, *Chlorella vulgaris var. viridis*, *Chlorella vulgaris var. vulgaris*, *Chlorella vulgaris var. vulgaris f. tertia*, *Chlorella vulgaris var. vulgaris f. viridis*, *Chlorella xanthella*, *Chlorella zofingiensis*, *Chlorella trebouxioides*, *Chlorella vulgaris*, *Chlorococcum infusionum*, *Chlorococcum sp.*, *Chlorogonium*, *Chroomonas sp.*, *Chrysosphaera sp.*, *Cricosphaera sp.*, *Cryptomonas sp.*, *ciclotella cryptica*, *ciclotella meneghiniana*, *ciclotella sp.*, *Dunaliella sp.*, *Dunaliella bardawil*, *Dunaliella bioculata*, *Dunaliella granulate*, *Dunaliella maritime*, *Dunaliella minuta*, *Dunaliella parva*, *Dunaliella peircei*, *Dunaliella primolecta*, *Dunaliella salina*, *Dunaliella terricola*, *Dunaliella tertiolecta*, *Dunaliella viridis*, *Dunaliella tertiolecta*, *Eremosphaera viridis*, *Eremosphaera sp.*, *Ellipsoidon sp.*, *Euglena*, *Franceia sp.*, *Fragilaria crotonensis*, *Fragilaria sp.*, *Gleocapsa sp.*, *Gloeothamnion sp.*, *Hymenomonas sp.*, *Isochrysis aff. galbana*, *Isochrysis galbana*, *Lepocinclis*, *Micractinium*, *Micractinium (UTEX LB 2614)*, *Monoraphidium minutum*, *Monoraphidium sp.*, *Nannochloris sp.*, *Nannochloropsis salina*, *Nannochloropsis sp.*, *Navicula acceptata*, *Navicula biskanterae*, *Navicula pseudotenelloides*, *Navicula pelliculosa*, *Navicula saprophila*, *Navicula sp.*, *Nephrochloris sp.*, *Nephroselmis sp.*, *Nitzschia communis*, *Nitzschia alexandrina*, *Nitzschia communis*, *Nitzschia dissipata*, *Nitzschia frustulum*, *Nitzschia hantzschiana*, *Nitzschia inconspicua*, *Nitzschia intermedia*, *Nitzschia microcephala*, *Nitzschia pusilla*, *Nitzschia pusilla elliptica*, *Nitzschia pusilla monoensis*, *Nitzschia quadrangular*, *Nitzschia sp.*, *Ochromonas sp.*, *Oocystis parva*, *Oocystis pusilla*, *Oocystis sp.*, *Oscillatoria limnetica*, *Oscillatoria sp.*, *Oscillatoria subbrevis*, *Pascheria acidophila*, *Pavlova sp.*, *Phagus*, *Phormidium*, *Platymonas sp.*, *Pleurochrysis carterae*, *Pleurochrysis dentate*, *Pleurochrysis sp.*, *Prototheca wickerhamii*, *Prototheca stagnora*, *Prototheca portoricensis*, *Prototheca moriformis*, *Prototheca zopfii*, *Pyramimonas sp.*, *Pyrobotrys*, *Sarcinoid chrysophyte*, *Scenedesmus armatus*, *Spirogyra*, *Spirulina platensis*, *Stichococcus sp.*, *Synechococcus sp.*, *Tetraedron*, *Tetraselmis sp.*, *Tetraselmis suecica*, *Thalassiosira weissflogii*, y *Viridiella fridericana*

1. *Chlorella* y *Prototheca*

En una modalidad preferida de la presente invención, el microorganismo que produce un lípido o un microorganismo a partir del cual se puede extraer, recuperar, u obtener un lípido, es del género *Chlorella*, de preferencia, *Chlorella protothecoides*.

La *Chlorella* es un género de algas verdes unicelulares, que pertenecen al phylum Chlorophyta. Es en forma esférica, aproximadamente de 2 a 10 µm de diámetro, y es sin flagelos. Algunas especies de *Chlorella* son naturalmente heterotróficas.

La *Chlorella*, en particular *Chlorella protothecoides*, es un microorganismo preferido para uso en la invención, debido a su alta composición de lípido (al menos aproximadamente 55% por peso), en particular lípido de cadena larga adecuado para biodiesel. Además, esta microalga que crece heterotróficamente, se puede diseñar genéticamente como se demuestra en los Ejemplos en la presente, y se puede lisar por el virus *Chlorella*.

En una modalidad preferida de la presente invención, el microorganismo usado para la expresión de un transgen es del género *Chlorella*, preferentemente, *Chlorella protothecoides*. Los ejemplos de expresión de transgenes en, por ejemplo, *Chlorella*, pueden encontrarse en la literatura (ver, por ejemplo, *Current Microbiology* Vol. 35 (1997), pp. 356–362; Sheng Wu Gong Cheng Xue Bao. 2000 jul;16(4):443-6; *Current Microbiology* Vol. 38 (1999), pp. 335–341; *Appl Microbiol Biotechnol* (2006) 72: 197–205; *Marine Biotechnology* 4, 63–73, 2002; *Current Genetics* 39:5, 365–370 (2001); *Plant Cell Reports* 18:9, 778–780, (1999); *Biologia Plantarum* 42(2): 209–216, (1999); *Plant Pathol. J* 21(1): 13–20, (2005)). Ver, además, los ejemplos en la presente descripción. Otros microbios que producen lípido pueden diseñarse también, incluyendo los microbios procariotas (ver Kalscheuer y otros, *Applied Microbiology and Biotechnology*, volumen 52, número 4/octubre, 1999).

En otra modalidad preferida de la presente invención, el microorganismo que produce un lípido o un microorganismo a partir del cual se puede extraer, recuperar, u obtener un lípido es del género *Prototheca*, preferentemente, *Prototheca moriformis*. Las especies del género *Prototheca* son adecuadas para producción de lípidos debido a que pueden producir altos niveles de lípidos, particularmente lípidos adecuados para producción de sustancias químicas y combustibles. Los lípidos producidos por *Prototheca* tienen cadenas de ácidos grasos de menor longitud y de un mayor grado de saturación que los lípidos producidos por la mayoría de las microalgas. Además, los lípidos de *Prototheca* están generalmente libres de pigmentos (niveles bajos a indetectables de clorofila y ciertos carotenoides) y en todo caso contienen mucho menos pigmentos que lípidos de otras microalgas. Las cepas ilustrativas de *Prototheca* para su uso en los métodos de la invención incluyen *Prototheca wickerhamii*, *Prototheca stagnora* (que incluyen UTEX 327),

Prototheca portoricensis, *Prototheca moriformis* (que incluyen las cepas UTEX 1441, 1435), y *Prototheca zopfii*. Las especies del género *Prototheca* son heterótrofos obligados.

2. Identificación de especies de microalgas

Las especies de microalgas, que incluyen *Chlorella* y *Prototheca*, para su uso en la invención pueden identificarse por amplificación de ciertas regiones objetivo del genoma. Por ejemplo, la identificación de una especie específica de *Chlorella* puede lograrse por amplificación y secuenciación de ADN nuclear y/o de cloroplastos usando cebadores y la metodología descrita en Wu y otros, Identification of *Chlorella spp.* isolates using any region of the genome. Los ejemplos incluyen secuencias de ADN ribosomal, *Bot. Bull. Acad. Sin.* (2001) 42:115-121. Los métodos bien establecidos para el análisis filogenético, tales como la amplificación y secuenciación del espaciador transcrito interno ribosomal (ITS1 y ITS2 ADNr), ARNr 18S, y otras regiones conservadas del genoma pueden ser usados por aquellos expertos en la materia para identificar especies no solamente de *Chlorella* o *Prototheca*, sino también otros organismos productores de lípidos e hidrocarburo capaces de usar los métodos descritos en la presente descripción. Para ejemplos de métodos de identificación y clasificación de algas ver además, por ejemplo, *Genetics*, agosto de 2005; 170(4): 1601-10 y *RNA*, abril de 2005;11(4):361-4.

[0021] La comparación de ADN genómico puede ser empleada para identificar especies de microalgas que se usan en la presente invención. Las regiones de ADN genómico conservado, tales como, aunque no limitados al ADN que codifica para ARNr 23S, puede ser amplificada a partir de especies de microalgas y comparadas con secuencias consenso con el objetivo de examinar especies de microalgas que estén taxonómicamente relacionadas con la microalgas de preferencia usadas en la presente invención. Abajo se muestran algunos ejemplos de la comparación de las secuencias de ADN para especies del género *Chlorella*.

[0022] La comparación del ADN genómico también puede ser útil para identificar especies de microalgas que han sido mal identificadas en un colección de cepas. A menudo, una colección de cepas identificará especies de microalgas en base a sus características fenotípicas y morfológicas. El uso de estas características puede conllevar a la categorización incorrecta de las especie o del género de una microalga. El uso de la comparación del ADN genómico puede ser un mejor método de categorización de especies de microalgas basado en su relación filogenética. Los ejemplos específicos de usar los datos

del genotipado para establecer relaciones filogenéticas de cepas de microalgas posiblemente mal identificadas se describen más abajo en los Ejemplos.

En algunos casos, las microalgas que se prefieren para usar en la presente invención tienen secuencias de ADN genómico que codifican para ARNr 23s que tienen al menos 99%, al menos 98%, al menos 97%, al menos 96%, al menos 95%, al menos 94%, al menos 93%, al menos 92%, al menos 91%, al menos 90%, al menos 89%, al menos 88%, al menos 87%, o al menos 86% identidad de nucleótidos con al menos una de las secuencias enumeradas en la sec. con núms. de ident.: 7-31. En otros casos, las microalgas que se prefieren para usar en la presente invención tienen secuencias de ADN genómico que codifican para ARNr 23s que tienen al menos 85%, al menos 80%, al menos 75% al menos 70% al menos 65% o al menos 60% identidad de nucleótidos con al menos una de las secuencias enumeradas en la sec. con núms. de ident..7-31, 32, 33, 34, 35, 36, 37, 38, y 39.

B. Levadura oleaginosa

En una modalidad preferida de la presente invención, un microorganismo que produce un lípido o un microorganismo a partir del cual se puede extraer, recuperar, u obtener un lípido es una levadura oleaginosa.

Los ejemplos de levadura oleaginosa que puede usarse para practicar la presente invención incluyen, pero sin limitarse a, las siguientes levaduras oleaginosas enumeradas en la Tabla 2.

Tabla 2. Ejemplos de levadura oleaginosa.

<i>Candida apicola</i> , <i>Candida sp.</i> , <i>Cryptococcus curvatus</i> , <i>Cryptococcus terricolus</i> , <i>Debaromyces hansenii</i> , <i>Endomycopsis vernalis</i> , <i>Geotrichum carabidarum</i> , <i>Geotrichum cucujoidarum</i> , <i>Geotrichum histeridarum</i> , <i>Geotrichum silvicola</i> , <i>Geotrichum vulgare</i> , <i>Hyphopichia burtonii</i> , <i>Lipomyces lipofer</i> , <i>Lypomyces orientalis</i> , <i>Lipomyces starkeyi</i> , <i>Lipomyces tetrasporous</i> , <i>Pichia mexicana</i> , <i>Rodosporidium sphaerocarpum</i> , <i>Rhodosporidium toruloides</i> <i>Rhodotorula aurantiaca</i> , <i>Rhodotorula dairenensis</i> , <i>Rhodotorula diffluens</i> , <i>Rhodotorula glutinus</i> , <i>Rhodotorula glutinis</i> var. <i>glutinis</i> , <i>Rhodotorula gracilis</i> , <i>Rhodotorula graminis</i> <i>Rhodotorula minuta</i> , <i>Rhodotorula mucilaginoso</i> , <i>Rhodotorula mucilaginoso</i> var. <i>mucilaginoso</i> , <i>Rhodotorula terpenoidalis</i> , <i>Rhodotorula toruloides</i> , <i>Sporobolomyces alborubescens</i> , <i>Starmerella bombicola</i> , <i>Torulaspora delbruekii</i> , <i>Torulaspora pretoriensis</i> , <i>Trichosporon behrend</i> , <i>Trichosporon brassicae</i> , <i>Trichosporon domesticum</i> , <i>Trichosporon</i>

laibachii, *Trichosporon loubieri*, *Trichosporon loubieri* var. *loubieri*, *Trichosporon montevideense*, *Trichosporon pullulans*, *Trichosporon* sp., *Wickerhamomyces Canadensis*, *Yarrowia lipolytica*, y *Zygoascus meyeræ*.

En una modalidad preferida de la presente invención, el microorganismo usado para la expresión de un transgen es una levadura oleaginosa. Los ejemplos de expresión de transgenes en levadura oleaginosa (por ejemplo, *Yarrowia lipolytica*) pueden encontrarse en la literatura (ver, por ejemplo, Bordes y otros, *J Microbiol Methods*, Jun 27 (2007)).

5 **C. Otros hongos**

En una modalidad preferida de la presente invención, un microorganismo que produce un lípido o un microorganismo a partir del cual se puede extraer, recuperar, u obtener un lípido es un hongo.

10 Los ejemplos de hongos que pueden usarse para practicar la presente invención incluyen, pero sin limitarse a, los siguientes hongos enumeradas en la Tabla 3.

Tabla 3. Ejemplos de hongos no levadura.

Mortierella, *Mortierella vinacea*, *Mortierella alpine*, *Pythium debaryanum*, *Mucor circinelloides*, *Aspergillus ochraceus*, *Aspergillus terreus*, *Penicillium lilacinum*, *Hansenula*, *Chaetomium*, *Cladosporium*, *Malbranchea*, *Rhizopus*, y *Pythium*

15 En una modalidad preferida de la presente invención, el microorganismo usado para la expresión de un transgen es un hongo. Los ejemplos de expresión de transgenes en hongos (por ejemplo, *Mortierella alpine*, *Mucor circinelloides*, y *Aspergillus ochraceus*) pueden encontrarse también en la literatura (ver, por ejemplo, *Microbiology*, jul; 153(Pt. 7):2013-25 (2007); *Mol Genet Genomics*, jun; 271(5):595-602 (2004); *Curr Genet*, mar;21(3):215-23 (1992); *Current Microbiology*, 30(2):83-86 (1995); Sakuradani, NISR Research Grant, "Studies of Metabolic Engineering of Useful Lipid-producing Microorganisms" (2004); y PCT/JP2004/012021).

20 **D. Bacterias**

En una modalidad preferida de la presente invención, un microorganismo que produce un lípido o un microorganismo a partir del cual se puede extraer, recuperar, u obtener un lípido es una bacteria.

Los ejemplos de expresión de genes exógenos en bacterias, tales como *E. coli*, son bien conocidos; ver, por ejemplo, *Molecular Cloning: A Laboratory Manual*, Sambrook y otros. (3ra edición, 2001, Cold Spring Harbor Press).

IV. MODIFICACIÓN DEL MICROBIO

En ciertas modalidades de la presente invención, se desea modificar genéticamente un microorganismo. De este modo, la presente solicitud describe las cepas genéticamente modificadas de microalgas, levadura oleaginosa, bacterias, u hongos con uno o más genes exógenos para producir diversos compuestos de hidrocarburos.

Los promotores, ADNc, y 3'UTR, así como otros elementos de vectores de expresión, se pueden generar a través de técnicas de clonación usando fragmentos aislados a partir de fuentes nativas (ver, por ejemplo, *Molecular Cloning: A Laboratory Manual*, Sambrook y otros. (3ra edición, 2001, Cold Spring Harbor Press; y la patente de Estados Unidos 4,683,202). Como alternativa, los elementos se pueden generar sintéticamente usando métodos conocidos (ver, por ejemplo, *Gene*, 16 de octubre de 1995; 164 (1) :49-53).

A. Optimización del codón para la expresión

El ADN que codifica un polipéptido que se expresa en un microorganismo, por ejemplo, una lipasa y marcador de selección son de preferencia ADNc con codón optimizado. Los métodos de grabación para la expresión de genes en las microalgas se describen en la patente de Estados Unidos 7,135,290. Información adicional para la optimización del codón está disponible, por ejemplo en la base de datos de uso de codones del GenBank. Como ejemplos no limitantes, el uso de codones en especies de *Prototheca*, *Dunaliella salina*, y *Chlorella protothecoides* se muestran en las Tablas 4, 5, y 6, respectivamente.

Tabla 4. Uso de codones en especies de *Prototheca*

Ala	GCG	345 (0.36)	Asn	AAT	8 (0.04)
	GCA	66 (0.07)		AAC	201 (0.96)
	GCT	101 (0.11)	Pro	CCG	161 (0.29)
	GCC	442 (0.46)		CCA	49 (0.09)
Cys	TGT	12 (0.10)		CCT	71 (0.13)
	TGC	105 (0.90)		CCC	267 (0.49)
Asp	GAT	43 (0.12)	Gln	CAG	226 (0.82)
	GAC	316 (0.88)		CAA	48 (0.18)
Glu	GAG	377 (0.96)	Arg	AGG	33 (0.06)
	GAA	14 (0.04)		AGA	14 (0.02)
				CGG	102 (0.18)

5	Phe	TTT	89 (0.29)	Ser	CGA	49 (0.08)
		TTC	216 (0.71)		CGT	51 (0.09)
	Gly	GGG	92 (0.12)		CGC	331 (0.57)
		GGA	56 (0.07)		AGT	16 (0.03)
GGT		76 (0.10)	AGC		123 (0.22)	
GGC		559 (0.71)	TCG		152 (0.28)	
10	His	CAT	42 (0.21)		TCA	31 (0.06)
		CAC	154 (0.79)		TCT	55 (0.10)
	Ile	ATA	4 (0.01)		TCC	173 (0.31)
		ATT	30 (0.08)		Thr	ACG
ATC		338 (0.91)	ACA		24 (0.05)	
15	Lys	AAG	284 (0.98)		ACT	21 (0.05)
		AAA	7 (0.02)	ACC	249 (0.52)	
	Leu	TTG	26 (0.04)	Val	GTG	308 (0.50)
		TTA	3 (0.00)	GTA	9 (0.01)	
CTG		447 (0.61)	GTT	35 (0.06)		
CTA		20 (0.03)	GTC	262 (0.43)		
CTT		45 (0.06)	Trp	TGG	107 (1.00)	
CTC		190 (0.26)		Tyr	TAT	10 (0.05)
25	Met	ATG	191 (1.00)	TAC	180 (0.95)	
				Terminación	TGA/TAG/TAA	

Tabla 5. Uso de codones preferidos en *Dunaliella salina*.

TTC (Phe)	TAC (Tyr)	TGC (Cys)	TAA (de terminación)
TGG (Trp)	CCC (Pro)	CAC (His)	CGC (Arg)
CTG (Leu)	CAG (Gln)	ATC (Ile)	ACC (Thr)
AAC (Asn)	AGC (Ser)	ATG (Met)	AAG (Lys)
GCC (Ala)	GAC (Asp)	GGC (Gly)	GTG (Val)
GAG (Glu)			

30 Tabla 6. Uso de codones preferidos en *Chlorella protothecoides*.

TTC (Phe)	TAC (Tyr)	TGC (Cys)	TGA (de terminación)
TGG (Trp)	CCC (Pro)	CAC (His)	CGC (Arg)
CTG (Leu)	CAG (Gln)	ATC (Ile)	ACC (Thr)
GAC (Asp)	TCC (Ser)	ATG (Met)	AAG (Lys)

GCC (Ala)	AAC (Asn)	GGC (Gly)	GTG (Val)
GAG (Glu)			

B. Promotores

[0023] Muchos promotores están activos en microalgas, lo que incluye promotores que son endógenos a las algas que se transforman, así como promotores que no son endógenos a las algas que se transforman (es decir, promotores de otras algas, promotores de plantas superiores, y los promotores de virus de plantas o virus de algas). Los promotores exógenos y/o endógenos que se activan en las microalgas, y los genes de resistencia a los antibióticos funcionales en microalgas son descritos por ejemplo, *por Curr Microbiol.* 1997 dic;35(6):356-62 (*Chlorella vulgaris*); *Mar Biotechnol* (NY). 2002 ene;4(1):63-73 (*Chlorella ellipsoidea*); *Mol Gen Genet.* 1996 oct 16;252(5):572-9 (*Phaeodactylum tricornutum*); *Plant Mol Biol.* 1996 abr;31(1):1-12 (*Volvox carteri*); *Proc Natl Acad Sci Estados Unidos.* 1994 nov 22;91(24):11562-6 (*Volvox carteri*); Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C, PMID: 10383998, Mayo 1999;1(3):239-251 (Laboratory of Molecular Plant Biology, Stazione Zoologica, Villa Comunale, I-80121 Naples, Italia) (*Phaeodactylum tricornutum* y *Thalassiosira weissflogii*); *Plant Physiol.* Mayo 2002;129(1):7-12. (*Porphyridium sp.*); *Proc Natl Acad Sci Estados Unidos.* 21 de enero 2003;100(2):438-42. (*Chlamydomonas reinhardtii*); *Proc Natl Acad Sci Estados Unidos.* Febrero 1990;87(3):1228-32. (*Chlamydomonas reinhardtii*); *ácidos nucleicos Res.* 25 de Junio 1992 ;20(12):2959-65; *Mar Biotechnol* (NY). Enero 2002;4(1):63-73 (*Chlorella*); *Biochem Mol Biol Int.* Agosto 1995;36(5):1025-35 (*Chlamydomonas reinhardtii*); *J Microbiol.* Agosto 2005;43(4):361-5 (*Dunaliella*); Yi Chuan Xue Bao. Abril 2005; 32(4):424-33 (*Dunaliella*); *Mar Biotechnol* (NY). Mayo 1999;1(3):239-251. (*Thalassiosira* y *Phaedactylum*); Koksharova, *Appl Microbiol Biotechnol* 2002 feb;58(2):123-37 (varias especies); *Mol Genet Genomics* febrero 2004; 271(1):50-9 (*Thermosynechococcus elongates*); *J. Bacteriol.* (2000), 182, 211-215; *FEMS Microbiol Lett.* 25 de abril 2003;221(2):155-9; *Plant Physiol.* Junio 1994;105(2):635-41; *Plant Mol Biol.* Diciembre 1995;29(5):897-907 (*Synechococcus* PCC 7942); *Mar Pollut Bull.* 2002;45(1-12):163-7 (*Anabaena* PCC 7120); *Proc Natl Acad Sci Estados Unidos.* Marzo 1984 ;81(5):1561-5 (*Anabaena* (varias cepas)); *Proc Natl Acad Sci Estados Unidos.* 27 de marzo 2001;98(7):4243-8 (*Synechocystis*); Wirth, *Mol Gen Genet* Marzo 1989; 216(1):175-7 (varias especies); *Mol Microbiol*, junio 2002; 44(6):1517-31; *Plasmid*, septiembre 1993;30(2):90-105 (*Fremyella diplosiphon*); Hall y otros. (1993) *Gene*

124: 75-81 (*Chlamydomonas reinhardtii*); Gruber y otros. (1991). *Current Micro.* 22: 15-20; Jarvis y otros. (1991) *Current Genet.* 19: 317-322 (*Chlorella*); para promotores adicionales ver también la Tabla 1 de la patente de Estados Unidos 6,027,900).

[0024] El promotor que se usa para expresar un gen exógeno puede ser el promotor que se une naturalmente a ese gen o puede ser un gen heterólogo. Algunos promotores están activos en más de una especie de microalgas. Otros promotores son específicos de especie. Los promotores preferidos incluyen promotores tales como RBCS2 de *Chlamydomonas reinhardtii* y promotores virales, tales como el de virus del mosaico de la coliflor (CMV) y virus de *Chlorella*, que demostraron que se activan en múltiples especies de microalgas (ver, por ejemplo, *Plant Cell Rep.* 2005 mar;23(10-11):727-35; *J Microbiol.* 2005 ago;43(4):361-5; *Mar Biotechnol* (NY). 2002 ene;4(1):63-73). En otras modalidades, se puede usar el promotor malato deshidrogenasa de *Botryococcus*, un ácido nucleico de ese tipo que comprende cualquier parte de la sec. con núm. de ident.: 3, o el promotor RBCS2 de *Chlamydomonas reinhardtii* (sec. con núm. de ident.: 4). Opcionalmente, se usan al menos 10, 20, 30, 40, 50, ó 60 nucleótidos o más de estas secuencias que contienen un promotor. Los promotores endógenos preferidos para las especies del género *Chlorella* son las sec. con núms. de ident.: 1 y 2.

[0025] Los promotores preferidos útiles para la expresión de genes exógenos en *Chlorella* se enumeran en el listado de secuencias de esta aplicación, tal como el promotor del gen HUP1 de la *Chlorella* (sec. con núm. de ident.: 1) y el promotor de la nitrato reductasa *Chlorella ellipsoidea* (sec. con núm. de ident.: 2). Los promotores del virus de la *Chlorella* también se pueden usar para expresar genes en *Chlorella*, tales como las secs. con núms. de ident.: 1-7 de la patente de Estados Unidos 6,395,965. Los promotores adicionales activos en *Chlorella* se pueden encontrar, por ejemplo, en *Biochem Biophys Res Commun*, 14 de octubre 1994 ;204(1):187-94 *Plant Mol Biol*, octubre 1994;26(1):85-93; *Virology*, 15 de agosto 2004;326(1):150-9; y *Virology*, 5 de enero 2004;318(1):214-23.

C. Marcadores de selección

Cualquiera de una amplia variedad de marcadores de selección se puede emplear en el constructo de un transgen útil para la transformación de un microorganismo, por ejemplo, *Chlorella*. Los ejemplos de marcadores de selección adecuados incluyen el gen de nitrato reductasa, el gen de la higromicina fosfotransferasa (HPT), el gen de la neomicina fosfotransferasa, y el gen ble, que confiere resistencia a bleomicina. Los métodos para

determinar la sensibilidad de las microalgas a los antibióticos son bien conocidos. Por ejemplo, *Mol Gen Genet* 16 de octubre de 1996; 252(5):572-9.

Más específicamente, Dawson y otros. (1997), *Current Microbiology* 35:356-362 (incorporada como referencia en la presente descripción en su totalidad), describen el uso del gen de la nitrato reductasa (NR) de *Chlorella vulgaris* como un marcador seleccionable para mutantes de *Chlorella sorokiniana* deficientes en NR. Kim y otros. (2002), *Mar. Biotechnol.* 4:63-73 (incorporada como referencia en la presente descripción en su totalidad), describen el uso del gen HPT como un marcador seleccionable para transformar la *Chlorella ellipsoidea*. Huang y otros (2007), *Appl. Microbiol. Biotechnol.* 72:197-205 (incorporada como referencia en la presente descripción en su totalidad), reportó el uso de Sh ble como un marcador seleccionable para *Chlorella sp. DT*.

D. Expresión inducible

La presente invención se proporciona también para el uso de un promotor inducible para expresar un gen de interés. En particular, el uso de un promotor inducible para expresar un gen lipasa permite la producción de la lipasa después del cultivo del microorganismo cuando las condiciones se han ajustado, si es necesario, para mejorar la transesterificación, por ejemplo, después de la ruptura de las células, reducción del contenido de agua de la mezcla de reacción, y/o adición de alcohol suficiente para conducir la conversión de los TAG a ésteres de ácidos grasos.

[0026] Los promotores inducibles útiles en la invención incluyen aquellos que median la transcripción de un gen unido operativamente en respuesta a un estímulo, tal como una molécula pequeña suministrada exógenamente (por ejemplo, la glucosa, como en la sec. con núm. de ident.: 1), temperatura (calor o frío), luz, etc. Los promotores adecuados pueden activar la transcripción de un gen esencialmente silente o regular ascendentemente, preferiblemente de forma sustancial, la transcripción de un gen unido operativamente que se transcribe a un nivel bajo. En este último caso, el nivel de transcripción del gen de interés, por ejemplo, el gen lipasa, de preferencia no interfiere significativamente con el crecimiento del microorganismo en el que se expresa.

La expresión de un transgen en *Chlorella* se puede realizar bajo condiciones inducibles, por ejemplo, usando promotores, tales como el promotor para el gen portador de hexosa *Chlorella* (sec. con núm. de ident.: 1). Este promotor se activa fuertemente por la presencia de glucosa en los medios de cultivo.

E. Expresión de dos o más genes exógenos

Además, un microorganismo genéticamente modificado, tales como una microalga, podría comprender y expresar dos o más genes exógenos, tales como, por ejemplo, una lipasa y un gen lítico, por ejemplo, uno que codifica una enzima de degradación de polisacáridos.

5 Uno o ambos genes se pueden expresar por medio del uso de un promotor inducible, el cual permite controlar el tiempo de expresión relativo de estos genes para mejorar el rendimiento de los lípidos y la conversión a ésteres de ácidos grasos. La expresión de los dos o más genes exógenos puede estar bajo el control del mismo promotor inducible o bajo el control de diferentes promotores inducibles. En la última situación, la expresión de un primer gen
10 exógeno se puede inducir por un primer período de tiempo (durante el cual la expresión de un segundo gen exógeno puede o no inducirse) y la expresión de un segundo gen exógeno se puede inducir por un segundo período de tiempo (durante el cual la expresión del primer gen exógeno puede o no inducirse). En la presente se proporcionan vectores y métodos para modificar los microbios que producen lípidos para metabolizar la sacarosa, el cual es un
15 rasgo ventajoso porque permite a las células modificadas convertir las materias primas de la caña de azúcar en lípidos apropiados para la producción de biodiesel.

Se proporcionan, además, cepas de microbios genéticamente modificados (por ejemplo, microalgas, levadura oleaginosa, bacteria u hongos) que expresan dos o más genes exógenos, tales como, por ejemplo, una acil graso-ACP tioesterasa y una acil graso-CoA/
20 aldehído reductasa, la acción combinada de los cuales produce un producto alcohólico. Además se proporcionan otras combinaciones de genes exógenos, que incluyen, sin limitación, una acil graso ACP tioesterasa y una acil graso-CoA reductasa para generar aldehídos. Además, esta solicitud proporciona la combinación de una acil graso ACP tioesterasa, una acil graso-CoA reductasa y una aldehído graso decarbonilasa para generar
25 alcanos. Uno o más de los genes exógenos se pueden expresar por medio del uso de un promotor inducible.

Los ejemplos de modificaciones adicionales más adecuadas para el uso en la presente invención se describen en la co-pendiente, propiedad común Solicitud núm. 60/837, 839, que se incorpora en la presente como referencia. Esta solicitud describe las cepas de
30 microalgas genéticamente modificadas para expresar dos o más genes exógenos, una que codifica un portador de una fuente de carbono fija (tal como sacarosa) y una segunda que codifica una enzima sacarosa invertasa. Los organismos fermentables resultantes producen

hidrocarburos a menor costo de fabricación que los que se han obtenidos por métodos previamente conocidos de producción biológica de hidrocarburo. Esta solicitud co-pendiente también enseña que la inserción de los dos genes exógenos descrita anteriormente se puede combinar con la interrupción de la biosíntesis de polisacáridos a través de mutagénesis dirigida y/o aleatoria, que dirige siempre el mayor flujo de carbono a la producción de hidrocarburos. Individualmente y en combinación, la conversión trófica, las modificaciones para alterar la producción de hidrocarburos y el tratamiento con enzimas exógenos alteran la composición de los hidrocarburos producida por un microorganismo. La alteración puede ser un cambio en la cantidad de hidrocarburos producidos, la cantidad de una o más especies de hidrocarburos producidos en relación con otros hidrocarburos, y/o los tipos de especies de hidrocarburos que se producen en el microorganismo. Por ejemplo, las microalgas se pueden modificar para producir una mayor cantidad y/o porcentaje de TAG.

F. Expresión compartimentada

La presente invención se proporciona también para la expresión compartimentada de un gen de interés. En particular, en modalidades particulares, puede ser ventajoso dirigir la expresión de un gen lipasa a uno o varios compartimentos celulares, donde se aísla a partir de la mayoría de los lípidos celulares hasta el inicio de la reacción de transesterificación. Los organelos preferidos para la orientación son los cloroplastos, mitocondrias, y el retículo endoplásmico.

1. Expresión en cloroplastos

En una modalidad preferida de la presente invención, la expresión de un polipéptido en un microorganismo se dirige a los cloroplastos. Los métodos para dirigir la expresión de un gen heterólogo al cloroplasto se conocen y se pueden emplear en la presente invención. Los métodos para dirigir los productos de genes extraños en los cloroplastos se describen en Shrier y otros, *EMBO J.* (1985) 4:25 32. Ver también Tomai y otros, *Gen. Biol. Chem.* (1988) 263:15104 15109 y Patente de Estados Unidos núm. 4,940,835 para el uso de péptidos de tránsito para la translocación nuclear de los productos de genes en el cloroplasto. Los métodos para dirigir el transporte de proteínas al cloroplasto se analizan también en Kenauf *TIBTECH* (1987) 5:40 47. Las secuencias dirigidas en cloroplasto endógenas a *Chlorella* se conocen, tales como genes en el genoma nuclear de *Chlorella* que codifica proteínas que se dirigen al cloroplasto; ver, por ejemplo, los números de acceso de GenBank AY646197 y AF499684.

Wageningen UR-Plant Research International vende un vector IMPACTVECTOR1.4, que usa la señal de secreción de la subunidad pequeña de la proteína del *Chrysanthemum morifolium* para entregar una proteína heteróloga en el ambiente del estroma del cloroplasto (citoplasmático), trasladándose a través de un sistema de doble
 5 membrana. La proteína se fusiona con los 11 primeros aminoácidos de la proteína madura rubisco para permitir el procesamiento adecuado del péptido señal (Wong y otros, *Plant Molecular Biology* 20: 81-93 (1992)). El péptido señal contiene un intrón natural del gen RbcS.

En otro enfoque, el genoma del cloroplasto es genéticamente modificado para
 10 expresar la proteína heteróloga. Se ha descrito la transformación estable de los cloroplastos de *Chlamydomonas reinhardtii* (un alga verde), usando el bombardeo de las células receptoras con microproyectiles de tungsteno de alta velocidad recubiertos con ADN extraño. Ver, por ejemplo, Boynton y otros, *Science* (1988) 240: 1534 1538; Blowers y otros, *Plant Cell* (1989) 1:123 132 y Debuchy y otros, *EMBO J.* (1989) 8: 2803 2809. La técnica de
 15 transformación, usando microproyectiles de tungsteno, se describe en Klein y otros, *Nature* (London) (1987) 7:70 73. Se conocen otros métodos de transformación del cloroplasto para plantas y microalgas. Ver, por ejemplo, la patente de Estados Unidos 5,693,507; 6,680,426; *Plant Physiol.* 2002 may;129(1):7-12; y *Plant Biotechnol J.* 2007 may;5(3):402-12.

Como se describe en la patente de Estados Unidos núm. 6,320,101 (publicada el 20
 20 de noviembre 2001 de Kaplan y otros; que se incorpora en la presente como referencia), las células se pueden tratar químicamente para reducir el número de cloroplastos por célula hasta aproximadamente uno. Después, el ácido nucleico heterólogo se puede introducir en las células a través del bombardeo de partículas con el objetivo de introducir al menos una molécula de ácido nucleico heteróloga en los cloroplastos. El ácido nucleico heterólogo se
 25 selecciona de manera que se integre en el genoma del cloroplasto a través de la recombinación homóloga que se efectúa fácilmente por las enzimas propias de los cloroplastos. Con este fin, el ácido nucleico heterólogo incluye, además de un gen de interés, al menos una secuencia de ácido nucleico que se deriva a partir del genoma del cloroplasto. Además, el ácido nucleico heterólogo típicamente incluye un marcador de selección. Detalles
 30 adicionales con relación a esta técnica se encuentran en la patente de Estados Unidos núms. 4,945,050 y 5,693,507, que se incorporan en la presente como referencia. Un polipéptido se puede producir de este modo por el sistema de expresión de la proteína del cloroplasto.

La patente de Estados Unidos núm. 7,135,620 (publicada el 14 de noviembre de 2006 de Daniell *y otros*, se incorpora en la presente como referencia) describe los vectores de expresión de cloroplastos y métodos relacionados. Los casetes de expresión son constructos de ADN que incluyen una secuencia de codificación y secuencias de control adecuadas para proporcionar la expresión correcta de la secuencia de codificación en el cloroplasto. Los casetes de expresión típicos incluyen los siguientes componentes: la región 5' sin traducir de un gen de microorganismo o gen de cloroplasto tal como psbA que se proveerá para la transcripción y traducción de una secuencia de ADN que codifica un polipéptido de interés en el cloroplasto; una secuencia de ADN que codifica un polipéptido de interés; y una región de terminación de la traducción y transcripción, tal como una región 3' de repetición invertida de un gen del cloroplasto que puede estabilizar el ARN de los genes introducidos, mejorando de ese modo la expresión de genes extraños. El casete puede incluir opcionalmente un gen de resistencia a antibiótico.

Típicamente, el casete de expresión está flanqueado por sitios de restricción convenientes para la inserción en un genoma adecuado. El casete de expresión se puede flanquear por secuencias de ADN del ADN del cloroplasto para facilitar la integración estable del casete de expresión en el genoma del cloroplasto, particularmente por recombinación homóloga. Alternativamente, el casete de expresión puede permanecer sin integrarse, en cuyo caso, el casete de expresión incluye, típicamente, un un origen de replicación del cloroplasto, el cual es capaz de proporcionar la replicación del ADN heterólogo en el cloroplasto.

El casete de expresión generalmente, incluye una región del promotor de un gen capaz de la expresión en el cloroplasto. La región del promotor podría incluir los promotores obtenibles de genes de cloroplastos, tales como el gen psbA de la espinaca o guisante, o la región del promotor rbcL y atpB del maíz y promotores ARNr. Los ejemplos de promotores se describen en Hanley-Bowdoin y Chua, *TIBS* (1987) 12:67 70; Mullet *y otros*, *Plant Molec Biol.* (1985) 4: 39 54; Hanley-Bowdoin (1986) PhD. Dissertation, the Rockefeller University; Krebbers *y otros*, *Nucleic Acids Res.* (1982) 10: 4985 5002; Zurawaki *y otros*, *Nucleic Acids Res.* (1981) 9:3251 3270; y Zurawski *y otros*, *Proc. Natl. Acad. Sci. Estados Unidos.* (1982) 79: 7699 7703. Otros promotores pueden identificarse y evaluar la fuerza relativa de los promotores de ese modo identificados, mediante la colocación de un promotor de interés 5' a un gen marcador menos el promotor y observar su eficacia en relación con la transcripción obtenida a partir de, por ejemplo, el promotor del gen psbA, un promotor del cloroplasto

relativamente fuerte. La eficacia de la expresión de genes heterólogos se puede mejorar, además, por cualquiera de una variedad de técnicas. Estas incluyen el uso de múltiples promotores insertados en tándem 5' al gen heterólogo, por ejemplo, un promotor doble psbA, la adición de secuencias potenciadoras y similares.

5 Numerosos promotores activos en el cloroplasto de *Chlorella* se pueden usar para la expresión de genes exógenos en el cloroplasto de *Chlorella*, tal como aquellos encontrados en el número de acceso de GenBank NC_001865 (genoma completo del cloroplasto de *Chlorella vulgaris*).

10 Donde se desea proporcionar la expresión inducible del gen heterólogo, se pueden incluir en el casete de expresión, un promotor inducible y/o una región 5' sin traducir que contiene secuencias que se proporcionan para la regulación a nivel de la transcripción y/o traducción (en el extremo 3'). Por ejemplo, la región 5' sin traducir puede ser de un gen donde la expresión es regulable por la luz. Del mismo modo, las regiones repetidas invertidas 3' se pueden usar para estabilizar el ARN de los genes heterólogos. Los genes inducibles se
15 pueden identificar por una expresión mejorada en la respuesta a un estímulo particular de interés y una expresión baja o ausente en ausencia del estímulo. Por ejemplo, un gen inducible por la luz se puede identificar donde ocurre la expresión mejorada durante la irradiación con luz, mientras que la expresión sustancialmente reducida o la no expresión ocurre en poca o sin luz. Los promotores regulados por la luz a partir de microalgas verdes se
20 conocen (ver, por ejemplo, *Mol Genet Genomics*, diciembre de 2005; 274 (6) :625-36).

La región de terminación que se emplea será principalmente una de conveniencia, ya que la región de terminación parece ser relativamente intercambiable entre los cloroplastos y la bacteria. La región de terminación puede ser nativa para la región de iniciación de la transcripción, puede ser nativa para la secuencia de ADN de interés, o se puede obtener de
25 otra fuente. Ver, por ejemplo, Chen y Orozco, *Nucleic Acids Res.* (1988) 16:8411.

Los casetes de expresión pueden transformarse en una célula de la planta de interés por cualquiera de un número de métodos. Estos métodos incluyen, por ejemplo, métodos biolísticos (Ver, por ejemplo, Sanford, *Trends In Biotech.* (1988) 6:299 302, patente de Estados Unidos núm. 4,945,050; electroporación (Fromm y otros, *Proc. Natl. Acad. Sci.* (Estados Unidos) (1985) 82:5824 5828); uso de un rayo láser, microinyección o cualquier
30 otro método capaz de introducir ADN en un cloroplasto.

Las descripciones adicionales de vectores de expresión de cloroplastos adecuados para uso en los microorganismos tales como las microalgas se encuentran en la patente de Estados Unidos núm. 7,081,567 (publicada el 25 de julio de 2006, de Xue *y otros.*); 6,680,426 (publicada el 20 de enero de 2004, de Daniell *y otros*) y 5,693,507 (publicado el 2 de diciembre de 1997, de Daniell *y otros*).

Las proteínas expresadas en el genoma nuclear de *Chlorella* se pueden dirigir al cloroplasto usando señales dirigidas al cloroplasto. Las secuencias dirigidas al cloroplasto endógenas a *Chlorella* se conocen, tal como los genes en el genoma nuclear de *Chlorella* que codifican proteínas que se dirigen al cloroplasto; ver, por ejemplo, los números de acceso de GenBank AY646197 y AF499684. Las proteínas también se pueden expresar en el cloroplasto de *Chlorella* por la inserción de genes directamente en el genoma del cloroplasto. La transformación del cloroplasto ocurre típicamente a través de la recombinación homóloga, y se puede realizar si las secuencias del genoma del cloroplasto se conocen por la creación de vectores dirigidos (ver, por ejemplo, la secuencia del genoma completo de un cloroplasto de *Chlorella*, número de acceso del GenBank NC_001865). Ver las secciones anteriores en la presente para los detalles de la transformación del cloroplasto.

2. Expresión en la mitocondria

En otra modalidad preferida de la presente invención, la expresión de un polipéptido en un microorganismo se dirige a la mitocondria. Se han descrito métodos para la selección de productos de genes extraños en las mitocondrias (Boutry *y otros*, *Nature* (Londres) (1987) 328:340-342), incluyendo en microalgas verdes (ver, por ejemplo, *Mol Gen Genet.*, 1993 ene;236(2-3):235-44).

Por ejemplo un vector de expresión que codifica una señal de secreción adecuada puede dirigir una proteína heteróloga a la mitocondria. Un vector de expresión ejemplar para dirigir la mitocondria es el vector IMPACTVECTOR1.5, de Wageningen UR-Plant Research International, que usa la señal de secreción CoxIV de levadura. Este vector de expresión demostró entregar proteínas en la matriz mitocondrial. La proteína se fusiona con los 4 primeros aminoácidos de la proteína CoxIV de levadura para permitir el procesamiento adecuado del péptido señal (Kohler *y otros*, *Plant J.* 11: 613-621 (1997)). Otras secuencias dirigidas mitocondriales se conocen, incluyendo aquellas funcionales en las microalgas verdes. Por ejemplo, ver *FEBS Lett.* 29 de enero de 1990 ;260(2):165-8; y *J. Biol. Chem.* 22 de febrero de 2002;277(8):6051-8.

Las proteínas expresadas en el genoma nuclear de *Chlorella* se pueden dirigir a la mitocondria usando señales mitocondriales dirigidas. Los detalles de la dirección y transformación de la proteína mitocondrial se proporcionan en la presente.

3. Expresión en el retículo endoplasmático

5 En otra modalidad preferida de la presente invención, la expresión de un polipéptido en un microorganismo se dirige al retículo endoplasmático (ER). La inclusión de una señal de retención o clasificación apropiada en un vector de expresión garantiza que las proteínas se retengan en el retículo endoplasmático (RE) y no vayan corriente abajo adentro de Golgi. Por ejemplo, el vector IMPACTVECTOR1.3, de Wageningen UR-Plant Research
10 International, incluye la señal KDEL de retención o clasificación bien conocida. Con este vector, la retención en el RE tiene una ventaja práctica ya que se reportó que mejora los niveles de expresión 5 veces o más. La razón principal de esto parece ser que el RE contiene concentraciones más bajas y/o proteasas diferentes responsables de la degradación post-traducciona

15 l de las proteínas expresadas que están presentes en el citoplasma. Se conocen las señales de retención del RE funcionales en microalgas verdes. Por ejemplo, ver *Proc. Natl. Acad. Sci. Estados Unidos*. 26 de abril de 2005;102(17):6225-30.

G. Transformación

[0027] Las células se pueden transformar por cualquier técnica adecuada que incluye, *por ejemplo*, electroporación biolística, (ver Maruyama y otros, (2004)
20 *Biotechnology Techniques* 8:821-826), transformación por perlas de vidrio y transformación por bigotes de carburo de silicio. Otro método que puede usarse involucra formar protoplastos y usar CaCl_2 y polietilenglicol (PEG) para introducir ADN recombinante en células de microalgas (ver Kim y otros, (2002), *Mar. Biotechnol.* 4:63-73, que reporta el uso de este método para la transformación de *Chlorella ellipsoidea*). La co-transformación de
25 microalgas se puede usar para introducir dos moléculas diferentes de vectores en una célula simultáneamente (ver, por ejemplo, *Protist* 2004 dic;155(4):381-93).

Los métodos de biolística (ver, por ejemplo, Sanford, *Trends in Biotech.* (1988) 6:299-302, patente de Estados Unidos núm. 4,945,050; electroporación (Fromm y otros, *PNAS*(1985) 82:5824-5828), uso de un rayo láser, microinyección, o cualquier otro método
30 capaz de introducir ADN en una microalga se puede usar también para la transformación.

H. Modificación de la vía lipídica

En ciertas modalidades de la presente invención, se prefiere modificar adicionalmente un microorganismo, tal como una microalga, por ejemplo, para proporcionar las características de crecimiento deseadas y/o para mejorar la cantidad y/o calidad de los lípidos producidos. Por ejemplo, las microalgas se pueden modificar para incrementar el flujo de carbono en la vía lipídica y/o modificar la vía lipídica para alterar beneficiosamente las proporciones o propiedades del lípido producido por las células. La vía es además, o alternativamente, modificada para alterar las propiedades y/o proporciones de varias moléculas de hidrocarburo que se producen a través del procesamiento enzimático de los lípidos.

1. **Alteración de las propiedades o porciones de lípidos o hidrocarburos producidos**

En algunas modalidades de la presente invención, se puede desear alterar las características, tales como el rendimiento de lípido por unidad de volumen y/o por unidad de tiempo, la longitud de cadena de carbono (por ejemplo, para la producción de biodiesel o para aplicaciones industriales que requieren materias primas de hidrocarburo), reducir el número de dobles o triples enlaces, opcionalmente a cero, quitar o eliminar anillos y estructuras cíclicas, e incrementar la proporción hidrógeno: carbono de una especie particular de lípido o de una población de lípidos distintos. Además, las microalgas que producen hidrocarburos adecuados pueden modificarse, además, para tener rendimiento de hidrocarburos aún más deseables. Los ejemplos de tales microalgas incluyen especies del género *Chlorella*.

a) **Regulación de enzimas que controlan los puntos de ramificación en la síntesis de ácidos grasos**

En modalidades particulares de la presente invención, una o más enzimas claves que controlan puntos de ramificación en el metabolismo de la síntesis de ácidos grasos se regularon ascendente o descendentemente para mejorar la producción de lípidos. La regulación ascendente se puede lograr, por ejemplo, mediante la transformación de células con constructos de expresión en las que un gen que codifica la enzima de interés se expresa, por ejemplo, por medio del uso de un promotor fuerte y/o elementos potenciadores que aumentan la transcripción. Tales constructos de expresión pueden incluir un marcador seleccionable de manera que los transformantes pueden ser objeto de selección, lo que puede resultar en una amplificación del constructo y un aumento en el nivel de expresión de la enzima codificada. Los ejemplos de enzimas adecuadas para la regulación ascendente de

acuerdo con los métodos de la invención incluyen la piruvato deshidrogenasa, que desempeña un papel en la conversión del piruvato a acetil-CoA (ejemplos, algunas de microalgas, incluyen números de acceso en Genbank NP_415392; AAA53047; Q1XDM1; y CAF05587). La regulación ascendente de la piruvato deshidrogenasa puede aumentar la producción de acetil-CoA, y de ese modo aumentar la síntesis de ácidos grasos. La acetil-CoA carboxilasa cataliza la etapa inicial en la síntesis de ácidos grasos. En correspondencia, en ciertas modalidades de la presente invención, esta enzima se regula ascendentemente para aumentar la producción de ácidos grasos (ejemplos, algunos de microalgas, incluyen los números de acceso al GenBank BAA94752; AAA75528; AAA81471; YP_537052; YP_536879; NP_045833; y BAA57908). En otra modalidad la producción de ácidos grasos también se incrementa por la regulación ascendente de la proteína portadora de acilo (ACP), la cual transporta las cadenas de acilo en crecimiento durante la síntesis de ácidos grasos (ejemplo, algunas de microalgas, incluyen los números de acceso en GenBank A0T0F8; P51280; NP_849041; YP_874433). La glicerol-3-fosfato acil transferasa cataliza el paso limitante de la velocidad de la síntesis de ácidos grasos. La regulación ascendente de esta enzima se desea para aumentar la producción de ácidos grasos (por ejemplo, algunas de microalgas, incluyen los números de acceso al GenBank AAA74319; AAA33122; AAA37647; P44857; y ABO94442). En algunas modalidades, dos o más de estos polipéptidos (piruvato deshidrogenasa, acetil-CoA carboxilasa, proteína portadora de acilos (ACP), glicerol-3-fosfato aciltransferasa) están regulados de forma ascendente. En ese caso, los dos o más genes que codifican los polipéptidos correspondientes podrían residir en un constructo de expresión individual o, alternativamente, en dos o más constructos de expresión. Las proteínas anteriores son candidatos para la expresión en microalgas, incluyendo las especies del género *Chlorella* y/o *Prototheca*.

La regulación descendente de una enzima de interés se puede lograr usando, por ejemplo, técnicas de mutagénesis antisentido, ARN/ADN catalítico, ARN de interferencia (ARNi), “knock-out,” “knock-down,” u otras. La expresión/función de la enzima se puede inhibir también usando intracuerpos. Los ejemplos de enzimas adecuadas para la regulación descendente de acuerdo con los métodos de la invención incluyen citrato sintasa, que consume la acetil-CoA como parte del ciclo de ácido tricarboxílico (TCA). La regulación descendente de la citrato sintasa puede forzar más acetil-CoA en la vía sintética de ácidos grasos.

b) Modulación de reguladores globales de los genes sintéticos de ácidos grasos

Los reguladores globales modulan la expresión de los genes de las vías biosintéticas de ácidos grasos. En consecuencia, uno o más reguladores globales de la síntesis de ácidos grasos se regula positiva o negativamente, según proceda, para inhibir o aumentar, respectivamente, la expresión de una pluralidad de genes sintéticos de ácidos grasos y, en última instancia, aumentar la producción de lípidos. Los ejemplos incluyen las proteínas de unión a elementos reguladores del estero, (SREBP), tales como SREBP-1a y SREBP-1c (para ejemplos ver los números de acceso en Genbank NP_035610 y Q9WTN3). Los reguladores globales pueden ser regulados de forma ascendente o descendente, por ejemplo, como se describió anteriormente con respecto a la regulación de las enzimas del punto de control.

c) Regulación de las enzimas de modificación de hidrocarburos

La presente solicitud describe las cepas genéticamente modificadas de microalgas, levadura oleaginosa, bacterias, u hongos con uno o más genes exógenos para producir diversos compuestos de hidrocarburos. De este modo, en ciertas modalidades de la presente invención, los métodos de la invención comprenden también la transformación de las células con uno o más genes que codifican enzimas para la modificación de hidrocarburos, tales como, por ejemplo, una acil graso-ACP tioesterasa (ver ejemplos en la Tabla 7 con números de acceso), una acil graso-CoA/aldehído reductasa (ver ejemplos en la Tabla 8, con números de acceso), una acil graso-CoA reductasa (ver ejemplos en la Tabla 9 con números de acceso), una aldehído graso decarbonilasa (ver ejemplos en la Tabla 10 con números de acceso), una aldehído graso reductasa, o un gen de escualeno sintasa (por ejemplo, ver el número de acceso AF205791 de GenBank). La estearoil-ACP desaturasa, por ejemplo, cataliza la conversión de estearoil-ACP a oleoil-ACP. La regulación ascendente de este gen puede aumentar la proporción de ácidos grasos monoinsaturados producida por una célula, mientras que la regulación descendente puede reducir la proporción de monoinsaturados. Del mismo modo, la expresión de una o más desaturasas de glicerolípidos se puede controlar para alterar la relación de ácidos grasos insaturados a saturados tales como ácido graso ω -6 desaturasa, ácidos grasos ω -3 desaturasa, o ω -6-oleato desaturasa.

Por ejemplo, se pueden modificar (o modificar adicionalmente) microalgas que podrían producir altos niveles de lípidos naturalmente, o mediante modificación genética

para expresar una acil-ACP tioesterasa exógena, que puede facilitar la escisión de los ácidos grasos de la proteína portadora de acilo (ACP) durante la síntesis lipídica. Estos ácidos grasos se pueden recuperar o, a través del procesamiento enzimático adicional dentro de la célula, producir otros compuestos de hidrocarburos. Opcionalmente, la acil graso-ACP tioesterasa se puede expresar a partir de un gen unido de forma operativa a un promotor inducible, y/o se puede expresar en un compartimiento intracelular.

Así, en una modalidad preferida de la presente invención, la enzimas modificadora de hidrocarburo adecuada para usar con los microorganismos y métodos de la invención es una acil graso-ACP tioesterasa. Las acil graso-ACP tioesteras incluyen, sin limitación, las enumeradas en la Tabla 7, cada una de las cuales se incorpora por este medio como referencia.

[0028] Tabla 7. Acil graso-ACP tioesterasas y números de acceso del GenBank.

<i>Umbellularia californica</i> acil graso-ACP tioesterasa (GenBank #AAC49001)
<i>Cinnamomum camphora</i> acil graso-ACP tioesterasa (GenBank #Q39473)
<i>Umbellularia californica</i> acil graso-ACP tioesterasa (GenBank #Q41635)
<i>Myristica fragrans</i> acil graso-ACP tioesterasa (GenBank #AAB71729)
<i>Myristica fragrans</i> acil graso-ACP tioesterasa (GenBank #AAB71730)
<i>Elaeis guineensis</i> acil graso-ACP tioesterasa (GenBank #ABD83939)
<i>Elaeis guineensis</i> acil graso-ACP tioesterasa (GenBank #AAD42220)
<i>Populus tomentosa</i> acil graso-ACP tioesterasa (GenBank #ABC47311)
<i>Arabidopsis thaliana</i> acil graso-ACP tioesterasa (GenBank #NP_172327)
<i>Arabidopsis thaliana</i> acil graso-ACP tioesterasa (GenBank #CAA85387)
<i>Arabidopsis thaliana</i> acil graso-ACP tioesterasa (GenBank #CAA85388)
<i>Gossypium hirsutum</i> acil graso-ACP tioesterasa (GenBank #Q9SQI3)
<i>Cuphea lanceolata</i> acil graso-ACP tioesterasa (GenBank #CAA54060)
<i>Cuphea hookeriana</i> acil graso-ACP tioesterasa (GenBank #AAC72882)
<i>Cuphea calophylla</i> subsp. <i>mesostemon</i> acil graso-ACP tioesterasa (GenBank #ABB71581)
<i>Cuphea lanceolata</i> acil graso-ACP tioesterasa (GenBank #CAC19933)
<i>Elaeis guineensis</i> acil graso-ACP tioesterasa (GenBank #AAL15645)
<i>Cuphea hookeriana</i> acil graso-ACP tioesterasa (GenBank #Q39513)
<i>Gossypium hirsutum</i> acil graso-ACP tioesterasa (GenBank #AAD01982)

Vitis vinifera acil graso-ACP tioesterasa (GenBank #CAN81819)
Garcinia mangostana acil graso-ACP tioesterasa (GenBank #AAB51525)
Brassica juncea acil graso-ACP tioesterasa (GenBank #ABI18986)
Madhuca longifolia acil graso-ACP tioesterasa (GenBank #AAX51637)
Brassica napus acil graso-ACP tioesterasa (GenBank #ABH11710)
Oryza sativa (cultivares grupo indica) acil graso-ACP tioesterasa (GenBank #EAY86877)
Oryza sativa (cultivares grupo japónica) acil graso-ACP tioesterasa (GenBank #NP_001068400)
Oryza sativa (cultivares grupo indica) acil graso-ACP tioesterasa (GenBank #EAY99617)
Cuphea hookeriana acil graso-ACP tioesterasa (GenBank #AAC49269)

Una acil graso-ACP tioesterasa puede seleccionarse en base a su especificidad por un ácido graso de crecimiento (durante la síntesis de ácido graso) que tiene una longitud particular de la cadena de carbono. Por ejemplo, la acil graso-ACP tioesterasa puede tener una especificidad por una longitud de la cadena de carbono en el intervalo de 8 a 34 átomos de carbono, preferentemente de 8 a 18 átomos de carbono, y con mayor preferencia de 10 a 14 átomos de carbono. Una especificidad por un ácido graso con 12 átomos de carbono es lo más preferido. Una especificidad por un ácido graso con 14 átomos de carbono es también preferido.

En otra modalidad preferida de la presente invención, las enzimas modificadoras de hidrocarburos adecuadas para usar con los microorganismos y métodos de la invención es una acil graso-CoA/aldehído reductasa. Las acil graso-CoA/aldehído reductasas incluyen, sin limitación, las enumeradas en la Tabla 8.

Tabla 8. Acil graso-CoA/aldehído reductasas enumeradas con los números de acceso al GenBank.

AAC45217, YP_047869, BAB85476, YP_001086217, YP_580344, YP_001280274, YP_264583, YP_436109, YP_959769, ZP_01736962, ZP_01900335, ZP_01892096, ZP_01103974, ZP_01915077, YP_924106, YP_130411, ZP_01222731, YP_550815, YP_983712, YP_001019688, YP_524762, YP_856798, ZP_01115500, YP_001141848, NP_336047, NP_216059, YP_882409, YP_706156, YP_001136150, YP_952365, ZP_01221833, YP_130076, NP_567936, AAR88762, ABK28586, NP_197634, CAD30694, NP_001063962, BAD46254, NP_001030809, EAZ10132, EAZ43639, EAZ07989,

NP_001062488, CAB88537, NP_001052541, CAH66597, CAE02214, CAH66590, CAB88538, EAZ39844, AAZ06658, CAA68190, CAA52019, y BAC84377

En otra modalidad preferida de la presente invención, las enzimas modificadoras de hidrocarburos adecuadas para usar con los microorganismos y métodos de la invención es una acil graso-CoA reductasa. Las acil graso-CoA reductasas incluyen, sin limitación, las enumeradas en la Tabla 9.

[0029] Tabla 9. Acil graso-CoA reductasas enumeradas por los números de acceso al GenBank.

NP_187805, ABO14927, NP_001049083, CAN83375, NP_191229, EAZ42242, EAZ06453, CAD30696, BAD31814, NP_190040, AAD38039, CAD30692, CAN81280, NP_197642, NP_190041, AAL15288, y NP_190042

En otra modalidad preferida de la presente invención, las enzimas modificadoras de hidrocarburos adecuadas para usar con los microorganismos y métodos de la invención es una aldehído graso decarboxilasa. Las aldehído graso decarboxilasas incluyen, sin limitación, las enumeradas en la Tabla 10.

[0030] Tabla 10. Aldehído graso decarboxilasas enumeradas por los números de acceso al GenBank.

NP_850932, ABN07985, CAN60676, AAC23640, CAA65199, AAC24373, CAE03390, ABD28319, NP_181306, EAZ31322, CAN63491, EAY94825, EAY86731, CAL55686, XP_001420263, EAZ23849, NP_200588, NP_001063227, CAN83072, AAR90847, y AAR97643

Los ejemplos adicionales de secuencias de aminoácidos para las enzimas modificadoras de hidrocarburos o ácidos nucleicos que las codifican se describen en las patentes de Estados Unidos núms. 6,610,527, 6,451,576, 6,429,014, 6,342,380, 6,265,639, 6,194,185, 6,114,160, 6,083,731, 6,043,072, 5,994,114, 5,891,697, 5,871,988, y 6,265,639, y se describen además con los números de acceso al GenBank AAO18435, ZP_00513891, Q38710, AAK60613, AAK60610, AAK60611, NP_113747, CAB75874, AAK60612, AAF20201, BAA11024, AF205791, y CAA03710.

En modalidades particulares, los microorganismos de la presente invención son genéticamente modificados para expresar uno o más genes exógenos seleccionados de una acil graso-ACP tioesterasa, una acil graso-CoA/aldehído reductasa, una acil graso-CoA reductasa, una aldehído graso reductasa, o una aldehído graso decarbonilasa. Los métodos de expresión adecuados se describen anteriormente en relación con la expresión de un gen de la lipasa, incluyendo, entre otros métodos, la expresión inducible y la expresión compartimentada.

[0031] Otras enzimas adecuadas para su uso con los microorganismos y los métodos de la invención incluyen aquellas que tienen al menos 70% de identidad de aminoácidos con una de las proteínas que se enumeran en las Tablas 7-10, y que exhiben la correspondiente actividad enzimática deseada (por ejemplo, la escisión de un ácido graso de una proteína portadora de acilo, la reducción de un acil-CoA a un aldehído o un alcohol, o la conversión de un aldehído a un alcano). En modalidades adicionales, la actividad enzimática está presente en una secuencia que tiene al menos aproximadamente 75%, al menos aproximadamente 80%, al menos aproximadamente 85%, al menos aproximadamente 90%, al menos aproximadamente 95%, o al menos aproximadamente 99% de identidad con una de las secuencias descritas anteriormente, las que se incorporan como referencia en la presente como se exponen completamente.

Las enzimas para la modificación de hidrocarburo descritas en la presente descripción son útiles en la producción de diferentes tipos de hidrocarburos a partir de un microorganismo (por ejemplo, una microalga, una levadura oleaginosa o un hongo) o la población de microorganismos, por el cual una acil graso-ACP tioesterasa escinde un ácido graso a partir de una proteína portadora de acilo (ACP) durante la síntesis de lípido. Mediante procesamiento enzimático adicional, el ácido graso escindido se combina con una coenzima para producir una molécula de acil-CoA. Este acil-CoA es el sustrato para la actividad enzimática de la acil graso-CoA reductasa para rendir un aldehído, así como para una acil graso-CoA/aldehído reductasa para producir un alcohol. El aldehído que se produce por la acción de la acil graso-CoA reductasa identificada arriba es el sustrato para la actividad enzimática adicional ya sea por una aldehído graso reductasa para producir un alcohol, o una aldehído graso decarbonilasa para producir un alcano.

Las enzimas modificadoras de hidrocarburos tienen una especificidad para actuar en un sustrato que incluye un número específico de átomos de carbono. Por ejemplo, una acil

graso-ACP tioesterasa puede tener una especificidad por la escisión de un ácido graso que tiene 12 átomos de carbono de ACP. Por lo tanto, en varias modalidades, el microorganismo puede contener un gen exógeno que codifica una proteína con especificidad para catalizar una actividad enzimática (por ejemplo, la escisión de un ácido graso de una ACP, la reducción de un acil-CoA a un aldehído o un alcohol, o la conversión de un aldehído a un alcano) con respecto al número de átomos de carbono contenidos en el sustrato. La especificidad enzimática puede, en varias modalidades, ser por un sustrato que tiene de 8 a 34 átomos de carbono, preferentemente de 8 a 18 átomos de carbono, y con mayor preferencia de 10 a 14 átomos de carbono. La especificidad más preferida es por un sustrato con 12 átomos de carbono. En aun otra modalidad, la especificidad preferida es por un sustrato con 14 átomos de carbono. En otras modalidades, la especificidad puede ser por 20 a 30 átomos de carbono.

[0032] Al seleccionar la combinación deseada de genes exógenos a expresarse, se puede adaptar el producto generado por el microorganismo, el cual se puede extraer de la biomasa acuosa. Por ejemplo, en ciertas modalidades, el microorganismo contiene: (i) un gen exógeno que codifica una acil graso -ACP tioesterasa; y, opcionalmente, (ii) un gen exógeno que codifica para una acil graso-CoA/aldehído reductasa o una acil graso-CoA reductasa; y, opcionalmente, (iii) un gen exógeno que codifica para una aldehído graso reductasa o una aldehído graso decarbonilasa. El microorganismo, cuando se cultiva como se describe en la presente descripción, sintetiza un ácido graso unido a una ACP y la acil graso-ACP tioesterasa cataliza la escisión del ácido graso de la ACP para producir, mediante procesamiento enzimático adicional, una molécula de acil-CoA. Cuando está presente, el acil graso-CoA/aldehído reductasa cataliza la reducción de la acil-CoA a un alcohol. Del mismo modo, la acil graso-CoA reductasa, cuando está presente, cataliza la reducción de la acil-CoA a un aldehído. En aquellas modalidades en las que un gen exógeno que codifica una acil graso-CoA reductasa está presente y se expresa para rendir un producto aldehído, una aldehído graso reductasa, codificada por el tercer gen exógeno, cataliza la reducción del aldehído a un alcohol. Del mismo modo, un aldehído graso decarbonilasa cataliza la conversión del aldehído a un alcano, cuando está presente en el microorganismo.

Los genes que codifican dichas enzimas se pueden obtener a partir de células ya conocidas por exhibir una producción significativa de lípido, tales como *Chlorella protothecoides*. Los genes ya conocidos tienen un papel en la producción de lípido, por ejemplo, un gen que codifica una enzima que satura los dobles enlaces, se puede transformar de forma individual en las células receptoras. Sin embargo, en la práctica de la invención no

es necesario hacer suposiciones a priori en cuanto a qué genes son necesarios. Una biblioteca de ADN que contienen genes diferentes, tales como los ADNc a partir de un organismo de buena producción lipídica, se puede transformar en células receptoras. El ADNc está de preferencia en unión operativa con un promotor activo en microalgas. Las células receptoras de microalgas diferentes transformadas mediante una biblioteca reciben genes diferentes a partir de la biblioteca. Los transformantes que tienen la producción mejorada de lípido se identifican a través de métodos de detección conocidos en la técnica, tales como, por ejemplo, HPLC, cromatografía de gases, y los métodos de espectrometría de masa de análisis de hidrocarburos (para ejemplos de esos análisis, ver *Biomass and Bioenergy* Vol. 6. núm. 4. pp. 269-274 (1994); *Experientia* 38; 47-49 (1982); y *Phytochemistry* 65 (2004) 3159-3165). Estos transformantes se someten después a una transformación adicional con la biblioteca original y/o de forma opcional se cruzan para generar una ronda de organismos que tienen la producción de lípido mejorada. Los procedimientos generales de la evolución de los organismos completos para adquirir una propiedad deseada se describen en, por ejemplo, US 6,716,631. Dichos métodos implican, por ejemplo, la introducción de una biblioteca de fragmentos de ADN en una pluralidad de células, a través de lo cual al menos uno de los fragmentos se somete a la recombinación con un segmento en el genoma o un episoma de las células para producir células modificadas. Las células modificadas se detectan después por las células modificadas que han evolucionado hacia la adquisición de la función deseada. Los vectores y métodos para la transformación son análogos a aquellos discutidos en relación con la expresión de los genes de lipasa.

Además, las bibliotecas sustractivas se pueden usar para identificar los genes cuya transcripción se induce en diferentes condiciones, especialmente las condiciones empleadas en el cultivo de microorganismos para la producción de biodiesel, o para la producción de hidrocarburos útiles tales como materia prima para aplicaciones industriales. Las bibliotecas sustractivas contienen las secuencias de nucleótidos que reflejan las diferencias entre las dos muestras diferentes. Dichas bibliotecas se preparan mediante procedimientos que incluyen las etapas de desnaturalización e hibridación de poblaciones de polinucleótidos (por ejemplo, ARNm, ADNc, secuencias amplificadas) a partir de cada muestra. Las secuencias comunes a ambas muestras hibridan y se eliminan, dejando las secuencias que difieren entre las muestras. De este modo, se pueden identificar las secuencias que se inducen bajo condiciones particulares. Esta técnica se puede emplear, por ejemplo, para identificar genes útiles para incrementar la producción lipídica (por ejemplo, ácido graso) y, en particular la producción

lipídica bajo cualquiera de las condiciones de cultivo deseadas. La técnica de hibridación sustractiva se puede emplear también para identificar promotores, por ejemplo, promotores inducibles, útiles en constructos de expresión de acuerdo con la invención.

De este modo, por ejemplo, las bibliotecas sustractivas se pueden preparar a partir de cultivos de microorganismo cultivados autotróficamente (en la luz sin una fuente de carbono fijo) o heterotróficamente (en la oscuridad en la presencia de una fuente de carbono fijo). En particular, los genes heterótrofos se pueden inducir durante el crecimiento en la oscuridad en presencia de una fuente de carbono fijo y por lo tanto, podrían estar presentes en una biblioteca generada, mediante secuencias sustractivas de células autotróficas a partir de secuencias de células heterotróficas en la oscuridad. Las bibliotecas sustractivas se pueden preparar también a partir de cultivos a los que se ha añadido un sustrato de carbono particular, tal como glucosa, para identificar los genes que juegan un papel en la metabolización del sustrato. Las bibliotecas sustractivas preparadas a partir de cultivos cultivados en presencia de exceso de nitrógeno contra nitrógeno limitado se pueden usar para identificar los genes que controlan la división celular en comparación con la producción de la acumulación de hidrocarburo. La preparación de una biblioteca sustractiva a partir de un cultivo al que se han añadido los lípidos (por ejemplo ácidos grasos) puede ayudar a identificar los genes cuya sobreexpresión incrementa la producción de ácidos grasos. Más específicamente, la adición de ácidos grasos a un cultivo de células que puede usar los ácidos grasos añadidos conducirá a la regulación descendente de los genes sintéticos del ácido graso para regular de forma descendente la producción de ácidos grasos. La sobreexpresión de uno o más genes, tendrá el efecto contrario.

2. Aumento del flujo de carbono en la vía lipídica

Algunas microalgas producen cantidades significativas de metabolitos no lipídicos, tales como, por ejemplo, los polisacáridos. Debido a que la biosíntesis de polisacáridos puede usar una proporción significativa del total de la energía metabólica disponible para las células, la mutagénesis de las células que producen lípidos seguido por la detección de la producción de polisacárido reducida o eliminada genera nuevas cepas que son capaces de producir altos rendimientos de lípidos.

El ensayo de fenol: ácido sulfúrico detecta los carbohidratos (ver Hellebust, *Handbook of Phycological Methods*, Cambridge University Press, 1978; y Cuesta G., y otros, *J. Microbiol. Methods*, 2003 Jan;52(1):69-73). El ensayo de 1,6 dimetilmetileno azul detecta

los polisacáridos aniónicos. (ver, por ejemplo, *Braz. J. Med. Biol. Res.* 1999 may;32(5):545-50; *Clin. Chem.* 1986 nov;32(11):2073-6).

Los polisacáridos pueden analizarse, además, por métodos como HPLC, cromatografía de exclusión por tamaño, y cromatografía de intercambio aniónico (ver, por ejemplo, Prosky L, Asp N, Schweizer TF, DeVries JW & Furda I (1988) Determination of insoluble, soluble and total dietary fiber in food and food products: Interlaboratory study. *Journal of the Association of Official Analytical Chemists* 71, 1017±1023; *Int J Biol Macromol.* 2003 nov;33(1-3):9-18). Los polisacáridos pueden detectarse, además, usando la electroforesis en gel (ver, por ejemplo, *Anal Biochem.* 2003 oct 15;321(2):174-82; *Anal Biochem.* 2002 ene 1;300(1):53-68).

V. MÉTODOS DE CULTIVO DE MICROORGANISMOS

A. Biorreactor

Los microorganismos se cultivan con dos propósitos, el de realizar manipulaciones genéticas y para la producción posterior de hidrocarburos (*por ejemplo*, lípidos, ácidos grasos, aldehídos, alcoholes, y alcanos). El primer tipo de cultivo se realiza a pequeña escala e inicialmente, al menos, bajo condiciones en las que el microorganismo inicial pueda crecer. Por ejemplo, si el microorganismo de partida es un fotoautótrofo el cultivo inicial se lleva a cabo en presencia de luz. Las condiciones de cultivo se pueden cambiar si el microorganismo se desarrolla o modifica para crecer de forma independiente de la luz. El cultivo para los propósitos de producción de hidrocarburos se lleva a cabo de preferencia a gran escala en un biorreactor (*por ejemplo*, 10,000 l, 40,000 l, 100,000 l o biorreactores mayores). Los microorganismos (*por ejemplo*, las microalgas) de la invención son típicamente cultivadas en los métodos de la invención en medios líquidos dentro de un biorreactor, típicamente en la ausencia de luz (crecimiento heterótrofo).

[0033] El biorreactor o fermentador se usa para cultivar células de microalgas durante las varias fases de su ciclo fisiológico. Los biorreactores ofrecen muchas ventajas para su uso en el crecimiento y la propagación heterotrófica. Para producir biomasa para la producción de lípidos, las microalgas se cultivan preferentemente en grandes cantidades de líquido, tales como cultivos en suspensión como un ejemplo. Los biorreactores, tales como los fermentadores de acero inoxidable pueden alojar volúmenes muy grandes de cultivo (biorreactores de 40,000 litros y más de capacidad son usados en varias modalidades de la invención). Los biorreactores típicamente permiten, además, el control de las condiciones de

cultivo tales como temperatura, pH, tensión de oxígeno, y niveles de dióxido de carbono. Por ejemplo, los biorreactores se configuran típicamente, por ejemplo usando puertos unidos a tuberías, para permitir que los componentes gaseosos como el oxígeno o el nitrógeno sean burbujeados a través del líquido de cultivo. Otros parámetros del cultivo, tales como el pH del medio de cultivo, la identidad y concentración de los oligoelementos, y otros constituyentes del medio, se pueden manipular fácilmente usando el biorreactor.

[0034] Los biorreactores se pueden configurar para que el medio de cultivo circule a través del biorreactor durante todo el período de tiempo en que las microalgas se reproducen e incrementan su número. En algunas modalidades, por ejemplo, el medio puede ser infundido dentro del biorreactor después de la inoculación, pero antes que las células alcancen la densidad deseada. En otros casos, el biorreactor se llena con el medio de cultivo al inicio de un cultivo, y no se infunde más medio de cultivo después que el cultivo es inoculado. En otras palabras, la biomasa de microalgas se cultiva en un medio acuoso durante un período de tiempo en el que las microalgas se reproducen e incrementan su número; sin embargo, durante este período de tiempo las cantidades del medio de cultivo acuoso no fluyen a través del biorreactor. Así, en algunas modalidades el medio de cultivo acuoso no fluye a través del biorreactor después de la inoculación.

[0035] Los puertos de los biorreactores se pueden usar para introducir, o extraer, de la cámara del reactor que contiene las microalgas, gases, sólidos, semisólidos y líquidos. Aunque muchos biorreactores tienen más de un puerto (por ejemplo, uno para la entrada del medio y otro para el muestreo), no es necesario que solamente una sustancia entre o salga por el puerto. Por ejemplo, se puede usar un puerto para que el medio de cultivo fluya o una fuente de carbono adicional en el biorreactor y después usarlo para muestrear la entrada de gas, salida de gas, u otros propósitos. Preferentemente, un puerto de muestreo se puede usar repetidamente sin que se altere o comprometa la naturaleza axénica del cultivo. Un puerto de muestreo puede configurarse con una válvula u otro dispositivo que permita detener o iniciar la circulación de la muestra o proporcionar un medio de muestreo continuo. Típicamente, los biorreactores tienen al menos un puerto que permite la inoculación de un cultivo y dicho puerto puede usarse, además, con otros fines tales como la entrada de medio o gas.

Los biorreactores permiten manipular el contenido de gas del cultivo del microorganismo (*por ejemplo*, microalgas). Para ilustrar, parte del volumen de un biorreactor

puede ser gas en lugar de líquido, y las entradas de gas del biorreactor permiten el bombeo de gases hacia dentro del biorreactor. Los gases que se pueden bombear beneficiosamente dentro de un biorreactor incluyen aire, mezclas de aire/CO₂, gases nobles, tales como argón y otros gases. Los biorreactores están típicamente equipados para permitir que el usuario controle la velocidad de entrada del gas en el biorreactor. Como se denotó anteriormente, al aumentar el flujo de gas dentro del biorreactor aumenta la mezcla del cultivo.

[0036] El incremento del flujo gaseoso afecta también la turbidez del cultivo. La turbulencia se puede lograr colocando un puerto de entrada de gas debajo del nivel del medio de cultivo acuoso, de manera tal que el gas que entre al biorreactor burbujee hacia la superficie del cultivo. Uno o más puertos de salida de gas permiten que el gas escape, impidiendo de este modo que la presión aumente gradualmente en el biorreactor. Preferiblemente, el puerto de salida de gas conduce hacia una válvula “en una sola dirección” que impide que microorganismos contaminantes entren al biorreactor.

B. Medio

El medio de cultivo para el cultivo de microorganismos, que incluye microalgas, contiene, típicamente, componentes tales como una fuente fija de nitrógeno, una fuente fija de carbono, elementos traza, opcionalmente un tampón para mantener el pH, y fosfato (típicamente proporcionado como una sal de fosfato). Otros componentes pueden incluir sales tales como cloruro de sodio, particularmente para microalgas marinas. Las fuentes de nitrógeno incluyen fuentes de nitrógeno orgánico e inorgánico, incluyendo, por ejemplo, pero sin limitación, nitrógeno molecular, nitrato, sales de nitrato, amoníaco (puro o en forma de sal, como (NH₄)₂SO₄ y NH₄OH), proteínas (y aminoácidos), harina de soja, licor de maíz fermentado, y extracto de levadura. Los ejemplos de oligoelementos incluyen zinc, boro, cobalto, cobre, manganeso, y molibdeno en, por ejemplo, las formas respectivas de ZnCl₂, H₃BO₃, CoCl₂·6H₂O, CuCl₂·2H₂O, MnCl₂·4H₂O y (NH₄)₆Mo₇O₂₄·4H₂O.

[0037] Los medios de crecimiento sólidos y líquidos generalmente están disponibles de una amplia variedad de fuentes, y las instrucciones para la preparación de un medio en particular que sea adecuado para una amplia gama de variedades de microorganismos, pueden encontrarse, por ejemplo, en el sitio web <http://www.utex.org/>, un sitio mantenido por la Universidad de Texas en Austin, 1 University Station A6700, Austin, Texas 78712-0183, para su colección de cultivo de algas (UTEX). Por ejemplo, varios

medios de agua dulce y agua salada incluyen los descritos en la publicación PCT núm. 2008/151149, incorporada en la presente como referencia.

En un ejemplo particular, el medio de proteosa es adecuado para cultivos axénicos, y un volumen de 1 litro de medio de proteosa (pH ~6.8) puede prepararse por la adición de 1g de proteosa peptona a 1 litro de medio Bristol. El medio Bristol comprende 2.94 mM NaNO_3 , 0.17 mM $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.3 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.43 mM K_2HPO_4 , 1.29 mM KH_2PO_4 y 1.43 mM NaCl en una solución acuosa. La solución se cubre y se somete a autoclave, y después se almacena a temperatura refrigerada antes del uso. Otro ejemplo es el medio de aislamiento de *Prototheca* (PIM), el cual está formado por 10g/l de hidrógeno ftalato de potasio (KHP), 0.9g/l de hidróxido de sodio, 0.1g/l de sulfato de magnesio, 0.2g/l de hidrógeno fosfato de potasio, 0.3g/l de cloruro de amonio, 10g/l de glucosa, 0.001g/l de hidrocloreuro de tiamina, 20g/l de agar, 0.25g/l de 5-fluorocitosina, en un intervalo de pH entre 5.0 y 5.2 (ver Pore, 1973, *App. Microbiology*, 26:648-649). Otros medios adecuados para usar con los métodos de la invención se pueden identificar fácilmente mediante la consulta de la URL identificada anteriormente, o mediante la consulta de otras organizaciones que mantienen cultivos de microorganismos, tales como SAG, CCAP o CCALA. SAG se refiere a la Colección de Cultivos de Algas en la Universidad de Göttingen (Göttingen, Alemania), CCAP se refiere a la colección de cultivos de algas y protozoos gestionado por la Asociación Escocesa de Ciencias Marinas (Escocia, Reino Unido), y CCLA se refiere a la colección de cultivos de algas de laboratorio en el Instituto de Botánica (Třeboň, República Checa). Adicionalmente, la patente de Estados Unidos núm. 5,900,370 describe formulaciones de medios y condiciones adecuadas para la fermentación heterótrofa de las especies de *Prototheca*.

[0038] Para la producción de aceite (lípidos), es importante la selección de una fuente fija de carbono, así como el costo de la fuente fija de carbono debe ser lo suficientemente bajo para que la producción de aceite resulte económica. Así, aunque las fuentes de carbono adecuadas incluyen, por ejemplo, acetato, floridosida, fructosa, galactosa, ácido glucurónico, glucosa, glicerol, lactosa, manosa, N-acetilglucosamina, ramnosa, sacarosa, glucosa y/o xilosa, la selección de materias primas que contengan dichos compuestos es un aspecto importante de los métodos de la invención. Algunas especies de microorganismos (*por ejemplo*, las microalgas) pueden crecer mediante la utilización de una fuente de carbono fijo, tales como la glucosa o acetato en la ausencia de luz. Este crecimiento se conoce como crecimiento heterótrofo. Para *Chlorella protothecoides*, por ejemplo, el

crecimiento heterótrofo resulta en alta producción de biomasa y acumulación de alto contenido en lípidos en las células. Otras materias primas adecuadas incluyen, por ejemplo, licor negro, almidón de maíz, material celulósico despolimerizado, suero de leche, melaza, jugo de caña espeso, papa, sorgo, sacarosa, azúcar de remolacha, caña de azúcar, arroz y trigo. Las fuentes de carbono pueden proporcionarse además como una mezcla de sacarosa y pulpa de azúcar de remolacha despolimerizada. Una o más fuentes de carbono pueden ser suministradas a una concentración de al menos aproximadamente 50 μ M, al menos aproximadamente 100 μ M, al menos aproximadamente 500 μ M, al menos aproximadamente 5mM, al menos aproximadamente 50 mM, al menos aproximadamente 500mM, de una o más fuentes fijas de carbono suministradas exógenamente. Las fuentes de carbono de particular interés para los propósitos de la presente invención incluyen celulosa (en forma despolimerizada), glicerina, sacarosa (en forma de jugo de caña o melazas) y sorgo.

Algunos microorganismos crecen de forma natural en o se pueden manipular para crecer en una fuente de carbono fijo que es una fuente heterogénea de compuestos tales como los residuos urbanos, aguas residuales tratadas, aguas residuales, y otras fuentes de carbono fijo y otros nutrientes, tales como sulfatos, fosfatos, y nitratos. El componente de las aguas residuales es una fuente de nutrientes en la producción de hidrocarburos, y el cultivo proporciona una fuente barata de hidrocarburos.

En algunos métodos de crecimiento heterótrofos, los microorganismos pueden cultivarse usando biomasa celulósica como una materia prima. La biomasa celulósica (por ejemplo, rastrojo como rastrojo de maíz) es barata y fácilmente disponible; sin embargo los intentos de usar este material como una materia prima para la levadura han fracasado. En particular, se ha encontrado que estas materias primas tienen un efecto inhibitor en el crecimiento de la levadura, y la levadura no puede usar los azúcares de 5 carbonos producida a partir de materiales celulósicos (por ejemplo, xilosa a partir de hemicelulosa). En contraste, las microalgas pueden crecer en material celulósico procesado. En correspondencia, los métodos de crecimiento heterótrofo incluyen un método de cultivar una microalga en presencia de un material celulósico y/o un azúcar de 5-carbonos. Los materiales celulósicos generalmente incluyen:

Componente	Porcentaje peso seco
Celulosa	40-60%

Hemicelulosa	20-40%
Lignina	10-30%

[0039] Los materiales celulósicos adecuados incluyen los residuos de cosechas energéticas herbáceas y leñosas, así como cultivos agrícolas, es decir, partes de plantas, principalmente tallos y hojas no eliminadas de los campos con el alimento principal o el producto de fibra. Los ejemplos incluyen rastrojo de maíz (tallos, hojas, cáscara, y mazorcas), paja de trigo, y paja de arroz. Los azúcares de 5 carbonos que se producen a partir de estos materiales incluyen la xilosa.

Se ha demostrado, por ejemplo, que la *Chlorella protothecoides* presenta niveles superiores de productividad cuando se cultiva en una combinación de glucosa y xilosa que cuando se cultiva en ya sea glucosa o solo en xilosa. Este efecto sinérgico proporciona una ventaja significativa, ya que permite el cultivo de *Chlorella* en combinaciones de xilosa y glucosa, tal como material celulósico.

En otra modalidad de los métodos de la invención, la fuente de carbono es sacarosa, incluyendo una materia prima compleja que contiene sacarosa, tal como el jugo espeso de la caña del procesamiento de la caña de azúcar o melazas. En una modalidad, el medio de cultivo incluye, además, al menos una enzima de utilización de la sacarosa. En algunos casos, el medio de cultivo incluye una sacarosa invertasa. En una modalidad, la enzima sacarosa invertasa es una enzima sacarosa invertasa secretable codificada por un gen sacarosa invertasa exógeno expresado por la población de microorganismos. De este modo, en algunos casos, las microalgas se han modificado genéticamente para expresar una enzima de utilización de la sacarosa, tal como un portador de sacarosa, una sacarosa invertasa, una hexoquinasa, una glucoquinasa, o una fructoquinasa.

Las materias primas complejas que contienen sacarosa incluyen a los desechos de melaza del procesamiento de la caña de azúcar; el uso de este producto de bajo valor del procesamiento de la caña de azúcar puede proporcionar ahorros significativos de los costos de producción de lípidos /aceite. El sorgo es otra materia prima compleja que contiene sacarosa que es útil en los métodos de la invención, incluyendo sirope de sorgo y sorgo puro. El sirope de sorgo se produce a partir del jugo de la caña de sorgo dulce, y su perfil de azúcar consiste de principalmente glucosa (dextrosa), fructosa y sacarosa.

C. Incremento del rendimiento de lípidos

Para la producción de lípidos/aceite de acuerdo con los métodos de la invención, se prefiere cultivar las células en la oscuridad, como es el caso, por ejemplo, cuando se usan fermentadores muy grandes (40,000 litros o de mayor capacidad) que no permiten a la luz alcanzar el cultivo. Como un ejemplo, un inóculo de las células de microalgas productoras de lípidos se introduce en el medio; hay un período de latencia (fase lag) antes de que las células comiencen a propagarse. Tras la fase de latencia, la velocidad de propagación aumenta a un ritmo constante y entra en la fase log o exponencial. La fase exponencial, a su vez, es seguida por una desaceleración de la propagación debido a la disminución de nutrientes como el nitrógeno, el aumento de sustancias tóxicas, y los mecanismos de detección de quórum. Después de esta desaceleración, se detiene la propagación y las células entran en una fase estacionaria o estado constante, dependiendo del ambiente particular proporcionado a las células. Para la obtención de biomasa rica en lípidos, el cultivo por regla general se cosecha bien después del final de la fase exponencial, proceso que puede terminar tempranamente al permitir que el nitrógeno u otro nutriente clave (que no sea carbono) se agote, lo que obliga a las células a convertir en lípidos las fuentes de carbono presentes en exceso. Los parámetros de cultivo pueden ser manipulados para optimizar la producción total de aceite, la combinación de especies de lípidos producidos y/o la producción de un aceite específico.

Las condiciones del proceso pueden ajustarse para aumentar el rendimiento de los lípidos adecuados para usar como biodiesel u otras moléculas objetivo, y/o para reducir el costo de producción. Por ejemplo, en ciertas modalidades, un microbio (*por ejemplo*, una microalga) se cultiva en presencia de una concentración limitante de uno o más nutrientes, tales como, por ejemplo, nitrógeno. Esta condición tiende a aumentar el rendimiento de lípidos microbianos, por encima del rendimiento de lípidos microbianos en un cultivo en el que el nitrógeno se suministra en exceso. En modalidades particulares, el aumento en el rendimiento de los lípidos es al menos aproximadamente: 10%, 20%, 30%, 40%, 50%, 75%, 100%, 200%, 300%, 400%, o 500%. El microbio puede ser cultivado en presencia de una cantidad limitante de un nutriente para una parte del período total de cultivo o para todo el período. En modalidades particulares, la concentración de nutrientes se mueve entre una concentración limitante y una concentración no limitante, al menos dos veces durante el período total de cultivo. Adicionalmente, como se describió anteriormente, ciertas materias primas de carbono fijo, tal como, glicerina, pueden emplearse para aumentar el porcentaje de

peso de la célula, o sea lípido, en relación con cantidades comparables de otras materias primas de carbono fijo.

Para incrementar el rendimiento lipídico, el ácido acético se puede emplear en la materia prima para un microbio que produce lípido (por ejemplo, una microalga). El ácido acético se alimenta directamente en el punto de metabolismo que inicia la síntesis de ácidos grasos (es decir, acetil-CoA); de este modo el suministro de ácido acético en el cultivo puede incrementar la producción de ácidos grasos. Generalmente, el microbio se cultiva en presencia de una cantidad suficiente de ácido acético para incrementar el rendimiento de lípidos microbianos y/o rendimiento de ácidos grasos microbianos, en particular, sobre el rendimiento de lípidos microbianos (por ejemplo, ácido graso) en la ausencia de ácido acético.

[0040] En otra modalidad, el rendimiento en la producción de lípidos se incrementa al cultivar un microorganismo productor de lípidos (por ejemplo, microalgas) en presencia de uno o más cofactor(es) de una enzima de la vía metabólica de los lípidos (por ejemplo, una enzima para la síntesis de ácidos grasos). En general, la concentración del(de los) cofactor(es) es suficiente para aumentar el rendimiento de lípidos microbianos (por ejemplo, ácidos grasos) por encima del rendimiento de lípidos microbianos en ausencia del (de los) cofactor(es). En una modalidad particular, el(los) cofactor(es) se proporcionan al cultivo mediante la inclusión de un microbio en el cultivo (por ejemplo, microalgas) que contiene un gen exógeno que codifica para el(los) cofactor(es). Como alternativa, el(los) cofactor(es) se pueden proporcionar a un cultivo mediante la inclusión de un microbio (por ejemplo, microalgas) que contiene un gen exógeno que codifica para una proteína que participa en la síntesis del cofactor. En ciertas modalidades, los cofactores adecuados incluyen cualquier vitamina requerida por una enzima de la vía de los lípidos, tales como, por ejemplo: biotina, pantotenato. Los genes que codifican para los cofactores adecuados para su uso en la invención o que participan en la síntesis de tales cofactores son bien conocidos y se pueden introducir en los microbios (por ejemplo, microalgas), usando constructos y técnicas como las descritas anteriormente.

D. Biomasa de microalgas con alto contenido de aceite

La biomasa de microalgas con un alto porcentaje de aceites/lípidos acumulados por peso seco de las células se ha generado utilizando diferentes métodos de cultivo, que se conocen en el arte. La biomasa de microalgas con un alto porcentaje de aceites/lípidos

acumulados es útil de acuerdo con la presente invención. Li y otros describen cultivos de *Chlorella vulgaris* con hasta 56.6% de lípido por peso seco celular (DCW) en cultivos estacionarios que crecen bajo condiciones autotróficas usando altas concentraciones de hierro (Li y otros, *Bioresource Technology* 99(11):4717-22 (2008)). Rodolfi y otros, describen cultivos de *Nanochloropsis sp.* y *Chaetoceros calcitrans* con 60% de lípido por DCW y 39.8% de lípido por DCW, respectivamente, que crecen en un fotobiorreactor bajo condiciones de deficiencia de nitrógeno (Rodolfi y otros, *Biotechnology & Bioengineering* (2008) [jun18 Epub antes de la impresión]). Solovchenko y otros, describen cultivos de *Parietochloris incise* con aproximadamente 30% de acumulación de lípidos (por DCW) cuando se cultivan fototrópicamente y bajo condiciones de bajo nitrógeno (Solovchenko y otros, *Journal of Applied Phycology* 20:245-251 (2008)). La *Chlorella protothecoides* puede producir hasta 55% de lípido (DCW) cultivada bajo ciertas condiciones heterótrofas con deficiencia de nitrógeno (Miao yd Wu, *Bioresource Technology* 97:841-846 (2006)). Otras especies de *Chlorella* que incluyen *Chlorella emersonii*, *Chlorella sorokiniana* y *Chlorella minutissima* se describen por tener acumulado hasta 63% de aceite (DCW) cuando se cultivan en biorreactores de tanque agitado bajo condiciones de medio bajo en nitrógeno (Illman y otros, *Enzyme and Microbial Technology* 27:631-635 (2000)). Aún un por ciento superior de acumulación de lípido por peso seco celular se reportó, incluyendo 70% de acumulación de lípidos (DCW) en cultivos de *Dumaliella tertiolecta* que crecen en condiciones de NaCl aumentado (Takagi y otros, *Journal of Bioscience and Bioengineering* 101(3): 223-226 (2006)) y 75% acumulación de lípidos en cultivos de *Botryococcus braunii* (Banerjee y otros, *Critical Reviews in Biotechnology* 22(3): 245-279 (2002)).

[0041] La biomasa de microalgas generada por los métodos de cultivo descritos en la presente y útiles de conformidad con la presente invención, comprende al menos 10% de aceite de microalgas por peso seco. En algunas modalidades, la biomasa de microalgas comprende al menos 15%, al menos 25%, al menos 35%, al menos 45%, al menos 55%, o al menos 60% de aceite de microalgas por peso seco. En algunas modalidades, la biomasa de microalgas contiene de 10-90% de aceite de microalgas, de 25-75% de aceite de microalgas, de 40-75% de aceite de microalgas, o de 50-70% de aceite de microalgas por peso seco.

E. Cultivo de microorganismos bajo condiciones inducidas

Como se describe en la presente, en ciertas modalidades de la presente invención, los microorganismos se cultivan bajo condiciones inducidas, es decir, por el suministro de un estímulo. Generalmente, este se lleva a cabo como sigue: El microorganismo se cultiva durante un primer período de tiempo suficiente para incrementar la densidad celular. Después, el estímulo se proporciona y los microorganismos se cultivan durante un segundo período de tiempo. Durante el segundo período de tiempo ocurre el efecto deseado del estímulo, por ejemplo, la inducción de un gen exógeno o producción de lipídica incrementada. El cultivo de los microorganismos durante el segundo período de tiempo puede ser en presencia continua del estímulo. Como alternativa, el estímulo no se podría proporcionar o sólo parcialmente, (es decir, por un tiempo limitado) durante el cultivo del microorganismo por el segundo período de tiempo.

F. Almacenamiento de microorganismos antes de la extracción de lípido

En algunos casos se prefiere almacenar los microorganismos cultivados por un período de tiempo antes de someterlos a los procesos de extracción descritos más abajo. En algunos métodos de la invención, los microorganismos, producidos mediante un proceso de cultivo como es descrito en la presente descripción se almacenan opcionalmente por un período de tiempo entre la terminación del proceso de cultivo y el lisado del microorganismo. En algunos casos, el microorganismo se almacena durante al menos una hora entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado. En otros casos, el microorganismo se almacena durante al menos dos horas, al menos tres horas, al menos cuatro horas, al menos cinco horas, al menos seis horas, al menos siete horas, al menos ocho horas, al menos nueve horas, al menos diez horas, al menos once horas, al menos doce horas, al menos trece horas, al menos catorce horas, al menos quince horas, al menos dieciséis horas, al menos diecisiete horas, al menos dieciocho horas, al menos diecinueve horas, al menos veinte horas, al menos veintiuna horas, al menos dos horas, al menos veintitrés horas, o durante al menos veinticuatro horas entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado. En algunos casos, el microorganismo se almacena durante al menos treinta y seis horas entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado. En algunos casos, el microorganismo se almacena durante al menos cuarenta y ocho horas entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado, o por períodos de tiempo más largos. En otros casos, el

microorganismo se almacena durante al menos sesenta o setenta y dos horas, o más, entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado.

Los microorganismos preparados en un proceso de cultivo se almacenan opcionalmente a una temperatura por debajo de 15 grados Celsius entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado. En algunos casos, los microorganismos se almacenan a una temperatura por debajo de 14°C, por debajo de 13°C, por debajo de 12°C, por debajo de 11°C, por debajo de 10°C, por debajo de 9°C, por debajo de 8°C, por debajo de 7°C, por debajo de 6°C, por debajo de 5°C, por debajo de 4°C, por debajo de 3°C, por debajo de 2°C, o por debajo de 1°C entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado. En algunos casos, el microorganismo se almacena a una temperatura por encima de 30 grados Celsius entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado. En algunos casos, el microorganismo se almacena a una temperatura por encima de 40 grados Celsius entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado. En otros casos, el microorganismo se almacena a una temperatura por encima de 31°C, por encima de 32°C, por encima de 33°C, por encima de 34°C, por encima de 35°C, por encima de 36°C, por encima de 37°C, por encima de 38°C, por encima de 39°C, por encima de 41°C, por encima de 42°C, por encima de 43°C, por encima de 44°C, por encima de 45°C, por encima de 46°C, por encima de 47°C, por encima de 48°C, por encima de 49°C, o por encima de 50°C entre la terminación del proceso de cultivo y el lisado del microorganismo cultivado.

En algunos casos, el almacenamiento o envejecimiento de los microorganismos cultivados se usa para romper las células para facilitar la extracción de petróleo. Con el almacenamiento, la estructura celular se podría debilitar lo suficiente como para provocar que los contenidos de las células comiencen a fugarse, o permitir el paso de reactivos y otros materiales dentro de las células para facilitar la extracción de los contenidos de lípidos. El almacenamiento se puede usar en este contexto, en combinación con otros métodos de lisis que se describen más abajo.

En algunos casos, el microorganismo durante el almacenamiento se somete a agitación. La agitación puede ser además, para cualquier combinación de las condiciones de almacenamiento establecidas anteriormente. La agitación se puede hacer en un agitador, vórtice, o similares. La agitación puede resultar también de las fuerzas de cizalla presentes durante la centrifugación de alta fuerza g. Como alternativa, el microorganismo no se agita

durante el almacenamiento. Los microorganismos se pueden almacenar en un biorreactor u otro recipiente de cultivo, u opcionalmente, transferir a un recipiente de almacenamiento separado.

VI. MÉTODOS DE EXTRACCIÓN DE LÍPIDO DEL MICROORGANISMO

5 En un aspecto, la presente invención se dirige a un proceso para la extracción, recuperación, aislamiento u obtención de lípidos a partir de microorganismos. El proceso de la presente invención se aplica a la extracción de una variedad de lípidos a partir de una variedad de microorganismos.

A. Lisis de las células

10 Los lípidos intracelulares producidos en los microorganismos se extraen después de la lisis de las células del microorganismo. Una vez extraídos, los lípidos pueden ser refinados adicionalmente para producir lípidos de alta pureza.

Después de finalizar el cultivo, los microorganismos se pueden separar del caldo de fermentación, preferentemente sin una etapa de secado tal como el secado de tambor, secado por pulverización, secado de bandeja, secado al vacío y otras etapas que eliminan prácticamente toda el agua extracelular e intracelular del caldo. Opcionalmente, la separación se efectúa por centrifugación para generar una pasta concentrada. La centrifugación no elimina las cantidades significativas de agua intracelular proveniente de los microorganismos y no es una etapa de secado. La biomasa se puede lavar después con una solución de lavado (por ejemplo, agua DI) para deshacerse del caldo de fermentación y los residuos celulares. Opcionalmente, la biomasa microbiana lavada también se puede secar (secado en el horno, liofilizada, etc.) antes de la ruptura celular. Por otra parte, las células se pueden lisar sin separación de parte o todo el caldo de fermentación, cuando se completa la fermentación. Por ejemplo, las células pueden estar en una proporción de menos de 1:1, v:v de células con respecto al líquido extracelular cuando ellas se lisan.

[0042] Los microorganismos que contiene un lípido se pueden lisar para producir un lisado. Como se detalla en la presente, la etapa de lisar un microorganismo (también conocida como lisis celular) se puede lograr por cualquier medio conveniente, incluyendo lisis inducida por calor, adición de una base, adición de un ácido, uso de enzimas como las proteasas y enzimas de degradación de polisacáridos tales como las amilasas, usando ultrasonido, lisis mecánica, usando choque osmótico, infección con un virus lítico, y/o

expresión de uno o más genes líticos. La lisis se realiza para liberar las moléculas intracelulares producidas por el microorganismo. Cada uno de estos métodos para la lisis de un microorganismo se puede usar como un método único o en combinación.

[0043] La magnitud de la ruptura celular se puede observar por análisis microscópico. Usando uno o varios de los métodos descritos en la presente, se observa típicamente más del 70% de ruptura celular. Preferentemente la ruptura celular es más del 80%, con más preferencia más del 90% y con la máxima preferencia aproximadamente 100%.

En modalidades particulares, el microorganismo se lisa después del crecimiento, por ejemplo para aumentar la exposición de lípido celular a un catalizador para la transesterificación, tal como una lipasa o un catalizador químico, expresado como se describió anteriormente. El tiempo de expresión de la lipasa (por ejemplo, a través de un promotor inducible), lisis celular, y el ajuste de las condiciones de reacción de transesterificación (por ejemplo, la eliminación de agua, la adición de alcohol, etc.) se puede ajustar para optimizar el rendimiento de los ésteres de ácidos grasos a partir de la transesterificación mediada por lipasa. A continuación se describen una serie de técnicas de lisis. Estas técnicas se pueden usar individualmente o en combinación.

1. Lisis inducida por calor

[0044] En una modalidad preferida de la presente invención, la etapa de lisis de un microorganismo comprende calentar una suspensión celular que contiene el microorganismo. En esta modalidad, el caldo de fermentación que contiene los microorganismos (o una suspensión de microorganismos aislados a partir del caldo de fermentación) se calienta hasta que los microorganismos, es decir, las paredes celulares y las membranas de los microorganismos se degraden o se rompan. Típicamente, las temperaturas aplicadas son al menos 50°C. Temperaturas más altas, tales como, al menos 60°C, al menos 70°C, al menos 80°C, al menos 90°C, al menos 100°C, al menos 110°C, al menos 120°C, al menos 130°C o superiores, se usan para una lisis celular más eficiente. El Ejemplo 7 describe una modalidad de lisis usando tratamiento térmico.

[0045] El lisado de las células mediante tratamiento térmico se puede realizar hirviendo al microorganismo. Como alternativa, el tratamiento por calor (sin hervir) se puede realizar en una autoclave (ver Ejemplo 6). El lisado tratado térmicamente se puede enfriar para su posterior tratamiento.

[0046] La ruptura celular también se puede realizar por tratamiento con vapor, es decir, mediante la adición de vapor a presión. El tratamiento con vapor de las microalgas para la ruptura celular se describe, por ejemplo, en la patente de los Estados Unidos núm. 6,750,048.

5 2. **Lisis usando una base**

[0047] En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo comprende la adición de una base a una suspensión celular que contiene el microorganismo.

10 [0048] La base debe ser lo suficientemente fuerte para hidrolizar al menos una porción de los compuestos proteicos de los microorganismos usados. Las bases que son útiles para solubilizar las proteínas se conocen en la materia de la química. Las bases ejemplares que son útiles en los métodos de la presente invención incluyen, pero sin limitarse a, hidróxidos, carbonatos y bicarbonatos de litio, sodio, potasio, calcio y mezclas de los mismos. Una base preferida es el KOH. Los ejemplos 6 y 7 describen modalidades de lisis
15 celular usando KOH.

3. **Lisis ácida**

 [0049] En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo comprende la adición de un ácido a una suspensión celular que contiene el microorganismo. La lisis ácida se puede llevar a cabo usando un ácido a una
20 concentración de 10 a 500 mM o preferentemente de 40 a 160 mM. La lisis ácida se realiza preferentemente a temperatura por encima de la temperatura ambiente (por ejemplo, 40-160°, y preferentemente una temperatura de 50-130°). Para temperaturas moderadas (por ejemplo, temperatura ambiente a 100°C y particularmente de temperatura ambiente a 65°C), el tratamiento con ácido puede ser útil en combinación con sonicación u otros métodos de o
25 ruptura celular. El Ejemplo 7 describe las modalidades de la lisis celular usando lisis ácida.

4. **Lisis de las células usando enzimas**

 [0050] En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo comprende la lisis del microorganismo mediante el uso de una enzima. Las enzimas de preferencia para la lisis de un microorganismo son las proteasas y las enzimas
30 que degradan polisacáridos tales como hemicelulasa (por ejemplo, la hemicelulasa de *Aspergillus niger*; Sigma Aldrich, St. Louis, MO; # H2125), pectinasa (por ejemplo, la

pectinasa de *Rhizopus sp.*; Sigma Aldrich, St. Louis, MO; # P2401), Mannaway 4.0 1 (Novozymes), celulasa (por ejemplo, la celulasa de *Trichoderma viride*, Sigma Aldrich, St. Louis, MO; #C9422), driselasa (por ejemplo, driselasa de *Basidiomycetes sp.*; Sigma Aldrich, St. Louis, MO; #D9515. El Ejemplo 7 describe modalidades de lisis celular usando enzimas.

5 a) Celulasas

En una modalidad preferida de la presente invención, una celulasa para lisar un microorganismo es una enzima de degradación de polisacáridos, opcionalmente de *Chlorella* o un virus de *Chlorella*. El Ejemplo 7 describe las modalidades de la lisis celular usando una celulasa. Otro ejemplo de una enzima de degradación de polisacáridos que se puede usar es
10 Mannaway, como se describe en los Ejemplos.

b) Proteasas

Las proteasas tal como la proteasa de *Streptomyces griseus*, quimotripsina, proteinasa K, proteasas enumeradas en Degradation of Polylactide by Commercial Proteases, Oda Y y otros, Journal of Polymers and the Environment, volumen 8, número 1, Enero 2000, pp. 29-32(4), y otras proteasas se pueden usar para lisar los microorganismos. Otras proteasas que se pueden usar incluyen Alcalasa 2.4 FG (Novozymes) y Flavourzima 100 L (Novozymes), como se describe en los Ejemplos.

c) Combinaciones

También se puede utilizar c[0051] cualquier combinación de una proteasa y una
20 enzima de degradación de polisacáridos, incluyendo cualquier combinación de las proteasas y las enzimas que degradan polisacáridos anteriores.

5. Lisis de células usando ultrasonido

[0052] En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo se realiza mediante ultrasonido, es decir, sonicación. De esta manera,
25 las células pueden también lisarse con sonido de alta frecuencia. El sonido se puede producir por vía electrónica y se transporta a través de una punta metálica a una suspensión celular apropiadamente concentrada. Esta sonicación (o ultrasonicación) altera la integridad celular basada en la formación de cavidades en la suspensión celular. El Ejemplo 6 describe un método para la lisis celular usando ultrasonido. El Ejemplo 7 describe las modalidades de
30 lisis celular usando sonicación.

6. Lisis mecánica

En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo se realiza mediante lisis mecánica. Las células se pueden lisar mecánicamente y, opcionalmente, homogeneizarse para facilitar la recolección de los hidrocarburos (por ejemplo, los lípidos). Por ejemplo, un disruptor de presión se puede utilizar para bombear una mezcla que contiene las células a través de una válvula de orificio restringido. Se aplica alta presión (hasta 1500 bar), seguido por una expansión instantánea a través de una boquilla de salida. La ruptura celular se lleva a cabo por tres mecanismos diferentes: impacto en la válvula, alta cizalla del líquido en el orificio, y la repentina caída de presión después de la descarga, causando una explosión de la célula. El método libera las moléculas intracelulares.

[0053] Alternativamente se puede usar un molino de bolas. En un molino de bolas, las células se agitan en suspensión con pequeñas partículas abrasivas, tales como gránulos. Las células se rompen debido a las fuerzas de cizalla, el molido entre los gránulos, y las colisiones con los gránulos. Los gránulos rompen las células para liberar los contenidos celulares. Las células también se pueden romper por las fuerzas de cizalla, tal como con el uso de mezclado (tal como con una licuadora de alta velocidad o Waring como ejemplos), prensa francesa, o incluso centrifugación en el caso de paredes celulares débiles, para romper las células.

7. Lisis de células mediante choque osmótico (citólisis)

En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo se realiza mediante la aplicación de un choque osmótico. Se puede lograr, por ejemplo, mediante la centrifugación del caldo de fermentación y la resuspensión de la pasta celular en agua desionizada.

8. Infección con un virus lítico

[0054] En una modalidad preferida de la presente invención, la etapa de lisis de un microorganismo comprende la infección del microorganismo con un virus lítico. Una amplia variedad de virus se conocen por lisar microorganismos adecuados para el uso en la presente invención, y la selección y uso de un virus lítico en particular para un microorganismo en particular se encuentra dentro del nivel de experiencia en la técnica.

Por ejemplo, el virus *paramecium bursaria chlorella* (PBCV-1) es el prototipo de un grupo (familia Phycodnaviridae, género Chlorovirus) de los virus de doble cadena de ADN, grandes, icosaédricos, que forman placa, que se replican en, y lisan, ciertas algas verdes similares a *chlorella*, unicelulares, eucariotas. En correspondencia, cualquier microalga susceptible, tal como *C. protothecoides*, puede lisarse mediante la infección del cultivo con un virus de *chlorella* adecuado. Se conocen los métodos de infectar las especies de *Chlorella* con un virus de *chlorella*. Ver, por ejemplo, *Adv. Virus Res.* 2006; 66: 293-336; *Virology*, 1999 abr 25; 257(1):15-23; *Virology*, 2004 ene 5; 318(1): 214-23; *Nucleic Acids Symp. Ser.* 2000; (44): 161-2; *J. Virol.* 2006 mar; 80(5): 2437-44; y *Annu. Rev. Microbiol.* 1999; 53: 447-94.

9. Autolisis (expresión de un gen lítico)

En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo comprende la autolisis. En esta modalidad, un microorganismo de la invención se modifica por ingeniería genética para producir un gen lítico que lisa al microorganismo. Este gen lítico puede expresarse con un promotor inducible, de manera que las células pueden ser cultivadas primero a una densidad deseable en un fermentador y después cosechadas, seguido de la inducción del promotor para expresar el gen lítico para lisar las células. En una modalidad, el gen lítico codifica una degradante del polisacárido.

En otras ciertas modalidades, el gen lítico es un gen de un virus lítico. De este modo, por ejemplo, un gen lítico de un virus de *Chlorella* se puede expresar en una *Chlorella*, tal como *C. protothecoides*.

Los métodos de expresión adecuados se describen en la presente con respecto a la expresión de un gen lipasa. La expresión de los genes líticos se realiza de preferencia usando un promotor inducible, tal como un promotor activo en microalgas que se induce mediante un estímulo, tal como la presencia de una molécula pequeña, luz, calor, y otros estímulos. Los genes líticos de los virus de *chlorella* se conocen. Por ejemplo, ver *Virology* 260, 308-315 (1999); *FEMS Microbiology Letters* 180 (1999) 45-53; *Virology* 263, 376-387 (1999); y *Virology* 230, 361-368 (1997).

10. Oscilación de presión

En otra modalidad preferida de la presente invención, la etapa de lisis de un microorganismo comprende someter el microorganismo a incrementos y disminuciones

rápidas en la presión. Dichos incrementos y disminuciones rápidas se realizan, de preferencia, a través de un diferencial de presión lo suficientemente amplio que las células no son capaces de autorregularse, y como resultado estallan.

11. Consideración adicional

5 Cuando los lípidos no se extraen inmediatamente después del aislamiento del microorganismo, los microorganismos aislados se secan típicamente. El secado de un microorganismo se puede hacer, por ejemplo, en un secador de tambor. Los microorganismos secos se pueden envasar en recipientes sellados al vacío para evitar la degradación de los lípidos.

10 B. Tratamiento del lisado celular con solvente orgánico

Un lípido producido por un microorganismo se refiere también en la presente como "primer lípido". Un primer lípido podría comprender uno o más lípidos producidos por el microorganismo.

15 Un lípido usado para extraer un primer lípido de un microorganismo se refiere como un "segundo lípido". Un segundo lípido preferido es un aceite. Un solvente orgánico de la presente invención comprende un segundo lípido.

Algunos de los métodos de la presente invención comprenden la etapa de tratamiento de un lisado con un solvente orgánico. Típicamente, el solvente orgánico se añade directamente al lisado sin separación previa de los componentes del lisado. Después de
20 la adición del solvente, el lisado se separa en diferentes capas ya sea por su propia cuenta o como resultado de la centrifugación o similares. Las capas pueden incluir en orden de densidad decreciente: un sedimento de sólidos pesados, una fase acuosa, una fase de emulsión, y una fase oleosa. La fase de emulsión es una emulsión de lípidos y fase acuosa. En
dependencia del porcentaje de solvente orgánico añadido con respecto al lisado (p/p o v/v), si
25 acaso la fuerza de centrifugación, el volumen de medios acuosos y otros factores, se pueden presentar una o ambas fases de emulsión y aceite.

La incubación o el tratamiento del lisado celular o fase de emulsión con el solvente orgánico se realiza durante un tiempo suficiente para permitir que el lípido producido por el microorganismo se solubilice en el solvente orgánico para formar una mezcla heterogénea.

1. Aceites

En una modalidad preferida de la presente invención, un solvente orgánico es un aceite seleccionado del grupo que consiste en aceite de soja, colza, canola, palma, almendra de palma, coco, maíz, aceite de residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de buey, sebo porcino, microalgas, macroalgas, *Cuphea*, lino, cacahuete, grasa blanca selecta (manteca de cerdo), *Camelina sativa* semilla de mostaza, nuez de marañón, avena, lupino, kenaf, caléndula, cáñamo, café, semilla de lino, avellana, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez, jojoba, jatrofa, macadamia, nueces del Brasil, y aguacate. Se incluyen, además, los combustibles fósiles tales como el aceite crudo o una fracción destilada del petróleo.

La cantidad de solvente orgánico añadido al lisado es típicamente mayor que 5% (medido por v/v y/o p/p) del lisado con el que se combina el solvente. Así, un v/v o p/p preferido del solvente orgánico es mayor que 5%, al menos 6%, al menos 7%, al menos 10%, al menos 20%, al menos 25%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, y al menos 95% del lisado celular.

2. Otros solventes orgánicos

Otros ejemplos no limitantes de solventes orgánicos que se pueden usar para practicar los métodos de la presente invención incluyen hexano, isohexano, metanol, dodecano, aceite crudo derivado de fósil y fracciones de destilación de éste, y el dióxido de carbono supercrítico. El metanol puede ser un solvente orgánico ventajoso para el uso en la fabricación de biodiesel, ya que se puede usar en el proceso de transesterificación, así como la extracción de aceite.

C. Extracción sin solventes

Los lípidos se pueden extraer también de un lisado por el enfriamiento del lisado, sin uso sustancial o ninguno de solventes orgánicos. En dichos métodos, el lisado se produce de preferencia por tratamiento con ácido en combinación con la temperatura ambiente por encima. La sonicación se puede usar también, particularmente si la temperatura está entre la temperatura ambiente y 65°C. Un lisado de este tipo en la centrifugación o sedimentación se puede separar en capas, una de las cuales es una capa acuosa:lipídica. Otras capas pueden incluir un sedimento sólido, una capa acuosa y una capa lipídica. El lípido se puede extraer de la capa de emulsión por descongelación o de lo contrario enfriamiento de la emulsión, como

se describe además más abajo. En dichos métodos, no es necesario añadir ningún solvente orgánico. Si algún solvente orgánico se añade, el solvente orgánico puede ser inferior al 5% del lisado v/v o p/p.

D. Agitación o no agitación de la mezcla heterogénea

- 5 En una modalidad preferida de la presente invención, después de producir un lisado celular y, opcionalmente, añadir el solvente orgánico al lisado celular o a la fase de emulsión, como se describió anteriormente, se agita la mezcla heterogénea. La agitación se puede hacer en un agitador, vórtice, o similares. La agitación puede resultar también de las fuerzas de cizalla presentes durante la centrifugación de alta fuerza g. Como alternativa, la mezcla
- 10 heterogénea no se agita.

E. Separación de la mezcla heterogénea en una composición lípido:solvente orgánico y una composición acuosa y, opcionalmente una composición emulsionada o una composición de sedimento celular

- 15 El lisado celular, opcionalmente tratado con un solvente orgánico, producido por los métodos anteriores es una mezcla heterogénea que incluye los lípidos, soluciones acuosas, restos celulares y el solvente orgánico (si se añaden). La mezcla heterogénea se puede separar en múltiples capas como se describe más abajo. Las múltiples capas pueden incluir en orden de densidad decreciente, un sedimento de restos celulares, una capa acuosa, una capa de emulsión lipídica:acuosa y una capa lipídica. La presencia y proporciones relativas de las
- 20 diferentes capas dependen de la técnica de lisis, la técnica de separación que incluyen fuerza centrífuga, concentración lipídica en el microorganismo y si se usa un solvente orgánico. El lípido deseado puede ocurrir tanto en una capa lipídica separada y en una capa lipídica:acuosa. El uso de un solvente orgánico favorece la formación de una capa lipídica separada. Mientras mayor sea la concentración del solvente orgánico en relación a la lisis, es
- 25 más probable que el lípido se encuentre en una capa lipídica separada en vez de en una capa de emulsión acuosa:lipídica. El solvente orgánico, si está presente, se asocia con el lípido ya sea en la capa lipídica o de emulsión o ambas. El lípido se puede obtener tanto de una capa lipídica separada como de una capa de emulsión, como se describe más abajo.

- 30 El lisado celular se puede separar en diferentes capas como se describió anteriormente mediante técnicas, tales como por ejemplo, centrifugación y sedimentación, es decir, permitiendo formar las capas de manera espontánea con el tiempo.

F. Separación del lípido

La separación del lípido de otros componentes presentes en el lisado depende de las capas que se forman como resultado de la separación. Si se forma una capa lipídica separada, esta capa se forma en la parte superior de las otras capas y se puede aspirar, pipetear o decantar o similares de la parte superior del recipiente que contiene las capas separadas. Esta capa lipídica se puede usar después en diversas aplicaciones que se describen más abajo. Del mismo modo las capas sólidas y/o acuosas se pueden drenar y excluir de la parte inferior del recipiente. Si una capa de emulsión se forma, la emulsión se puede someter a la extracción adicional para separar los lípidos en la emulsión del líquido acuoso. En una circunstancia de ese tipo, las capas acuosa y/o de sedimento por separado, si se presentan, se pueden drenar antes de separar la emulsión. Una capa lipídica separada, si se presenta, se puede decantar también, sin embargo, no es necesario, y la capa lipídica separada puede facilitar la separación de la emulsión en sus componentes.

La emulsión se puede separar en sus componentes mediante el lavado con un volumen adicional de solvente orgánico para extraer los lípidos de la emulsión. Como alternativa, la emulsión se puede lavar con una solución acuosa para eliminar los fluidos acuosos de la emulsión. Como alternativa, la emulsión se puede someter al enfriamiento por debajo de la temperatura ambiente. De preferencia, la emulsión se congela y descongela. La congelación es de preferencia a una temperatura de -5°C a -30°C , y se mantiene durante al menos una hora.

Después de cualquiera de los tratamientos anteriores, la emulsión se vuelve a separar en capas de componentes mediante las mismas técnicas, como se describió previamente. Las capas de componentes pueden incluir un sedimento celular, capa acuosa, capa de emulsión y capa lipídica como se describió previamente. Las capas de componentes se pueden separar, como se describió previamente. Si una capa de emulsión está presente, la capa de emulsión se puede someter a una ronda adicional de extracción usando los mismos procedimientos descritos anteriormente.

Después de uno o más ciclos de separación, el lisado finalmente produce una capa lipídica. La capa lipídica se puede usar en diversas aplicaciones, como se describe más abajo. Cuando cualquiera de los ciclos de extracción se realiza con un solvente orgánico, los restos de solvente orgánico permanecen asociados a la capa lipídica. Típicamente, no es necesario separar la capa lipídica del solvente orgánico. Por ejemplo, cuando el solvente orgánico es un

aceite o lípido, el lípido presente en los microorganismos se puede referir como un primer lípido y el lípido o aceite usado para la extracción se puede referir como un segundo lípido. Por ejemplo, el primer lípido puede ser el triacilglicerol de una microalga y el segundo lípido puede ser aceite de soja. Después de la separación de una capa lípido:solvente orgánico, (en este caso, lípido de microalgas: aceite de soja), la mezcla se puede, por ejemplo transesterificar o hidrotratar para producir biodiesel o diesel renovable, respectivamente.

Uno de los métodos de separación que se puede usar en un lisado, incluyendo un lisado que contiene una emulsión, es la aplicación de fuerzas de cizalla. A continuación de la lisis celular, el aceite se podría presentar como una emulsión de aceite en agua con un tamaño de gota pequeña (<5 o <10 micrón). La emulsión se podría estabilizar por cualquier número de emulsionantes anfipáticos (por ejemplo oleosinas, proteínas desnaturalizadas, fosfolípidos, ácidos grasos etc.) En igualdad de circunstancias, la estabilidad de una emulsión se relaciona inversamente al tamaño de la gota de la fase incluida. Existen un número de procedimientos establecidos en la presente que se pueden aplicar para causar que las gotas de aceite se unan e incrementen de tamaño, de ese modo volviendo la emulsión más susceptible a la rotura (por ejemplo, agitación suave en la presencia de ciertos detergentes, o ácidos o bases). Una vez que una emulsión de aceite en agua se ha tratado para debilitar la emulsión (por un incremento en el tamaño de gota, como se describió anteriormente, y/o por algún tratamiento químico o físico que disminuye la eficacia del (de los) agente(s) emulsionante(s)), las gotas necesitan estar físicamente juntas a la fuerza para lo cual se pueden unir y podría ocurrir la fase de separación.

Un método de forzar las gotas de aceite incluidas juntas es mediante la centrifugación que causa la aglomeración física, al forzar la emulsión a través de una película delgada e introducir la cizalla concomitante que puede requerirse para romper la emulsión. Un dispositivo que se puede usar en los métodos proporcionados en la presente para introducir la fuerza-g, película delgada y cizalla es una centrífuga de discos apilados en la que la fase ligera se expulsa rápidamente como una película delgada entre discos giratorios. La película es de preferencia entre 1 micrómetro y 1000 micrómetros. La película es de preferencia menor o similar al diámetro de las gotas de aceite incluidas.

En una modalidad, las células microbianas portadoras de aceite en el caldo de fermentación se centrifugan para reducir el contenido de agua de la composición. Las células se rompen por homogenización y/o tratamiento químico/enzimático. Opcionalmente, las

gotas de aceite se fuerzan a unirse por una agitación suave del lisado. La fase de separación de la composición se alcanza después mediante la centrifugación de discos apilados.

Otro método de separación que se puede usar sólo o en combinación con otros métodos descritos en la presente es el uso de uno o más surfactantes para desestabilizar una emulsión. La desestabilización resulta en la unión de los glóbulos de aceite en grandes agrupamientos y eventualmente en una composición de fase separada de una capa ligera de aceite, y una fase acuosa más pesada. Los surfactantes preferidos para desestabilizar las emulsiones de la biomasa microbiana que contiene aceite son la oleamida DEA (dietanolamina) (Stepan Chemical Co.), que incluyen pero no se limitan a, NINOL 201, laurelamida DEA que incluyen pero no se limitan a, NINOL 96L y NINOL 55L y cocoamida DEA que incluyen pero no se limitan a, NINOL 40CO. Los métodos de recuperación del aceite a partir de una emulsión de lípidos usando surfactantes se proporcionan en los ejemplos más abajo.

La composición lipídica obtenida se puede analizar por un número de métodos, que incluyen HPLC como se describe en el Ejemplo 6.

VII. Método de producción de combustibles adecuados para el uso en vehículos diesel y motores de reacción

Un interés creciente se dirige al uso de componentes de hidrocarburos de origen biológico en los combustibles, tales como biodiesel, diesel renovable y combustible de reactor, ya que las materias primas biológicas renovables que pueden sustituir a las materias primas derivadas de los combustibles fósiles están disponibles, y su uso, por lo tanto, es deseable. Existe una necesidad urgente de métodos para la producción de componentes de hidrocarburos a partir de materiales biológicos. La presente invención satisface esta necesidad al proporcionar métodos para la producción de biodiesel, diesel renovable y combustible de reactor usando la composición lípido:solvente orgánico o los lípidos descritos en la presente descripción como un material biológico para producir biodiesel, diesel renovable y combustible de reactor.

Después de la extracción, la presente invención proporciona la ventaja de que el aceite microbiano y el solvente orgánico (preferiblemente un aceite vegetal), pueden juntos ser sometidos a un tratamiento químico para la fabricación de un combustible para uso en vehículos diesel y motores de reacción. La capacidad de evitar la separación del solvente orgánico usado para la extracción del aceite microbiano proporciona una ventaja significativa

sobre los métodos tradicionales de extracción de aceite, tal como la extracción de hexano en la que el hexano se debe destilar separado del aceite microbiano previo a cualquiera de las etapas adicionales de procesamiento.

[0055] Los combustibles diesel tradicionales son destilados del petróleo ricos en hidrocarburos parafínicos. Ellos tienen intervalos de ebullición tan amplios como de 370° a 780°F, los cuales son adecuados para la combustión en un motor de arranque por compresión, tal como un vehículo de motor diesel. La American Society of Testing and Materials (ASTM), establece el grado de diesel de acuerdo con el intervalo de ebullición, junto con los intervalos permisibles de otras propiedades del combustibles, tales como el número de cetano, punto de nube, punto de inflamación, viscosidad, punto de anilina, contenido de azufre, contenido de agua, contenido de cenizas, corrosión de la tira de cobre, y residuos de carbono. Técnicamente, cualquier hidrocarburo destilado del material derivado a partir de la biomasa que reúna la especificación ASTM conveniente se puede definir como diesel, o biodiesel.

El combustible diesel se puede producir de la biomasa a través de varios tipos de tecnologías. Las materias primas para los combustibles diesel derivados de la biomasa incluyen, pero no se limitan a, soja, semilla de colza, canola, palma, y aceites residuales de cocina, junto con las grasas de animales. Los aceites de partida pueden ser también de origen de algas. La capa lípido:solvente orgánico producida por el método de la presente invención puede servir como materia prima para producir biodiesel y diesel renovable.

A. Biodiesel

[0056] El biodiesel es un líquido que varía en color – entre dorado y marrón oscuro – en dependencia de la materia prima usada en la producción. Es prácticamente inmiscible en agua, tiene un alto punto de ebullición y baja presión de vapor. El biodiesel se refiere a un combustible procesado equivalente al diesel para su uso en vehículos con motor diesel. El biodiesel es biodegradable y no tóxico. Un beneficio adicional del biodiesel con respecto a los combustibles diesel convencionales, es el menor desgaste del motor.

Típicamente, el biodiesel comprende ésteres de alquilo de cadena corta. Varios procesos convierten la biomasa o un lípido producido y aislado según lo descrito en la presente, en los combustibles diesel. Un método preferido para producir biodiesel es la transesterificación de un lípido como se describe en la presente. Un éster de alquilo de cadena corta preferido para su uso como biodiesel es un éster metílico o éster etílico.

[0057] El biodiesel producido por un método descrito en la presente se puede usar solo o mezclado con combustible diesel convencional en cualquier concentración en la mayoría de los vehículos modernos de motores a diesel. Cuando se mezcla con el combustible diesel convencional (diesel), el biodiesel puede estar presente en

5 aproximadamente de 0.1% a aproximadamente 99.9%. Gran parte del mundo usa un sistema conocido como el factor "B", para indicar la cantidad de biodiesel en cualquier mezcla de combustible. Por ejemplo, el combustible que contiene 20% de biodiesel se etiqueta como B20. El biodiesel puro se refiere como B100.

[0058] El biodiesel también puede usarse como combustible de calefacción en

10 calderas domésticas y comerciales. Las calderas de aceite que existen pueden contener partes de caucho, y pueden requerir conversión para funcionar con biodiesel. El proceso de conversión es relativamente sencillo, el cual involucra el intercambio de las partes de caucho por partes sintéticas debido a que el biodiesel es un fuerte solvente. Debido a su fuerte poder solvente, el quemado del biodiesel incrementará la eficiencia de las calderas.

[0059] El biodiesel puede usarse como un aditivo en las formulaciones de diesel para incrementar la lubricidad del combustible Ultra-Low Sulfur Diesel (ULSD) puro, lo cual es ventajoso, ya que éste prácticamente no tiene contenido de azufre.

[0060] El biodiesel es un mejor solvente que el petrodiesel y se puede usar para separar los depósitos de los residuos en las líneas de combustible de los vehículos que

20 previamente han funcionado con el petrodiesel.

1. Producción de biodiesel

El biodiesel se puede producir por transesterificación de los triglicéridos contenidos en la biomasa rica en aceite y grasas animales. La capa lípido:solvente orgánico o lípidos producidos por el método de la presente invención pueden servir como materia prima para

25 producir el biodiesel. Por lo tanto, en otro aspecto de la presente invención se proporciona un método para producir biodiesel. En una modalidad preferida, el método para la producción del biodiesel comprende las etapas de (a) lisar un microorganismo que contiene lípido para producir un lisado; (b) tratar el lisado con un solvente orgánico durante un período de tiempo suficiente para permitir que el lípido del microorganismo se solubilice en el solvente

30 orgánico, donde mediante el lisado tratado con solvente orgánico se forma una mezcla heterogénea; (c) separar la mezcla heterogénea en capas que comprenden una capa lípido:solvente orgánico, y una capa acuosa, (d) eliminar la composición lípido:solvente

orgánico de la composición acuosa, composición de emulsión, o composición de sedimento celular, y (e) transesterificar la composición lípido:solvente orgánico, por la cual el biodiesel se produce. La composición lípido:solvente orgánico comprende el solvente orgánico y los lípidos microbianos.

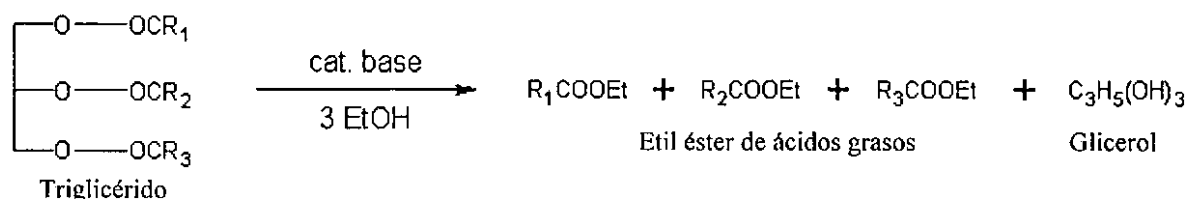
5 Los métodos para cultivar un microorganismo, lisar un microorganismo para producir un lisado, tratar el lisado en un medio que comprende un solvente orgánico para formar una mezcla heterogénea y separar el lisado tratado en una composición lípido:solvente orgánico y una composición acuosa o emulsionada se han descrito anteriormente y también se pueden usar en el método de producción de biodiesel.

10 Los lípidos y la composición lípido:solvente orgánico, donde el solvente orgánico es un triacilglicérido seleccionado como se describió anteriormente se pueden someter a la transesterificación para producir ésteres de ácidos grasos de cadena larga útiles como biodiesel. Las reacciones de transesterificación preferidas se describen más abajo e incluyen transesterificación catalizada por bases y transesterificación usando lipasas recombinantes.

15 [0061] En un proceso de transesterificación catalizada por bases, los triglicéridos se hacen reaccionar con un alcohol, tal como metanol o etanol, en presencia de un catalizador alcalino, típicamente hidróxido de potasio. Esta reacción forma ésteres metílicos o etílicos y glicerina (glicerol) como un subproducto.

a) Proceso químico general

20 [0062] Los aceites animales y vegetales se hacen típicamente de triglicéridos que son ésteres de ácidos grasos libres con el alcohol trihídrico, glicerol. En la transesterificación, el glicerol en un triacilglicérido (TAG) se sustituye por un alcohol de cadena corta tal como el metanol o el etanol. Un esquema de la reacción típica es el siguiente:



25 [0063] En este esquema, el alcohol se desprotona con una base para hacerlo un nucleófilo más fuerte. Comúnmente, el etanol o metanol se usan en gran exceso (hasta 50 veces). Normalmente, esta reacción se llevará a cabo de forma excesivamente lenta o no ocurrirá en lo absoluto. El calor, así como un ácido o una base, se pueden usar para ayudar a

la reacción a que se lleve a cabo más rápidamente. El ácido o la base no se consumen por la reacción de transesterificación, por lo que no son reactivos sino catalizadores. Casi todo el biodiesel se produce usando la técnica de catalización con base, ya que ésta sólo requiere bajas temperaturas y presiones y produce por encima del 98% de rendimiento de conversión
5 (siempre y cuando el aceite de partida sea bajo en humedad y libre de ácidos grasos).

Cualquiera de los ácidos grasos libres en el aceite base se convierten en jabón y se eliminan del proceso, o se esterifican (produciendo más biodiesel) usando un catalizador ácido.

La forma más común de transesterificación usa el metanol para producir ésteres metílicos, ya que es el alcohol más barato disponible. El etanol se usa para producir biodiesel
10 de éster etílico. Los alcoholes superiores, tales como isopropanol y butanol se pueden usar también.

Un subproducto del proceso de transesterificación es la producción de glicerol. Aproximadamente por cada tonelada de biodiesel producido, se producen 100 kg de glicerol.
15 Este glicerol se puede usar como un componente químico, y se puede usar también como una fuente de carbono para fermentar los microorganismos.

b) Usando lipasas recombinantes

[0064] La transesterificación también se ha llevado a cabo usando una enzima, tal como una lipasa en lugar de una base. La transesterificación catalizada por lipasa puede
20 realizarse, por ejemplo, a una temperatura entre la temperatura ambiente y 80°C, y una proporción molar de TAG al alcohol inferior mayor que 1:1, preferentemente 2:1, con mayor preferencia 3:1, y con la máxima preferencia aproximadamente 3:1.

Las lipasas adecuadas para usar en la transesterificación incluyen, pero sin limitarse a, las enumeradas en la Tabla 11 más abajo. Otros ejemplos de lipasas útiles para la
25 transesterificación se encuentran en, por ejemplo las patentes de Estados Unidos núms. 4,798,793; 4,940,845 5,156,963; 5,342,768; 5,776,741 y WO89/01032.

[0065] Tabla 11. Lipasas adecuadas para su uso en la transesterificación.

La lipasa de *Aspergillus niger* ABG73614, lipasa B de *Candida antarctica* (novozym-435) CAA83122, lipase de *Candida cylindracea* AAR24090, lipasa de *Candida lipolytica* (Lipasa L; Amano Pharmaceutical Co., Ltd.), lipasa de *Candida rugosa* (por ejemplo, Lipasa-OF;

Meito Sangyo Co., Ltd.), lipasa de *Mucor miehei* (Lipozyme IM 20), lipasa de *Pseudomonas fluorescens* AAA25882, lipasa de *Rhizopus japonicas* (Lilipase A-10FG) Q7M4U7_1, lipasa de *Rhizomucor miehei* B34959, lipasa de *Rhizopus oryzae* (Lipasa F) AAF32408, lipasa de *Serratia marcescens* (Enzima SM) ABI13521, lipasa de *Thermomyces lanuginosa* CAB58509, Lipasa P (Nagase ChemteX Corporation), y Lipasa QLM (Meito Sangyo Co., Ltd., Nagoya, Japón)

Un reto para usar una lipasa para la producción de ésteres de ácidos grasos adecuados para el biodiesel es que el precio de la lipasa es mucho mayor que el precio del hidróxido de sodio (NaOH) usado por los procesos de base fuerte. Este reto se abordó usando una lipasa inmovilizada, que se puede reciclar. Sin embargo, la actividad de la lipasa inmovilizada debe mantenerse después de reciclarse por un número mínimo de ciclos para permitir un proceso basado en lipasa que compita con el proceso de base fuerte en términos de costo de producción. Las lipasas inmovilizadas se someten a envenenamiento por los alcoholes inferiores típicamente usados en la transesterificación. La patente de los Estados Unidos núm. 6,398,707 (concedida el 04 de junio 2002 a Wu y otros) describe los métodos para incrementar la actividad de las lipasas inmovilizadas y la regeneración de las lipasas inmovilizadas con actividad reducida.

2. Normas

[0066] La norma internacional común para el biodiesel es la EN 14214. ASTM D6751, la norma de biodiesel más común que se referencia en los Estados Unidos y Canadá. Alemania usa la norma DIN EN 14214 y el Reino Unido exige el cumplimiento de la norma BS EN 14214.

Las pruebas industriales básicas para determinar si los productos se ajustan a estas normas típicamente incluyen cromatografía de gases, HPLC, y otras. El biodiesel que cumple las normas de calidad no es muy tóxico, con un índice de toxicidad (DL₅₀) mayor de 50 ml/kg.

B. Diesel Renewable

El diesel renovable comprende una mezcla de alcanos, tales como C10:0, C12:0, C14:0, C16:0 y C18:0 y así, se distinguen del biodiesel. El diesel renovable de alta calidad se obtiene conforme a la norma ASTM D 975.

La capa lípido:solvente orgánico o lípidos producidos por los métodos de la presente invención pueden servir como materia prima para producir el diesel renovable. Los lípidos producidos por los métodos de la presente invención pueden servir como materia prima para producir el diesel renovable. El diesel renovable se puede producir mediante al menos tres procesos procesamiento hidrotérmico (hidrotratamiento) hidroprocesamiento y licuefacción indirecta. Estos procesos producen destilados no ésteres. Durante estos procesos, los triglicéridos producidos y aislados como se describe en la presente, se descomponen a alquenos de C16 y C18.

Así, en otro aspecto de la presente invención se proporciona un método para producir diesel renovable. En una modalidad preferida, el método para la producción de diesel renovable comprende las etapas de un método de producción de diesel renovable que comprende las etapas de: (a) lisar un microorganismo que contiene lípido para producir un lisado; (b) tratar el lisado con un solvente orgánico durante un período de tiempo suficiente para permitir que el lípido del microorganismo se solubilice en el solvente orgánico para formar una mezcla heterogénea; (c) separar la mezcla heterogénea en una composición lípido:solvente orgánico, y una composición acuosa, y opcionalmente, una composición emulsionada o composición de sedimento celular; (d) eliminar la composición lípido:solvente orgánico de la composición acuosa, composición de emulsión, o composición de sedimento celular, y (e) tratar la composición lipídica:orgánica para producir un alcano de cadena lineal, por el cual el biodiesel se produce. La composición lípido:solvente orgánico comprende el solvente orgánico y los lípidos celulares.

1. Hidrotratamiento

En una modalidad preferida del método para la producción de diesel renovable, el tratamiento de la composición lípido:solvente orgánico o los lípidos producidos y aislados para producir un alcano de cadena lineal, se realiza mediante el hidrotratamiento de la composición lípido:solvente orgánico. En el procesamiento hidrotérmico, típicamente la biomasa se hace reaccionar en agua a una temperatura y presión elevadas para formar aceites y residuos sólidos. Las temperaturas de conversión son típicamente de 570° a 660°F, con presión suficiente para mantener el agua primariamente como un líquido, de 100 a 170 atmósferas como estándar. Los tiempos de reacción son del orden de 15 a 30 minutos. Después que la reacción se completa, los orgánicos se separan del agua. De tal modo se produce un destilado apropiado para el diesel.

2. Hidroprocesamiento

Un diesel renovable, denominado "diesel verde" se puede producir a partir de ácidos grasos por la tecnología de hidrotratamiento tradicional. Los aceites que contienen triglicéridos se pueden hidroprocesar ya sea como co-alimento con petróleo o como un alimento específico. El producto es un combustible diesel Premium que no contiene azufre y que tiene un número de cetano de 90-100. De este modo, en otra modalidad preferida del método para la producción de diesel renovable, el tratamiento de la composición lípido:solvente orgánico o los lípidos producidos y aislados para producir un alcano de cadena lineal, se realiza mediante el hidrotratamiento de la composición lípido:solvente orgánico.

[0067] Las refinerías de petróleo usan el hidroprocesamiento para eliminar las impurezas mediante el tratamiento de las fuentes con hidrógeno. Las temperaturas de conversión del hidroprocesamiento son típicamente de 600° a 700° F. Las presiones son normalmente de 40 a 100 atmósferas. Los tiempos de reacción son del orden de 10 a 60 minutos.

[0068] Los catalizadores sólidos se emplean para aumentar ciertas velocidades de reacción, mejorar la selectividad de determinados productos, y optimizar el consumo de hidrógeno.

[0069] El hidrotratamiento y el hidroprocesamiento en última instancia conducen a una reducción en el peso molecular de la fuente. En el caso de los aceites que contienen triglecéridos, la molécula de triglicéridos se reduce a cuatro moléculas de hidrocarburos bajo las condiciones de hidroprocesamiento: una molécula de propano y tres moléculas de hidrocarburo, típicamente en el intervalo de C12 a C18. En algunos métodos, la primera etapa del tratamiento de un triglicérido es el hidroprocesamiento para saturar los dobles enlaces, seguido por desoxigenación a temperatura elevada en presencia de hidrógeno y un catalizador. En algunos métodos, la hidrogenación y desoxigenación ocurren en la misma reacción. En otros métodos la desoxigenación ocurre antes de la hidrogenación. La isomerización se realiza entonces opcionalmente, también en presencia de hidrógeno y un catalizador. Finalmente, los gases y los componentes de nafta se pueden eliminar si se desea. Por ejemplo, ver las patentes de los Estados Unidos núms. 55,475,160 (hidrogenación de triglicéridos); 5,091,116 (desoxigenación, hidrogenación y eliminación de gas); 6,391,815 (hidrogenación); y 5,888,947 (isomerización).

3. Licuefacción indirecta

[0070] Un diesel tradicional ultra bajo en azufre se puede producir a partir de cualquier forma de biomasa mediante un proceso de dos pasos. Primero, la biomasa se convierte en un gas sintético, una mezcla gaseosa rica en hidrógeno y monóxido de carbono.

5 Entonces el gas sintético se convierte catalíticamente en líquido. Típicamente, la producción de líquidos se logra usando la síntesis de Fischer-Tropsch (FT). Esta tecnología se aplica al carbón, gas natural y aceites pesados. De este modo, aún en otra modalidad preferida del método para la producción de diesel renovable, el tratamiento de la composición
10 lípido:solvente orgánico o los lípidos producidos y aislados para producir un alcano de cadena lineal, se realiza mediante la licuefacción indirecta de la composición lípido:solvente orgánico.

C. Combustible de reactor

El uso anual de combustible de reactor en los Estados Unidos en el 2006 fue aproximadamente de 21 mil millones de galones (aproximadamente 80 millones de litros). El
15 combustible del avión es de transparente a color pajizo. El combustible más común es un combustible a base de aceite sin plomo/parafina clasificado como Aeroplane A-1, el cual se produce para un conjunto especificaciones normalizadas a nivel internacional. El combustible de avión es una mezcla de un gran número de diferentes hidrocarburos, posiblemente tanto como mil o más. El intervalo de su tamaño (peso molecular o números de carbono) está
20 limitado por los requerimientos para el producto, por ejemplo, el punto de congelación o el punto de humo. El combustible de avión tipo Kerosene (incluyendo el Jet A y el Jet A-1) tiene una distribución del número de carbonos entre aproximadamente 8 y 16 números de carbono. El combustible de avión de tipo nafta o wide-cut (incluyendo el Jet B) típicamente tiene una distribución del número de carbonos entre aproximadamente 5 y 15 números de
25 carbonos.

Ambos aviones (Jet A y jet B) pueden contener un número de aditivos. Los aditivos útiles incluyen, pero sin limitarse a, antioxidantes, agentes antiestáticos, inhibidores de la corrosión, y agentes inhibidores de la congelación del sistema de combustible (FSII). Los
30 antioxidantes previenen el engomado y usualmente se basan en fenoles alquilados, por ejemplo, AO-30, AO-31, o AO-37. Los agentes antiestáticos disipan la electricidad estática y evitan las chispas. El Stadis 450 con ácido dinonilnaftilsulfónico (DINNSA) como ingrediente activo, es un ejemplo. Los inhibidores de la corrosión, por ejemplo, el DNI-4A,

se usa para combustibles civiles y militares y el ICD-6A se usa para combustibles militares. Los agentes FSII incluyen, por ejemplo, el Di-EGME.

Una solución es mezclar los combustibles de algas con el combustible de reactor existente. La presente invención proporciona una solución. La capa lípido:solvente orgánico o lípidos producidos por el método de la presente invención pueden servir como materia prima para producir el combustible de reactor. Por lo tanto, en otro aspecto de la presente invención, se proporciona un método para producir combustible de reactor. Por consiguiente, se proporcionan dos métodos para la producción del combustible de reactor a partir de la capa lípido:solvente orgánico o lípidos producidos por el método de la presente invención, el craqueo catalítico fluido (FCC) y la hidrodesoxigenación (HDO).

1. Craqueo catalítico fluido

[0071] El craqueo catalítico fluido (FCC) es un método que se usa para producir olefinas, especialmente propileno a partir de fracciones pesadas del crudo. Existen reportes en la literatura de que los aceites vegetales tal como el aceite de canola se pueden procesar usando el FCC para dar una corriente de hidrocarburos útiles como un combustible de gasolina.

La capa lípido:solvente orgánico o lípidos producidos por el método de la presente invención se puede convertir en olefinas C_2-C_5 . El proceso involucra el flujo de la capa orgánica lípido:solvente o los lípidos producidos a través de una zona de FCC y se recoge una corriente del producto compuesta de olefinas, que es útil como un combustible de reactor. La capa lípido:solvente orgánico o lípidos producidos se ponen en contacto con un catalizador de craqueo en condiciones de craqueo para proporcionar una corriente de producto que comprende las olefinas C_2-C_5 e hidrocarburos útiles como combustible de reactor.

Por lo tanto, en aún otro aspecto de la presente invención, se proporciona un método para producir combustible de reactor. En una modalidad preferida, el método para la producción de combustible de reactor comprende las etapas de (a) lisar un microorganismo que contiene lípido para producir un lisado; (b) tratar el lisado con un solvente orgánico durante un período de tiempo suficiente para permitir que el lípido del microorganismo se solubilice en el solvente orgánico para formar una mezcla heterogénea; (c) y una composición acuosa, y opcionalmente, una composición emulsionada o composición de sedimento celular; (d) eliminar la composición lípido:solvente orgánico de la composición acuosa, composición de emulsión, o composición de sedimento celular; y (e) tratar la

composición lípido:solvente orgánico para producir una olefina C₂-C₅, a través de la cual se produce el biodiesel. La composición lípido:solvente orgánico comprende el solvente orgánico y los lípidos celulares.

En una modalidad preferida del método para la producción de un combustible de reactor, la etapa (f) se realiza mediante la fluidez de la composición lípido:solvente orgánico a través de una zona de craqueo catalítico fluido. La etapa (f) podría comprender, además contactar la composición lípido:solvente orgánico con un catalizador de craqueo en condiciones de craqueo para proporcionar una corriente de producto que comprende las olefinas C₂-C₅.

[0072] En ciertas modalidades de este método, puede ser deseable eliminar los contaminantes que pueden estar presentes en la composición lípido:solvente orgánico. De este modo, antes de la etapa (f) fluir la composición lípido:solvente orgánico a través de una zona de craqueo catalítico fluido, la composición lípido:solvente orgánico se trata previamente. El tratamiento previo podría involucrar el contacto de la composición lípido:solvente orgánico con una resina de intercambio iónico. La resina de intercambio iónico es una resina de intercambio iónico ácida, tal como AmberlystTM-15 y se puede usar como lecho en un reactor a través del cual fluye la composición lípido:solvente orgánico, ya sea hacia arriba o hacia abajo. Otros tratamientos previos pueden incluir lavados con ácidos leves por contacto de la composición lípido:solvente orgánico con un ácido, tal como el ácido sulfúrico, acético, nítrico o clorhídrico. El contacto se realiza con una solución diluida del ácido, por lo general a temperatura ambiente y a presión atmosférica.

La composición lípido:solvente orgánico, opcionalmente sometida a un tratamiento previo, se hace fluir a una zona de FCC, donde los componentes hidrocarburos se craquean a olefinas. El craqueo catalítico se lleva a cabo poniendo en contacto la composición lípido:solvente orgánico en una zona de reacción con un catalizador compuesto de un material particulado finamente dividido. La reacción es de craqueo catalítico, a diferencia del hidrocrqueo, y se lleva a cabo en ausencia de hidrógeno añadido o consumo de hidrógeno. A medida que avanza la reacción de craqueo, cantidades sustanciales de coque se depositan en el catalizador. El catalizador se regenera a altas temperaturas por la quema del coque del catalizador en una zona de regeneración. El catalizador que contiene coque, que se refiere en la presente como "catalizador coquizado", se transporta continuamente desde la zona de reacción a la zona de la regeneración para regenerarse y reemplazarse por un catalizador

regenerado esencialmente libre de coque a partir de la zona de regeneración. La fluidización de las partículas del catalizador mediante varias corrientes gaseosas permite el transporte del catalizador entre la zona de reacción y la zona de regeneración. Los métodos para el craqueo de los hidrocarburos, tal como los de la composición lípido:solvente orgánico descritos en la presente, en una corriente fluidizada del catalizador, el transporte del catalizador entre las zonas de reacción y de regeneración, y la combustión de coque en el regenerador, son bien conocidos por los expertos en la técnica de los procesos de FCC. Las aplicaciones del FCC ilustrativos y los catalizadores útiles para craquear la composición lípido:solvente orgánico para producir olefinas de C_2 - C_5 son descritos en las patentes de los Estados Unidos núms. 6,538,169, 7,288,685, las cuales se incorporan en su totalidad como referencia.

En una modalidad, el craqueo de la composición lípido:solvente orgánico de la presente invención, se lleva a cabo en la sección del elevador o, alternativamente, la sección del elevador, de la zona de FCC. La composición lípido:solvente orgánico se introduce en el elevador por una tobera lo que resulta en la vaporización rápida de la composición lípido:solvente orgánico. Antes de contactar el catalizador, la composición lípido:solvente orgánico normalmente tendrá una temperatura de aproximadamente 149°C a aproximadamente 316°C (300°F a 600°F). El catalizador se hace fluir desde un recipiente de mezclado hasta el elevador, donde entra en contacto con la composición lípido:solvente orgánico durante un tiempo de aproximadamente 2 segundos o menos.

El catalizador mezclado y los vapores de la composición lípido:solvente orgánico reaccionada se descargan desde la parte superior del elevador a través de una salida y se separan en una corriente de vapor del producto craqueado que incluye las olefinas y una colección de partículas del catalizador cubiertas con cantidades sustanciales de coque que se refiere generalmente como "catalizador coquizado". En un esfuerzo por minimizar el tiempo de contacto de la composición lípido:solvente orgánico y el catalizador, que puede promover una conversión adicional de los productos deseados a otros productos no deseados, se puede usar cualquier disposición de separadores, tal como una disposición de brazo en espiral, para eliminar el catalizador coquizado de la corriente del producto rápidamente. El separador, por ejemplo, el separador de brazo en espiral, se sitúa en una porción superior de una cámara con una zona de extracción situada en la porción inferior de la cámara. El catalizador separado por el arreglo del brazo en espiral, gotea en la zona de extracción. La corriente de vapor del producto craqueado que comprende los hidrocarburos craqueados que incluyen olefinas ligeras y algún catalizador, salen de la cámara a través de un conducto que está en

comunicación con los ciclones. Los ciclones eliminan los restos de partículas del catalizador de la corriente de vapor del producto para reducir las concentraciones de partículas a niveles muy bajos. La corriente de vapor del producto entonces sale de la parte superior del recipiente de separación. El catalizador, separado por los ciclones, se devuelve al recipiente de separación y luego a la zona de extracción. La zona de extracción elimina los hidrocarburos adsorbidos de la superficie del catalizador por el contacto a contracorriente con el vapor.

[0073] La baja presión parcial de los hidrocarburos opera a favor de la producción de olefinas ligeras. En consecuencia, la presión del elevador se establece a aproximadamente de 172 a 241 kPa (25 a 35 psia) con una presión parcial de hidrocarburos de aproximadamente 35 a 172 kPa (5 a 25 psia), con una presión parcial de hidrocarburos preferida de aproximadamente 69 a 138 kPa (10 a 20 psia). Esta presión parcial, relativamente baja para los hidrocarburos, se logra mediante el uso de vapor como un diluyente en la medida en que el diluyente es 10 a 55 % en peso de la composición lípido:solvente orgánico y preferentemente aproximadamente 15 % en peso de la composición lípido:solvente orgánico. Otros diluyentes, tales como gas seco, se pueden usar para alcanzar presiones parciales de hidrocarburos equivalentes.

[0074] La temperatura de la corriente de craqueo en la salida del elevador será de aproximadamente 510°C a 621°C (950°F a 1150°F). Sin embargo, las temperaturas de salida del elevador por encima de 566°C (1050°F) producen más gas seco y más olefinas mientras que las temperaturas de salida del elevador por debajo de 566°C (1050°F) producen menos etileno y propileno. Por consiguiente, se prefiere para llevar a cabo el proceso de FCC a una temperatura preferida de aproximadamente 566°C a aproximadamente 630°C, presión preferida de aproximadamente de 138 kPa a aproximadamente 240 kPa (20 a 35 psia). Otra condición para el proceso es la relación del catalizador con respecto a la composición lípido:solvente orgánico la cual puede variar de aproximadamente 5 a aproximadamente 20 y preferentemente de aproximadamente 10 a aproximadamente 15.

En una modalidad del método para producir un combustible de reactor, la composición lípido:solvente orgánico se introduce en la sección del elevador de un reactor de FCC. La temperatura en la sección del elevador estará muy caliente y se encuentra en el intervalo de aproximadamente 700°C (1292°F) a aproximadamente 760°C (1400°F) con una relación del catalizador con respecto a la composición lípido:solvente orgánico de

aproximadamente 100 a aproximadamente 150. Se prevé que la introducción de la composición lípido:solvente orgánico en la sección del elevador producirá cantidades considerables de propileno y etileno.

Los productos hidrocarburos líquidos y gaseosos producidos pueden analizarse por cromatografía de gases, HPLC, etc.

2. Hidrodesoxigenación

En otra modalidad del método para producir un combustible de reactor usando la composición lípido:solvente orgánico o los lípidos producidos como se describe en la presente, la estructura de la composición lípido:solvente orgánico o de los lípidos se rompe por un proceso conocido como hidrodesoxigenación (HDO). Como tal, la etapa (f) se realiza por hidrodesoxigenación de la composición lípido:solvente orgánico.

[0075] La HDO significa la eliminación de oxígeno por medio de hidrógeno, es decir, el oxígeno se elimina mientras que se rompe la estructura del material. Los dobles enlaces olefínicos se hidrogenan y todos los compuestos de azufre y nitrógeno se eliminan. La eliminación de azufre se llama hidrodesulfuración (HDS). El tratamiento previo y la pureza de las materias primas (composición lípido:solvente orgánico o los lípidos) contribuyen a la vida útil del catalizador.

Generalmente en la etapa de HDO/HDS, el hidrógeno se mezcla con la materia prima de alimentación (composición lípido:solvente orgánico o los lípidos) y después la mezcla se pasa a través de un lecho del catalizador, como un flujo de co-corriente, ya sea como una materia prima de alimentación en una sola fase, una fase doble. Después de la HDO/etapa de HDS, la fracción del producto se separa y se pasa a un reactor de isomerización separado. Un reactor de isomerización para material de partida biológico se describe en la literatura (FI 100 248) como un reactor de co-corriente.

El proceso para producir un combustible por hidrogenación de una alimentación de hidrocarburo, por ejemplo, la composición lípido:solvente orgánico o los lípidos en la presente descripción, también se puede realizar pasando la composición lípido:solvente orgánico o los lípidos como un flujo de co-corriente con gas de hidrógeno a través de una primera zona de hidrogenación, y posteriormente el efluente del hidrocarburo se hidrogena adicionalmente en una segunda zona de hidrogenación, pasando gas de hidrógeno a la segunda zona de hidrogenación como un flujo a contracorriente con respecto al efluente del

hidrocarburo. Las aplicaciones ilustrativas de la HDO y de los catalizadores útiles para craquear la composición lípido:solvente orgánico para producir olefinas de C_2-C_5 se describen en la patente de los Estados Unidos núm. 7,232,935 la cual se incorpora en su totalidad como referencia.

5 Típicamente, en la etapa de hidrodeshidrogenación, la estructura del componente biológico, tal como la composición lípido:solvente orgánico o los lípidos de la presente descripción, se descompone, el oxígeno, el nitrógeno, los compuestos de fósforo y de azufre y los hidrocarburos ligeros como gas se eliminan, y los enlaces olefínicos se hidrogenan. En la
10 isomerización para la ramificación de la cadena del hidrocarburo y se mejora el rendimiento de la parafina a bajas temperaturas.

En la primera etapa, es decir, en la etapa de HDO del proceso de craqueo, el gas de hidrógeno y la composición lípido:solvente orgánico o los lípidos de la presente descripción que deben hidrogenarse se pasan a un sistema de lecho del catalizador de HDO ya sea como
15 flujos de co-corriente o contracorriente, dicho sistema de lecho del catalizador comprende uno o más lechos de catalizadores, preferentemente 1-3 lechos de catalizador. La etapa de HDO típicamente funciona de una manera en co-corriente. En el caso de un sistema de lecho del catalizador de HDO, el cual comprende dos o más lechos de catalizadores, uno o más de los lechos pueden funcionar usando el principio de flujo a contracorriente.

20 [0076] En la etapa de HDO, la presión varía entre 20 y 150 bar, preferentemente entre 50 y 100 bar, y la temperatura varía entre 200 y 500°C, preferentemente en el intervalo de 300-400°C.

[0077] En la etapa de HDO, se pueden utilizar los catalizadores de hidrogenación conocidos que contienen metales del grupo VII y/o VIB del sistema periódico.
25 Preferentemente, los catalizadores de hidrogenación son catalizadores soportados de Pd, Pt, Ni, NiMo o un catalizador CoMo, siendo el soporte alúmina y/o sílice. Típicamente se usan los catalizadores NiMo/ Al_2O_3 y CoMo/ Al_2O_3 .

Antes de la etapa de HDO, la composición lípido:solvente orgánico los lípidos de la presente descripción pueden tratarse, opcionalmente, por prehidrogenación en condiciones
30 más suaves, evitándose así reacciones secundarias de los dobles enlaces. Dicha prehidrogenación se lleva a cabo en presencia de un catalizador de prehidrogenación a temperaturas de 50 a 400°C y a presiones de hidrógeno de 1 a 200 bar, preferentemente a una

temperatura entre 150 y 250°C y a una presión de hidrógeno entre 10 y 100 bar. El catalizador puede contener metales del grupo VIII y/o VIB del sistema periódico. Preferentemente, el catalizador de prehidrogenación es un catalizador soportado de Pd, Pt, Ni, NiMo o un catalizador CoMo, el soporte es alúmina y/o sílice.

5 [0078] Una corriente gaseosa a partir de la etapa de HDO que contiene hidrógeno se enfría y luego los compuestos de monóxido de carbono, dióxido de carbono, nitrógeno, fósforo y azufre, hidrocarburos ligeros gaseosos y otras impurezas, se eliminan de la misma. Después de la compresión, el hidrógeno purificado o el hidrógeno reciclado se retornan de vuelta al primer lecho del catalizador y/o entre los lechos de catalizadores para
10 compensar la corriente del gas eliminado. El agua se elimina a partir del líquido condensado. El líquido se pasa al primer lecho del catalizador o entre los lechos de los catalizadores.

[0079] Después de la etapa de HDO, el producto se somete a una etapa de isomerización. Es importante para el proceso que las impurezas se eliminen tan completamente como sea posible antes de que los hidrocarburos se pongan en contacto con el
15 catalizador de isomerización. La etapa de isomerización comprende una etapa opcional de separación, en la cual el producto de reacción de la etapa de HDO se puede purificar por separación con vapor de agua o un gas adecuado, tal como hidrocarburos ligeros, nitrógeno o hidrógeno. La etapa opcional de separación se lleva a cabo en forma de contracorriente en una unidad ubicada corriente arriba del catalizador de isomerización, en la cual el gas y el
20 líquido están en contacto uno con el otro, o antes del reactor de isomerización existente en una unidad de separación independiente que utiliza el principio de contracorriente.

Después de la etapa de separación, el gas de hidrógeno y la composición lípido:solvente orgánico hidrogenada o los lípidos de la presente descripción y opcionalmente una mezcla de n-parafina, se pasan a una unidad de isomerización reactiva que
25 comprende uno o varios lechos de los catalizadores. Los lechos de los catalizadores de la etapa de isomerización pueden operar en forma de co-corriente o contracorriente.

[0080] Es importante para el proceso que el principio de flujo a contracorriente se aplique en la etapa de isomerización. En la etapa de isomerización, esto se hace llevando a cabo cualquier etapa de separación opcional o la etapa de la reacción de
30 isomerización o ambas en forma de contracorriente.

[0081] La etapa de isomerización y la etapa HDO se pueden llevar a cabo en el mismo recipiente de presión o en recipientes de presión separados. La prehidrogenación

opcional se puede llevar a cabo en un recipiente de presión separado o en el mismo recipiente de presión que las etapas de HDO y de isomerización.

[0082] En la etapa isomerización, la presión varía en el intervalo de 20-150 bar, de preferencia en el intervalo de 20-100 bar, la temperatura es entre 200 y 500°C, preferentemente entre 300 y 400°C.

[0083] En la etapa de isomerización, los catalizadores de isomerización conocidos en la materia se pueden usar. Los catalizadores de isomerización adecuados contienen el tamiz molecular y/o un metal del grupo VII y/o un vehículo. Preferentemente, el catalizador de isomerización contiene SAPO-11 o SAPO41 o ZSM-22 o ZSM-23 o ferrierita y Pt, Pd o Ni y Al₂O₃ y SiO₂. Los catalizadores de isomerización típicos son, por ejemplo, Pt/SAPO-11/Al₂O₃, Pt/ZSM-22/Al₂O₃, Pt/ZSM-23/Al₂O₃ y Pt/SAPO-11/SiO₂.

Como producto, se obtiene un componente hidrocarburo de alta calidad de origen biológico, útil como un combustible diesel o un componente de éste, la densidad, el número de cetano y el rendimiento a baja temperatura de dicho componente hidrocarburo es excelente.

VIII. COMPOSICIONES

Otro propósito de la presente invención es proporcionar composiciones que comprenden lípidos aislados mediante el uso de los métodos descritos en la presente. Una composición preferida comprende (i) un primer lípido aislado a partir de un microorganismo, y (ii) un segundo lípido, donde el segundo lípido se obtiene a partir de una fuente que no sea el microorganismo.

El primer lípido se puede aislar de cualquiera de los microorganismos descritos en la presente.

A. Primeros Lípidos

Los métodos de la presente invención son aplicables a la extracción de una variedad de lípidos a partir de una variedad de microorganismos. Los microorganismos descritos en la presente producen una variedad de lípidos, tales como fosfolípidos, ácidos grasos libres, ésteres de ácidos grasos, incluyendo triglicéridos de ácidos grasos, esteroides, pigmentos (por ejemplo, carotenoides y oxicarotenoides) y otros lípidos, compuestos asociados a lípidos, tales como fitoesteroides, ergotionina, ácido lipoico y antioxidantes que incluyen betacaroteno y tocoferol. Los primeros lípidos ejemplares incluyen, pero no se limitan a, triglicéridos C8,

C10, C12, C14, C16 y C18, lípidos que contienen ácidos grasos omega-3 altamente insaturados tales como el ácido docosahexanoico (DHA), ácido eicosapentanoico (EPA), y/o ácido docosapentanoico (DPA). Los primeros lípidos incluyen también el ácido araquidónico, ácido estearidónico, colesterol, desmesterol, astaxantina, cantaxantina, y ácidos grasos altamente insaturados n-6 y n-3 tales como el ácido eicosapentanoico, ácido docosapentanoico y el ácido docosahexanoico. Otros lípidos y microorganismos que podrían ser adecuados para el uso en la presente invención serán fácilmente evidentes para aquellas personas expertos en la técnica.

Los primeros lípidos preferidos son lípidos que contienen una cantidad relativamente grande de ácidos grasos C18 y C16.

:1 C18:1

En una modalidad preferida de la presente invención, el primer lípido de una composición comprende al menos 50% de un lípido C18:1, de preferencia al menos 60%, y con mayor preferencia al menos 80%.

2. C10, C12 y C14

En otra modalidad preferida de la presente invención, el primer lípido de una composición comprende al menos 10% de un lípido C10:0, C12:0 y C14:0 combinados, de preferencia al menos 15%, con mayor preferencia al menos 20%, y con máxima preferencia al menos 30%.

B. Segundos lípidos

El segundo lípido puede ser cualquier aceite, opcionalmente seleccionado del grupo que consiste en aceite de soja, colza, canola, palma, coco, maíz, residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de buey, sebo porcino, microalgas, macroalgas, *Cuphea*, lino, cacahuate, grasa blanca selecta, manteca de cerdo, *Camelina sativa*, semilla de mostaza, nuez de marañón, avena, lupino, kenaf, caléndula, cáñamo, café, semilla de lino, avellanas, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez pacana, jojoba, jatrofa, nueces de macadamia, nueces del Brasil, y aguacate. Un segundo lípido preferido es un aceite de coco. Otro segundo lípido preferido es un aceite de palma o aceite de soja. Un segundo lípido puede ser, además, el combustible fósil tales como el aceite crudo o una fracción destilada del petróleo.

C. Composiciones que comprenden un primer lípido y un segundo lípido

El primer lípido y el segundo lípido se pueden proporcionar en diferentes proporciones. En una modalidad preferida, la proporción del primer lípido al segundo lípido está entre 1 y 100. En otra modalidad preferida, la proporción del primer lípido al segundo lípido está entre 1 y 10. En otra modalidad preferida, la proporción del primer lípido al segundo lípido es de aproximadamente 1. En otra modalidad preferida, la proporción del segundo lípido al primer lípido está entre 10 y 1.

Las modalidades adicionales de la presente invención incluyen los componentes funcionales opcionales que podrían permitir a un experto en la técnica realizar cualquiera de las variaciones del método descrito en la presente.

A pesar de que la invención anterior se describió en algún detalle a modo de ilustración y ejemplo para claridad y comprensión, será evidente para un experto en la técnica a la luz de las enseñanzas de esta invención que ciertas variaciones, cambios, modificaciones y sustitución de equivalentes pueden realizarse sin necesidad de apartarse del espíritu y alcance de esta invención. Como resultado, las modalidades descritas en la presente están sujetas a diversas modificaciones, cambios, y similares, siendo el alcance de esta invención determinado solamente con referencia a las reivindicaciones adjuntas a ésta. Aquellos expertos en la técnica reconocerán fácilmente una variedad de parámetros poco críticos que se pueden cambiar, alterar o modificar para producir resultados esencialmente similares

Aunque cada uno de los elementos de la presente invención se describen en la presente descripción conteniendo múltiples modalidades, debe entenderse que, a menos que se indique de cualquier otra forma, cada una de las modalidades de un determinado elemento de la presente invención es susceptible a ser usado con cada una de las modalidades de los otros elementos de la presente invención y se pretende que cada uso de ese tipo forme una modalidad diferente de la presente invención.

Las patentes, solicitudes de patentes y literatura científica referenciadas incluyendo los números de acceso a las secuencias de la base de datos del GenBank, referidos en la presente descripción se incorporan como referencia en su totalidad como si cada publicación individual, patente o solicitud de patente se indicaran específica e individualmente para ser incorporadas como referencia. Cualquier discrepancia entre las referencias citadas en la presente descripción y las enseñanzas específicas de esta especificación se deberá resolver a favor de esta última. Del mismo modo, cualquier discrepancia entre una definición de una

palabra o una frase comprendida en la técnica y una definición de la palabra o frase como se enseña específicamente en esta especificación se deberá resolver a favor de esta última. Las publicaciones mencionadas en la presente descripción se citan con el fin de describir y divulgar los reactivos, metodologías y conceptos que se pueden usar en relación con la presente invención. Nada en la presente se debe interpretar como una admisión de que estas referencias son materia anterior en relación con las invenciones descritas en la presente. En particular, las solicitudes de patente siguientes se incorporan como referencia en su totalidad a todos los efectos: la solicitud provisional de Estados Unidos núm. 61/028,493, presentada el 13 de febrero de, 2008, titulada “Extracción de lípidos a partir de microorganismos”; solicitud provisional de Estados Unidos núm. 61/036,918, presentada el 14 de marzo de 2008, titulada “Métodos de separación de aceite”; solicitud provisional de Estados Unidos núm. 61/043,318, presentada el 8 de abril de 2008, titulada “Fraccionamiento de biomasa que contiene aceite”; solicitud de patente PCT núm. PCT/US2009/066142, presentada el 30 de noviembre de 2009, titulada “Producción de aceites adaptados en heterótrofo microorganismos”; y solicitud de patente PCT núm. PCT/US2009/066141, presentada el 30 de noviembre de 2009, titulada “Fabricación de aceites adaptados en microorganismos heterótrofos recombinantes”.

[0084] A pesar de que esta invención se describió en relación con modalidades específicas, se entenderá que es capaz de modificaciones adicionales. La presente invención tiene además una amplia variedad de aplicaciones. Esta solicitud pretende cubrir cualquier variación, los usos o adaptaciones de la invención siguiendo, en general, los principios de la invención y se incluyen tales desviaciones de la presente descripción que vienen dentro de la práctica conocida o habitual en la materia a la que pertenece la invención y como se puede aplicar a las características esenciales anteriormente expuestas.

IX. EJEMPLOS

Ejemplo 1: Cultivo de microalgas para lograr alto contenido de aceite

[0085] Las cepas de microalgas se cultivaron para lograr un alto porcentaje de aceite por peso seco de células. Las células criopreservadas se descongelaron a temperatura ambiente y 500 µl de células se añadieron a 4.5 ml de medio (4.2 g/l de K₂HPO₄, 3.1 g/l de NaH₂PO₄, 0.2 g/l de MgSO₄ • 7H₂O, 0.25 g/l de ácido cítrico monohidrato, 0.025 g/l de CaCl₂ 2H₂O, 2 g/l de extracto de levadura) más 2% de glucosa y crecieron durante 7 días a 28°C con agitación (200 rpm) en una placa de 6 pocillos. El peso seco de las células se

determinó por centrifugación de 1 ml de cultivo a 14000 rpm durante 5 min en un tubo Eppendorf pre-pesado. El sobrenadante de cultivo se desechó y el sedimento celular resultante se lavó con 1 ml de agua desionizada. El cultivo se centrifugó de nuevo, se desechó el sobrenadante, y el sedimento de células se colocó a -80°C hasta que se congeló. Las muestras se liofilizaron después por 24 horas y se calcularon los pesos secos de las células. Para la determinación de lípidos totales en los cultivos, se retiraron 3 ml de cultivo y se sometieron a análisis mediante el uso de un sistema de Ankom (Ankom Inc., Macedon, NY), según el protocolo del fabricante. Las muestras se sometieron a extracción por solventes con un extractor Ankom XT10 según el protocolo del fabricante. Los lípidos totales se determinaron como la diferencia de masa entre las muestras secas hidrolizadas con ácido y las muestras secas extraídas con disolvente. Las mediciones del porcentaje de aceite del peso seco de las células se muestran en la Tabla 12.

Tabla 12. Porcentaje de aceite por peso seco celular para microalgas.

Especie	Cepa	% Aceite	Cepa #
<i>Chlorella kessleri</i>	UTEX 387	39.42	4
<i>Chlorella kessleri</i>	UTEX 2229	54.07	5
<i>Chlorella kessleri</i>	UTEX 398	41.67	6
<i>Parachlorella kessleri</i>	SAG 11.80	37.78	7
<i>Parachlorella kessleri</i>	SAG 14.82	50.70	8
<i>Parachlorella kessleri</i>	SAG 21.11 H9	37.92	9
<i>Prototheca stagnora</i>	UTEX 327	13.14	10
<i>Prototheca moriformis</i>	UTEX 1441	18.02	11
<i>Prototheca moriformis</i>	UTEX 1435	27.17	12
<i>Chlorella minutissima</i>	UTEX 2341	31.39	13
<i>Chlorella protothecoides</i>	UTEX 250	34.24	1
<i>Chlorella protothecoides</i>	UTEX 25	40.00	2
<i>Chlorella protothecoides</i>	CCAP 211/8D	47.56	3
<i>Chlorella</i> sp.	UTEX 2068	45.32	14
<i>Chlorella</i> sp.	CCAP 211/92	46.51	15
<i>Chlorella sorokiniana</i>	SAG 211.40B	46.67	16
<i>Parachlorella beijerinckii</i>	SAG 2046	30.98	17
<i>Chlorella luteoviridis</i>	SAG 2203	37.88	18
<i>Chlorella vulgaris</i>	CCAP 211/11K	35.85	19
<i>Chlorella reisi</i>	CCAP 11/8	31.17	20
<i>Chlorella ellipsoidea</i>	CCAP 211/42	32.93	21
<i>Chlorella saccharophila</i>	CCAP 211/31	34.84	22
<i>Chlorella saccharophila</i>	CCAP 211/32	30.51	23

Se realizaron tres procesos de fermentación con tres formulaciones diferentes del medio con el objetivo de generar biomasa de algas con alto contenido de aceite. La primera formulación (Medio 1) se basó en el medio descrito en Wu y otros. (1994 *Science in China*, vol. 37, núm. 3, pp. 326-335) y consistió por litro de: KH_2PO_4 , 0.7g; K_2HPO_4 , 0.3g; $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$, 0.3g; $\text{FeSO}_4\cdot 7\text{H}_2\text{O}$, 3mg; hidrocloreto de tiamina, 10 μg ; glucosa, 20g; glicina, 0.1g; H_3BO_3 , 2.9mg; $\text{MnCl}_2\cdot 4\text{H}_2\text{O}$, 1.8mg; $\text{ZnSO}_4\cdot 7\text{H}_2\text{O}$, 220 μg ; $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$, 80 μg ; y $\text{NaMoO}_4\cdot 2\text{H}_2\text{O}$, 22.9mg. El segundo medio (Medio 2) se derivó a partir de el medio del frasco descrito en el Ejemplo 1 y consistió por litro de: K_2HPO_4 , 4.2g; NaH_2PO_4 , 3.1g; $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$, 0.24g; ácido cítrico monohidrato, 0.25g; cloruro cálcico anhidro, 25mg; glucosa, 20g; extracto de levadura, 2g. El tercer medio (Medio 3) fue un híbrido y consistió por litro de: K_2HPO_4 , 4.2g; NaH_2PO_4 , 3.1g; $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$, 0.24g; ácido cítrico monohidrato, 0.25g; cloruro cálcico anhidro, 25mg; glucosa, 20g; extracto de levadura, 2g; H_3BO_3 , 2.9mg; $\text{MnCl}_2\cdot 4\text{H}_2\text{O}$, 1.8 mg; $\text{ZnSO}_4\cdot 7\text{H}_2\text{O}$, 220 μg ; $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$, 80 μg ; y $\text{NaMoO}_4\cdot 2\text{H}_2\text{O}$, 22.9mg. Las formulaciones de los tres medios se prepararon y esterilizaron en autoclave en recipientes fermentadores a escala de laboratorio por 30 minutos a 121°C. Se añadió glucosa estéril a cada recipiente siguiendo con la esterilización en autoclave con enfriamiento posterior.

El inóculo para cada fermentador fue *Chlorella protothecoides* (UTEX 250), preparada en dos etapas del frasco usando las condiciones de medio y temperatura del fermentador inoculado. Cada fermentador se inoculó con 10% (v/v) del cultivo semi-logarítmico. Los tres fermentadores a escala de laboratorio se mantuvieron a 28°C por la duración del experimento. El crecimiento celular de microalgas en el Medio 1 se evaluó también a una temperatura de 23°C. Para todas las evaluaciones de fermentación, el pH se mantuvo en 6.6-6.8, agitaciones a 500 rpm, y flujo de aire en 1 vvm. Los cultivos de fermentación se cultivaron durante 11 días. La acumulación de biomasa se midió mediante la densidad óptica a 750 nm y el peso seco celular.

La concentración de lípido/aceite se determinó mediante la transesterificación directa con los métodos estándares de cromatografía de gases. Brevemente, las muestras de caldo de fermentación con biomasa se secaron en un papel secante y transfirieron a tubos de centrifuga y secaron en un horno de vacío a 65-70°C durante 1 hora. Cuando las muestras se secaron, 2 ml de H_2SO_4 al 5% en metanol se añadió a los tubos. Los tubos se calentaron después en un bloque térmico a 65-70°C durante 3.5 horas, mientras se agitaron y sonicaron de forma intermitente. 2ml de heptano se añadieron después y los tubos se agitaron vigorosamente. 2Ml de K_2CO_3 al 6% se añadió y los tubos se agitaron vigorosamente para

mezclar y centrifugar después a 800 rpm durante 2 minutos . El sobrenadante se transfirió después a viales de GC que contienen el agente de secado Na₂SO₄ y corrió usando los métodos de cromatografía de gases. El por ciento de aceite/lípido se basó en una base de peso seco celular. El peso seco celular por células cultivadas usando: Medio 1 a 23°C fue 9.4g/l; Medio 1 a 28°C fue 1.0g/l, Medio 2 a 28°C fue 21.2g/l; y Medio 3 a 28°C fue 21.5g/l. La concentración lípido/aceite por células cultivadas usando: Medio 1 a 23°C fue 3g/l; Medio 1 a 28°C fue 0.4g/l; Medio 2 a 28°C fue 18 g/l; y Medio 3 a 28°C fue 19g/l. El porcentaje de aceite basado en el peso seco celular por células cultivadas usando: Medio 1 a 23°C fue 32%; Medio 1 a 28°C fue 40%; Medio 2 a 28°C fue 85%; y Medio 3 a 28°C fue 88%. Los perfiles de lípido (en % del área, después de normalizar al estándar interno) para la biomasa de algas generadas usando las formulaciones de los tres medios diferente a 28°C se resumen más abajo en la Tabla 13.

Tabla 13. Perfiles de lípido para *Chlorella protothecoides* cultivadas bajo diferentes condiciones del medio.

	Medio 1 28°C (en % del área)	Medio 2 28°C (en % del área)	Medio 3 28°C (en % del área)
C14:0	1.40	0.85	0.72
C16:0	8.71	7.75	7.43
C16:1	--	0.18	0.17
C17:0	--	0.16	0.15
C17:1	--	0.15	0.15
C18:0	3.77	3.66	4.25
C18:1	73.39	72.72	73.83
C18:2	11.23	12.82	11.41
C18:3 alfa	1.50	0.90	1.02
C20:0	--	0.33	0.37
C20:1	--	0.10	0.39
C20:1	--	0.25	--
C22:0	--	0.13	0.11

Ejemplo 3: Genotipado de microalgas con alto contenido de aceite

Las muestras de microalgas de las 23 cepas enumeradas en la Tabla 12 anteriormente se genotiparon. El ADN genómico se aisló a partir de la biomasa de algas de la siguiente manera. Las células (aproximadamente 200 mg) se centrifugaron a partir de cultivos líquidos por 5 minutos a 14,000 x g. Las células se resuspendieron después en agua destilada estéril, se centrifugaron por 5 minutos a 14,000 xg y se desechó el sobrenadante. Una sola perla de vidrio de ~ 2 mm de diámetro se añadió a la biomasa y los tubos se colocaron a -80 °C durante al menos 15 minutos. Las muestras se retiraron y se añadieron 150 µl de tampón de molturación (Sarkosil 1%, sacarosa 0.25 M, NaCl 50 mM, EDTA 20

- mM, Tris-HCl 100 mM, pH 8.0, ARNasa A 0.5 µg/µl). Los sedimentos se resuspendieron con una breve agitación, seguido por la adición de 40 µl de NaCl 5M. Las muestras se agitaron brevemente, seguido por la adición de 66 µl de CTAB 5% (bromuro cetil trimetilamonio) y una agitación breve final. Las muestras se incubaron después a 65°C
- 5 durante 10 minutos después de lo cual se centrifugaron a 14,000 x g durante 10 minutos. El sobrenadante se transfirió a un tubo nuevo y se extrajo una vez con 300 µl de fenol:cloroformo:alcohol isoamílico a 12:12:1, seguido por la centrifugación durante 5 minutos a 14,000 x g. La fase acuosa resultante se transfirió a un tubo nuevo que contenía 0.7 vol de (~ 190 µl), se mezcló por inversión y se incubó a temperatura ambiente durante 30
- 10 minutos o toda la noche a 4 °C. El ADN se recuperó a través de la centrifugación a 14,000 x g durante 10 minutos. El precipitado resultante se lavó dos veces con 70% etanol, seguido de un lavado final con 100% etanol. Los sedimentos se secaron al aire durante 20-30 minutos a temperatura ambiente, seguido de la resuspensión en 50 µl de TrisCl 10 mM, EDTA 1 mM (pH 8.0).
- 15 [0086] Cinco µl del ADN total de algas, que se preparó como se describió anteriormente, se diluyeron 1:50 en Tris 10 mM, pH 8.0. Las reacciones de PCR, de volumen final de 20 µl, se establecieron de la siguiente manera. Diez µl de la mezcla maestra iProof HF 2 x (BIO-RAD) se añadieron a 0.4 µl del cebador SZ02613 (5'-TGTTGAAGAATGAGCCGCGAC-3' (sec. con núm. de ident.: 5) a una concentración
- 20 patrón de 10 mM). Esta secuencia del cebador corre de la posición 567-588 en el núm. de acceso L43357 del Gen Bank y está altamente conservada en los genomas de plantas superiores y plastidios de algas. Esto se siguió con la adición de 0.4 µl del cebador SZ02615 (5'-CAGTGAGCTATTACGCACTC-3' (sec. con núm. de ident.: 6) a la concentración patrón de 10 mM). Esta secuencia del cebador es complementaria a la posición de 1112-1093
- 25 en el núm. de acceso L43357 del Gen Bank y es muy conservada en los genomas de plantas superiores y plastidios de algas. A continuación, se añadieron 5 µl del ADN total diluido y 3.2 µl dH₂O. Las reacciones de PCR se llevaron a cabo de la siguiente manera: 98°C, 45", 98°C, 8", 53 °C, 12", 72°C, 20" por 35 ciclos seguidos de 72°C durante 1 min y se mantuvo a 25°C. Para la purificación de los productos de la PCR, se añadieron a cada reacción 20 µl de
- 30 Tris 10 mM, pH 8.0, seguido por la extracción con 40 µl de fenol:cloroformo:alcohol isoamílico 12:12: 1, la agitación y centrifugación a 14,000 x g durante 5 minutos. Las reacciones de PCR se aplicaron a columnas S-400 (GE Healthcare) y se centrifugaron durante 2 minutos a 3,000 x g. Los productos de la PCR purificados posteriormente se clonaron en

TOPO en el PCR8/GW/TOPO y los clones positivos se seleccionaron en placas LB /Spec. El ADN del plásmido purificado se secuenció en las dos direcciones usando cebadores M13 directo e inverso. Las secuencias de las cépas 1-23 (designadas en el Ejemplo 1, Tabla 12) se enumeran como sec. con núms. de ident.: 7-29 en el listado de secuencias anexo.

Ejemplo 4: ADN genómico del ARNr 23S de *Chlorella protothecoides* y *Prototheca*

El ADN genómico de 8 cepas de *Chlorella protothecoides* (UTEX 25, UTEX 249, UTEX 250, UTEX 256, UTEX 264, UTEX 411, CCAP 211/17 y CCAP 211/8d) se aisló y realizó el análisis de ADN genómico del ARNr 23S de acuerdo con los métodos descritos anteriormente en el Ejemplo 3. Todas las cepas de *Chlorella protothecoides* probadas fueron idénticas en secuencia a excepción de UTEX 25. Las secuencias de las ocho cepas se listan como sec. con núm. de ident.: 30 y 31 en el listado de secuencias adjunto.

La secuencia genómica del ARNr 23s para *Prototheca moriformis* UTEX 1436 (sec. con núm. de ident.: 32) se comparó también con otras especies de *Prototheca* (UTEX 1435, UTEX 1437, y UTEX 1439) y las cepas de *Chlorella protothecoides* descritas anteriormente. La comparación mostró que la secuencia genómica del ARNr 23S para *Prototheca moriformis* UTEX 1436 fue diferente a los otros genotipos de *Prototheca* (sec. con núm. de ident.: 33).

Ejemplo 5: Diversidad de cadenas de lípidos en especies de microalgas

Las muestras de lípido de un subconjunto de cepas cultivadas en el Ejemplo 1, Tabla 12 se analizaron para el perfil de lípidos usando HPLC. Los resultados se muestran más abajo en la Tabla 14.

[0087] Tabla 14. Diversidad de cadenas de lípidos en especies de microalgas

Cepa # de la Tabla 12	C:14:0	C:16:0	C:16:1	C:18:0	C:18:1	C:18:2	C:18:3	C:20:0	C:20:1
1	0.57	10.30	0	3.77	70.52	14.24	1.45	0.27	0
2	0.61	8.70	0.30	2.42	71.98	14.21	1.15	0.20	0.24
4	0.68	9.82	0	2.83	65.78	12.94	1.46	0	0
5	1.47	21.96	0	4.35	22.64	9.58	5.2	3.88	3.3
10	0	12.01	0	0	50.33	17.14	0	0	0
11	1.41	29.44	0.70	3.05	57.72	12.37	0.97	0.33	0
12	1.09	25.77	0	2.75	54.01	11.90	2.44	0	0

Ejemplo 6: Extracción de lípidos de microalgas usando aceite de coco

1. Producción celular

Un lote de tanque-F de *Chlorella protothecoides* (aproximadamente 1,200 galones) se usó para generar biomasa durante los procesos de extracción. El lote (# ZA07126) se dejó correr durante 100 horas, controlando a la vez los niveles de glucosa a 16 g/l, después de cuyo tiempo se terminó la alimentación de jarabe de maíz. Los niveles residuales de glucosa descendieron a <0 g/l dos horas después. Esto resultó en una edad final de 102 horas. El volumen final del caldo fue de 1,120 galones. Tanto los controles de contaminación en el proceso como un análisis exhaustivo de una muestra final de caldo no pudieron mostrar cualquiera de los signos de contaminación.

2. Ruptura celular

Las células liofilizadas de *Chlorella protothecoides* se resuspendieron a 200 g/l con agua DI, que contiene 20 g/l de KOH. Esta suspensión celular se autoclaveó a 130°C durante 30 minutos y se enfrió a temperatura ambiente. Una muestra de 10 ml de este material se sonicó con un sonicador Misonix 3000 equipado con una micropunta en el nivel 7. La suspensión se sonicó durante 6 minutos en un ciclo de 30 segundos de encendido/apagado para un total de 3 ciclos. Más del 70% de la ruptura se observó por análisis microscópico. Una alícuota de esta suspensión (1 ml de muestra) se centrifugó a 14 K rpm durante 15 min. Tres capas se observaron: un sedimento de sólidos pesados, una fase acuosa y una fase de emulsión. No se observó la capa de aceite.

Se realizó un segundo conjunto de experimentos en una variedad de tiempos de sonicación buscando la eficacia de la ruptura celular para lograr la máxima ruptura celular. Las muestras de 6 ml de células de *Chlorella protothecoides* preparadas como se describió anteriormente se sonicaron usando un sonicador Misonix 3000 equipado con una micropunta en el nivel 8 con intervalos de tiempo de 20 segundos de encendido y 20 segundos de apagado. Los tiempos de sonicación se extendieron desde 10 minutos, 20 minutos y 30 minutos y todos se realizaron con las muestras en hielo. Las muestras fueron centrifugadas después a 4300 rpm durante 30 minutos. Tres capas se observaron: un sedimento de sólidos pesados, una fase acuosa y una fase de emulsión. No se observó la capa de aceite. Cada intervalo de tiempo de muestreo se analizó también cualitativamente para la ruptura celular mediante la búsqueda de células enteras en el lisado bajo el aumento de 100x. No se observaron células enteras en ninguno de los intervalos de tiempo de muestreo, indicando que

10 minutos de sonicación fue suficiente para lograr la máxima ruptura celular. Las micrografías del lisado bajo el aumento mostraron gotas de aceite y restos celulares consistentes de organelos y membranas celulares rotas. En la inspección adicional de los sedimentos celulares de cada una de las muestras se demostró que no se observaron células enteras, sólo restos celulares.

3. Purificación del aceite de algas para el control

El aceite de *Chlorella protothecoides* se obtuvo a partir de biomasa seca a través de la extracción con hexano usando métodos estándares conocidos en la técnica (ver, por ejemplo, las patentes de Estados Unidos 5,567,732; 6,255,505; y Yamada y otros, Industrial applications of single cell oils, Eds. Kyle y Ratledge, 118-138 (1992)).

4. Análisis HPLC de los perfiles de ácidos grasos

Las concentraciones de ácidos grasos y los perfiles de (i) células de *Chlorella*, (ii) aceite de algas puras, (iii) aceite de coco puro, y (iv) muestras de aceite extraído se determinaron por HPLC.

Una muestra de 50 µl de la suspensión celular de *Chlorella* lavada (200 g/l) (antes del tratamiento cáustico y de calor; ver anteriormente) se hidrolizó por la incubación a 80°C en una solución saturada de isopropanol/KOH durante 4 horas. Las células se centrifugaron después a 14 K rpm durante 10 minutos y el hidrolizado se cargó en una columna de HPLC. Las muestras se analizaron mediante un HPLC Agilent 1100, usando el siguiente método. Las muestras se derivatizaron con bromuro de bromofenacilo (60 mg/ml) y se cargaron en una columna Luna 5u C8 (2) 100A 150x2 mm (Phenomenex). Las muestras se eluyeron de la columna usando un gradiente de agua a 100% de acetonitrilo:tetrahidrofurano (95:5). Las señales se detectaron usando un detector de matriz DAD a una longitud de onda de 254 nm. El resultado de dicho análisis detallando los perfiles de lípidos para los aceites puros (aceite de algas y coco), establecidos por separado (es decir, sin mezclarlos entre sí), se muestra en la Figura 1. Este resultado demuestra que los perfiles lipídicos del aceite puro de algas y aceite de coco son diferentes. Tal que, el aceite de coco se puede usar como un solvente orgánico para extraer los lípidos a partir de microalgas y el por ciento de lípidos extraídos se puede determinar (ver más abajo).

5. Construcción y validación de las curvas teóricas

Los ácidos grasos de las mezclas de aceite puro de algas extraídas de *Chlorella* y el aceite de coco se midieron como porcentajes del total de lípido y se muestran en la Tabla 15.

Tabla 15. Contenido de ácido graso como porcentaje del lípido total.

% Aceite de coco	% Aceite de algas	C18:1	C12	C14	C18:1/C12	C18:1/C14
0	100	59	0.01	1.00	5.900	59.00
2	98	57.9	1.09	8.14	53.13	7.12
5	95	56.3	2.71	8.39	20.76	6.71
10	90	53.5	5.41	8.81	9.89	6.08
20	80	48.0	10.81	9.64	4.44	4.98
25	75	45.3	13.51	10.06	3.35	4.5
40	60	37.0	21.61	11.31	1.71	3.27
50	50	31.5	27.01	12.15	1.17	2.59
60	40	26.0	32.40	12.98	0.80	2.00
27	25	17.8	40.5	14.23	0.44	1.25
80	20	15.0	43.2	14.65	0.35	1.02
89.75	10.25	9.6	48.47	15.46	0.20	0.62
94.75	5.25	6.9	51.17	15.88	0.13	0.43
100	0	4	54	17.00	0.07	0.24

En el aceite puro de algas, el 59% del total de ácidos grasos son del tipo C18:1, el 0.01% son del tipo C12, y el 1% son del tipo C14. El 39.99% de ácidos grasos de algas restantes son C10:1, C10, C18:3, C18:2, C16 o ácidos grasos C18, la mayoría siendo ácidos grasos C18:2 (ver Figura 1). En el aceite puro de coco, el 4% de los ácidos grasos son del tipo C18:1, el 54% son del tipo C12, y el 17% son del tipo C14. El 25% restante son C10:1, C10, C18:3, C18:2, C16 o ácidos grasos C18 (ver Figura 1). Al mezclar el aceite de algas con aceite de coco, cambia la proporción de los respectivos ácidos grasos. Por ejemplo, en una mezcla 1:1 de aceite puro de algas y aceite puro de coco, el 31.5% del total de ácidos grasos son del tipo C18:1 (donde la mayoría es a partir del aceite de algas), 27.01% son del tipo C12 (que son casi exclusivamente a partir del aceite de coco), y 12.15% son del tipo C14 (donde la mayoría es a partir del aceite de coco). Esta tabla de referencia es útil para determinar el por ciento de ácidos grasos de algas y el por ciento de ácidos grasos de coco en el análisis por HPLC de una composición lípido:solvente orgánico obtenida a partir de la capa más ligera de un lisado de biomasa de microalgas que se ha sometido a centrifugación.

Una curva se construyó basada en el total esperado de C18:1 vs C12 o C14 en una mezcla de los dos aceites puros, es decir, aceite de algas y aceite de coco. Las medidas teóricas para ambas curvas se validaron mediante el análisis por HPLC de una mezcla 1:1 de aceite puro de algas para aceite puro de coco que debería resultar en un 50% de cada

proporción. El valor medido del C18:1 para C12 y C14 fue del 51% y 65% respectivamente cuando se ajusta a la curva correspondiente (Figura 2).

6. Separación del aceite microbiano a partir de una emulsión usando un solvente orgánico (Ejemplo del aceite de coco)

5 Con el fin de liberar el aceite del lisado de *Chlorella* el aceite de coco se añadió a un 1 ml de material celular de *Chlorella* desestabilizado en proporciones de 3:1, 1:1, 1:3 y 1:10 (aceite de algas para aceite de coco), se mezcló bien y calentó durante 15 min a 70°C y se centrifugó después durante 15 min a 14 K rpm en una microcentrífuga. Tras la centrifugación se observaron cuatro capas: un sedimento de sólidos pesados, una fase acuosa, una fase de emulsión, una fase de aceite. Después, la fase de aceite se separó de las otras capas con pipeta. La fase de aceite incluyó una mezcla de aceite de algas y aceite de coco de los que una alícuota se sometieron a análisis por HPLC. La medición de aceite de algas en la suspensión se basó en el contenido de aceite de las células. Los controles incluyeron (i) 3 mg de aceite puro de algas, (ii) 3 mg de aceite puro de coco, y (iii) una mezcla 1:1 de aceite puro de algas y aceite puro de coco. Una muestra de 5 µl de cada uno se analizó mediante HPLC.

7. Resultados

El contenido de aceite de la masa celular de partida (CM) fue aproximadamente 30% determinado por HPLC. El aceite de coco se usó para extraer los lípidos a partir de una suspensión celular de *Chlorella* desestabilizada (200 g/l fue de suspensión) en mezclas de 3:1, 1:1, 1:3 y 1:10 (aceite de algas para coco). Las muestras de este experimento incluyeron: (1) 3 mg de aceite puro de algas (control); (2) 3 mg de aceite puro de coco (control); (3) mezcla 1:1 de 3 mg de cada aceite puro (control); (4) CM (1 ml) + aceite de coco (3:1); (5) CM (1 ml) + aceite de coco (1:1); (6) CM (1 ml) + aceite de coco (1:3); (7) CM (1 ml) + aceite de coco (1:10); y (8) CM (1 ml).

25 Ninguna capa de aceite se presentó en la muestra sin aceite de coco (8). La composición de ácidos grasos (expresado como el % de ácido grasos en el total lipídico) de los controles de aceite puro y cada una de las mezclas de extracción (experimentos (1) (7)) se muestra en la Tabla 16.

30 Tabla 16. Composición de ácido graso de las mezclas de extracción y controles de aceite puro.

% Ácido graso del total	Accite de algas (1)	Accite de coco (2)	1:1 Mezcla de accite de algas y accite	C.p./C.aceite(3:1) (4)	C.p./C.aceite (1:1) (5)	C.p./C.aceite(1:3) (6)	C.p./C.aceite (1:10) (7)
-------------------------	---------------------	--------------------	--	-------------------------	-------------------------	-------------------------	--------------------------

			de coco (3)				
C10:1	0%	1%	0%	0%	0%	0%	0%
C10	0%	10%	5%	9%	10%	9%	10%
C12	0%	56%	27%	50%	53%	53%	54%
C18:3	4%	0%	1%	0%	0%	0%	0%
C14	1%	18%	9%	18%	17%	18%	18%
C18:2	25%	1%	11%	1%	1%	1%	1%
C16	11%	8%	9%	8%	8%	8%	8%
C18:1	59%	4%	33%	11%	8%	8%	6%
C18:0	0%	2%	4%	3%	3%	3%	2%

C.p., *Chlorella protothecoides*; C.aceite, aceite de coco

Las proporciones de C18:1 para C12 y C14 se calcularon y ajustaron en el gráfico teórico (ver anteriormente y en la Figura 2). La eficacia de extracción del aceite de coco se determinó mediante la división del por ciento de aceite de algas calculado (de la curva teórica) por el por ciento esperado (contenido de aceite en el material de partida). La eficacia de extracción se extendió de 44%. en la mezcla 1:10 (aceite de algas para aceite de coco) (ver experimento (7) anterior) a un 16% en la mezcla de 3:1 (ver experimento (4) anterior) como se determinó a partir de los datos de C18 1 a C12 (Tabla 17). De este modo, cuanto mayor sea el por ciento de aceite de coco en el procedimiento de extracción, mayor es el rendimiento de la extracción de los lípidos a partir de *Chorella*.

Tabla 17. Eficiencia de extracción como una función de la fracción de aceite de coco.

% aceite de algas en la extracción	Medido C18:1 a C12	% Medido	% extracción (C18:1/C12)
9 (7)	0.12	4.0	44%
25 (6)	0.14	6.0	24%
50 (5)	0.16	7.0	14%
75 (4)	0.22	12.0	16%

El intervalo de la eficiencia de extracción fue del 33% (ver experimento (7) anterior) a un 14% (ver experimento (4) anterior) como se determinó a partir de los datos C18:1 a C14 (Tabla 18). De este modo, cuanto mayor sea el por ciento de aceite de coco en el procedimiento de extracción, mayor es el rendimiento de la extracción de los lípidos a partir de *Chorella*.

Tabla 18. Eficiencia de extracción como una función de la fracción de aceite de coco.

% aceite de algas en la extracción	Medido C18:1 a C14	% Medido	% extracción (C18:1/C14)
9 (7)	0.35	3.0	33%
25 (6)	0.43	5.3	21%
50 (5)	0.47	6.3	13%
75 (4)	0.62	10.3	14%

Los experimentos anteriores demostraron que un solvente orgánico (en este caso el aceite de coco) extrajo de manera eficiente los lípidos de una masa celular de algas para producir una composición lípido:solvente orgánico.

5 Un conjunto de experimentos adicionales se realizaron para observar la eficacia de la extracción de lípidos de una suspensión celular desestabilizada de *Chlorella protothecoides* (sonicación) usando cuatro surfactantes para desestabilizar la emulsión. Los surfactantes que se usaron en este estudio fueron: oleamida DEA (NINOL 201), laurelamida DEA (NINOL 55LL), laurelamida DEA (NINOL 96L) y cocoamida DEA (NINOL 40CO). Cada surfactante se añadió en las concentraciones de 0.2%, 0.5% y 1.0% de las suspensiones de lisado celular y se colocaron en una estufa a 55°C durante 15 horas. Cada tubo se centrifugó después a 4200 rpm durante 20 minutos. En cada muestra se pudo observar cuatro fases: una fase sólida, una fase acuosa, y la capa de emulsión y una capa de aceite en la parte superior. Las muestras de control negativo sin añadir surfactante no produjeron fase de aceite visible, pero produjeron sólo las capas sólida, acuosa y de emulsión. Las capas de aceite de cada una de las muestras tratadas con surfactante se pipetearon cuidadosamente y pesaron para determinar los rendimientos globales de aceite. Los resultados se resumen más abajo en la Tabla 19. Estos resultados muestran que un solvente orgánico, tal como un surfactante, extrae eficientemente los lípidos a partir de la biomasa de algas que contiene lípido para producir una composición lípido: solvente orgánico.

Tabla 19. Uso de surfactantes para extraer lípidos a partir de biomasa de algas.

Nombre del surfactante	NINOL 201	NINOL 201	NINOL 201	NINOL 55LL	NINOL 55LL	NINOL 55LL	NINOL 40CO	NINOL 40CO	NINOL 40CO	NINOL 96SL	NINOL 96SL	NINOL 96SL
Surfactante	Oleamida DEA	Oleamida DEA	Oleamida DEA	Laurelamida DEA	Laurelamida DEA	Laurelamida DEA	Cocoamida DEA	Cocoamida DEA	Cocoamida DEA	Laurelamida DEA	Laurelamida DEA	Laurelamida DEA
Porcentaje v/v	1.0%	0.5%	0.2%	1.0%	0.5%	0.2%	1.0%	0.5%	0.2%	1.0%	0.5%	0.2%
Aceite en ferm g/l	62	62	62	62	62	62	62	62	62	62	62	62
(g) Aceite en 5 ml de muestra (teórico)	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
Temperatura (C)	55	55	55	55	55	55	55	55	55	55	55	55
Tiempo (h)	15	15	15	15	15	15	15	15	15	15	15	15
TOTAL												
Aceite en la muestra (g)	0.183	0.237	0.041	0.262	0.012	0.026	0.276	0.082	0.000	0.286	0.035	0.000

% del teórico	59.07%	76.55%	13.11%	84.44%	4.03%	8.40%	89.01%	26.45%	0.00%	92.30%	11.41%	0.00%
---------------	--------	--------	--------	--------	-------	-------	--------	--------	-------	--------	--------	-------

Ejemplo 7: Lisis de *Chlorella Protothecoides*

1. Producción de biomasa

Un tanque-F de *Chlorella protothecoides* (aproximadamente 1,200 galones) se usó para generar biomasa durante los procesos de extracción. El lote (# ZA07126) se dejó correr durante 100 horas, controlando a la vez los niveles de glucosa a 16 g/l, después de cuyo tiempo se terminó la alimentación de jarabe de maíz. Los niveles residuales de glucosa descendieron a <0 g/l dos horas después. Esto resultó en una edad final de 102 horas. El volumen final del caldo fue de 1,120 galones. Tanto los controles de contaminación en el proceso como un análisis exhaustivo de una muestra final de caldo no pudieron mostrar cualquiera de los signos de contaminación.

2. Tratamiento químico y térmico

Las células se resuspendieron en agua a una concentración de biomasa de 150 g/l y alicuotaron en tubos de 27 x 5 ml. Cada tubo se condicionó por la matriz más abajo en la Tabla 20.

Tabla 20. Tratamiento químico/térmico de la matriz.

Tubo	Condición	[KOH] mN	[H ₂ SO ₄] mN	Temp, 30 min
1	Control	0	0	25 °C
2	40 mN KOH	40	0	25 °C
3	80 mN KOH	80	0	25 °C
4	120 mN KOH	120	0	25 °C
5	160 mN KOH	160	0	25 °C
6	40 mN H ₂ SO ₄	0	40	25 °C
7	80 mN H ₂ SO ₄	0	80	25 °C
8	120 mN H ₂ SO ₄	0	120	25 °C
9	160 mN H ₂ SO ₄	0	160	25 °C
10	Control	0	0	65 °C
11	40 mN KOH	40	0	65 °C
12	80 mN KOH	80	0	65 °C
13	120 mN KOH	120	0	65 °C
14	160 mN KOH	160	0	65 °C
15	40 mN H ₂ SO ₄	0	40	65 °C
16	80 mN H ₂ SO ₄	0	80	65 °C
17	120 mN H ₂ SO ₄	0	120	65 °C
18	160 mN H ₂ SO ₄	0	160	65 °C
19	Control	0	0	130 °C
20	40 mN KOH	40	0	130 °C
21	80 mN KOH	80	0	130 °C

22	120 mN KOH	120	0	130 °C
23	160 mN KOH	160	0	130 °C
24	40 mN H ₂ SO ₄	0	40	130 °C
25	80 mN H ₂ SO ₄	0	80	130 °C
26	120 mN H ₂ SO ₄	0	120	130 °C
27	160 mN H ₂ SO ₄	0	160	130 °C

Después del tratamiento, una muestra de 1.5 ml se centrifugó a 14 K rpm durante 15 minutos, y se midió el tamaño de cada una de las capas resultantes.

3. Lavado con agua de la emulsión

5 Un nuevo conjunto de muestras se generaron bajo condiciones para las muestras de 21-27 usando un total de 10 ml de suspensión celular, y alicuotaron en muestras de 3 x 3 ml. Las emulsiones resultantes se aislaron en tubos nuevos de 15 ml, y se sometieron al posterior esquema de lavado con agua de acuerdo con la Tabla 21.

Tabla 21. Matriz de lavado con el agua de la emulsión.

Lavado 1		Lavado 2	
Muestra	Proporción volumen de lavado con agua:emulsión	Proporción volumen de lavado con agua:emulsión	
21	1 a 1	1 a 1	
22	1 a 1	1 a 1	
23	1 a 1	1 a 1	
24	1 a 1	1 a 1	
25	1 a 1	1 a 1	
26	1 a 1	1 a 1	
27	1 a 1	1 a 1	
Muestra	Proporción volumen de lavado con agua:emulsión	Proporción volumen de lavado con agua:emulsión	
21	5 a 1	5 a 1	
22	5 a 1	5 a 1	
23	5 a 1	5 a 1	
24	5 a 1	5 a 1	
25	5 a 1	5 a 1	
26	5 a 1	5 a 1	
27	5 a 1	5 a 1	
Muestra	Proporción volumen de lavado con agua:emulsión	Proporción volumen de lavado con agua:emulsión	
21	10 a 1	10 a 1	
22	10 a 1	10 a 1	
23	10 a 1	10 a 1	
24	10 a 1	10 a 1	
25	10 a 1	10 a 1	

26	10 a 1	10 a 1
27	10 a 1	10 a 1

El agua se añadió a la emulsión, se mezcló y centrifugó a 4,400 rpm durante 10 minutos. A continuación del primer lavado, la fase acuosa se eliminó, y se realizó un segundo lavado con agua en la misma proporción, se mezcló y centrifugó a 4,400 rpm durante 10 minutos. A continuación del tratamiento, las emulsiones se observaron para la apariencia de una capa visible de aceite.

4. **Sonicación**

Un muestreo de las condiciones (muestras # 1, 9, 10, 14, 18, 19, 23, 27) a partir de la matriz de tratamiento químico/calor se sonicaron utilizando un sonicador Misonix 3000 equipado con una micro punta en el nivel 7. La suspensión se sonicó durante 3 minutos en un ciclo de 30 segundos de encendido/20 segundos de apagado. Después del tratamiento, una muestra de 1.5 ml se centrifugó a 14 K rpm durante 15 minutos y se midió el tamaño de cada una de las capas resultantes.

5. **Congelación de la emulsión**

Las muestras que generan emulsiones se sometieron a congelación a – 20°C durante >24 horas. A continuación de la congelación, las muestras se descongelaron a temperatura ambiente, se mezclaron y centrifugaron a 4,400 rpm durante 10 minutos. Las muestras estudiadas se enumeran en la Tabla 22.

Tabla 22. Emulsiones usadas en el cliclo de congelación.

Muestra	Sonicada
160 mN HCl, 130 °C por 30 min	Sí, ciclo de 30 min
120 mN HCl, 130 °C por 30 min	No
160 mN KOH 130 °C por 30 min	Sí, ciclo de 30 min
Sin químicos, 130 °C por 30 min	Sí, ciclo de 30 min
Sin químicos, sin calor	Sí, ciclo de 30 min

A continuación del tratamiento, las emulsiones se observaron para la apariencia de una capa visible de aceite.

6. **Experimentos de tratamiento enzimático #1**

Varias enzimas se evaluaron como un método alternativo de ruptura celular. Estas enzimas, todas obtenidas de Sigma Aldrich, fueron Hemicelulasa (P/N H2125), Pectinasa (P/N P2401), celulasa (C9422), y Driselasa (D9515). Cinco ml de material celular se preparó

de diferentes concentraciones de biomasa, e intercambió en el buffer. Las concentraciones enzimáticas se prepararon después como se describe en la Tabla 23.

Tabla 23. Matriz para el tratamiento de la enzima.

Enzima	Enzima conc. (%)	Temp. de incubación, °C	Tiempo de incubación, h	Tampón de incubación	[BM] g/l
Hemicelulasa	1%	25 °C	12	100 mM Acetato, pH 4.6	150
Pectinasa	1%	25 °C	12	100 mM Acetato, pH 4.6	150
Celulasa	1%	25 °C	12	100 mM Acetato, pH 4.6	150
Driselasa	1%	25 °C	12	100 mM Acetato, pH 4.6	150
Cóctel de las 4 enzimas	1% de cada una	25 °C	12	100 mM Acetato, pH 4.6	150
Hemicelulasa	1%	37 °C	1	100 mM Acetato, pH 4.6	150
Pectinasa	1%	37 °C	1	100 mM Acetato, pH 4.6	150
Celulasa	1%	37 °C	1	100 mM Acetato, pH 4.6	150
Driselasa	1%	37 °C	1	100 mM Acetato, pH 4.6	150
Cóctel de las 4 enzimas	1% de cada una	37 °C	1	100 mM Acetato, pH 4.6	150
Hemicelulasa/Driselasa	4%/2%	37 °C	12	50 mM Citrato, pH 4.5	150
Hemicelulasa/Driselasa	4%/2%	50 °C	12	50 mM Citrato, pH 4.5	150
Hemicelulasa/Driselasa	4%/2%	37 °C	12	50 mM Citrato, pH 4.5	100
Hemicelulasa/Driselasa	4%/2%	50 °C	12	50 mM Citrato, pH 4.5	100
Hemicelulasa/Driselasa	4%/2%	37 °C	12	50 mM Citrato, pH 4.5	50
Hemicelulasa/Driselasa	4%/2%	50 °C	12	50 mM Citrato, pH 4.5	50
Hemicelulasa/Driselasa	4%/2%	37 °C	12	50 mM Citrato, pH 4.5	25
Hemicelulasa/Driselasa	4%/2%	50 °C	12	50 mM Citrato, pH 4.5	25

5

7. Experimentos de tratamiento enzimático #2

Las enzimas adicionales se evaluaron solas y en combinación con otras, y se compararon con la ruptura de células mediante el almacenamiento y el envejecimiento de las células con o sin tratamiento enzimático. Una enzima manosa que degrada los polisacáridos y una proteasa se evaluaron por su capacidad para romper las células de *Chlorella protothecoides*. Además, se evaluó el efecto del almacenamiento y envejecimiento de una muestra de células cultivadas en el debilitamiento de la estructura celular para facilitar la extracción de aceite.

10

Se realizaron tres experimentos. Las células de *Chlorella protothecoides* se generaron a través del crecimiento heterótrofo usando la glucosa como única fuente de carbono. En experimentos en los que se usaron células secas, el caldo de fermentación se centrifugó y la pasta celular se sometió a secado en tambor. En experimentos en los que se almacenaron las células durante 7 días, el almacenamiento fue a 4°C. Una enzima de degradación de polisacáridos (Mannaway 4.0L) y la proteasa (Alcalase 2,4 FG) (ambos de

15

Novozymes) se usaron en los experimentos, como se indica. Las enzimas se añadieron a una concentración de 0.2% peso/volumen.

En el primer experimento, las células secas se reconstituyeron en agua desionizada a un peso seco celular de 155 gramos/litro en frascos, y mantuvieron a pH 7.5. Los frascos se colocaron en un agitador de incubación durante 22 horas a 50°C. La lisis celular enzimática para la muestra 1 en el Experimento 1 fue muy efectiva al combinar las enzimas, como se evidenció por el 90,4% de aceite total que se encuentra en la emulsión y no en el sedimento. Ninguna enzima de forma individual fue tan eficaz como la combinación de ambas, ni estuvo el control sin tratamiento enzimático. Los resultados se muestran más abajo en la Tabla 27.

En el segundo experimento, el caldo de fermentación fresco se separó en dos alícuotas para fresco y 7 días experimentos caldo. Las células frescas se concentraron por centrifugación y reconstituyeron en agua desionizada a 170 g/l en los frascos. Las muestras celulares resuspendidas se subdividieron después en dos grupos de 4 frascos, la mitad de las que se almacenaron durante 7 días a 4° C antes del tratamiento adicional. Para el tratamiento experimental, 3 frascos contenían las combinaciones de una proteasa y una manasa que degrada los polisacáridos, mientras que el 4^{to} no contenía enzimas (ver la Tabla 27 para las condiciones experimentales). Los frascos se colocaron en un agitador de incubación durante 22 horas a 50°C. Este segundo experimento indicó una lisis celular extensa con tratamiento enzimático doble en el día 0 y día 7, así como lisis en ausencia de las enzimas después de 7 días. Los resultados se muestran más abajo en la Tabla 28.

En el tercer experimento, el caldo de fermentación se recogió a partir de fermentaciones frescas, se concentró por centrifugación y se resuspendió en agua desionizada a dos concentraciones diferentes de 71 y 115 g/l y se agitó a 150 rpm en fermentadores Applikon durante 22 horas a 50°C. El pH se mantuvo en el 7.5 mediante la retroalimentación del control ácido/base. Los resultados de este tercer experimento indicaron que la ruptura celular y/o formación de la emulsión se facilita mediante la adición de enzimas y el envejecimiento de las células. Los resultados se muestran más abajo en la Tabla 29.

8. Análisis y caracterización del aceite

El TLC se realizó en placas TLC de sílice reforzada con aluminio (Sigma # 60805). El solvente de desarrollo (fase móvil) fue hexano/dietiléter/ácido acético (80:20:1, por volumen). Las manchas se visualizaron mediante pulverización con ácido sulfúrico al 10% y la carbonización en una placa de calentamiento. Las muestras de aceite recuperadas a partir

del control, condiciones ácidas y básicas se analizaron para determinar si ocurrió alguna degradación y/o modificación significativa del producto debido a las condiciones de procesamiento.

9. Resultados

5 (a) Tratamiento químico y térmico

Los resultados a partir de la matriz de tratamiento químico/térmico se informan en la Figura 3 y la Tabla 24 más abajo.

Tabla 24. Medición de las capas de la matriz de tratamiento químico/ térmico.

Tubo	Condición	[KOH] mN	[H ₂ SO ₄] mN	Temp, 30 min	Sedimento	Acuoso, mm	Emulsión, mm
1	Control	0	0	25 °C	15	15	0
2	40 mN KOH	40	0	25 °C	15	15	0
3	80 mN KOH	80	0	25 °C	15	15	0
4	120 mN KOH	120	0	25 °C	15	15	0
5	160 mN KOH	160	0	25 °C	15	15	0
6	40 mN H ₂ SO ₄	0	40	25 °C	15	15	0
7	80 mN H ₂ SO ₄	0	80	25 °C	15	15	0
8	120 mN H ₂ SO ₄	0	120	25 °C	15	15	0
9	160 mN H ₂ SO ₄	0	160	25 °C	15	15	0
10	Control	0	0	65 °C	10	19	1
11	40 mN KOH	40	0	65 °C	10	19	1
12	80 mN KOH	80	0	65 °C	10	19	1
13	120 mN KOH	120	0	65 °C	9	19	2
14	160 mN KOH	160	0	65 °C	8	19	3
15	40 mN H ₂ SO ₄	0	40	65 °C	10	19	1
16	80 mN H ₂ SO ₄	0	80	65 °C	10	19	1
17	120 mN H ₂ SO ₄	0	120	65 °C	10	19	1
18	160 mN H ₂ SO ₄	0	160	65 °C	10	19	1
19	Control	0	0	65 °C	5	15	10
20	40 mN KOH	40	0	130 °C	10	18	2
21	80 mN KOH	80	0	130 °C	5	15	10
22	120 mN KOH	120	0	130 °C	5	15	10
23	160 mN KOH	160	0	130 °C	5	15	10
24	40 mN H ₂ SO ₄	0	40	130 °C	2	23	5
25	80 mN H ₂ SO ₄	0	80	130 °C	1	19	10
26	120 mN H ₂ SO ₄	0	120	130 °C	1	19	10
27	160 mN H ₂ SO ₄	0	160	130 °C	1	19	10

10 A 25°C, ninguna diferencia se observó entre las muestras tratadas químicamente y el control, y no se formó la emulsión.

A 65°C, una emulsión comenzó a aparecer en las muestras tratadas con KOH, incrementando de tamaño con concentraciones mayores de sosa cáustica. La fase acuosa se tornó también más turbia con concentraciones crecientes de KOH. Parece existir una disminución proporcional en el tamaño del sedimento cuando se incrementa el tamaño de la emulsión. Para las muestras tratadas con ácido a 65°C, una emulsión más pequeña se forma en relación con las muestras tratadas con cáustica, y ningún aumento en el tamaño de la emulsión se observó con el incremento del nivel de ácido, y la fase acuosa fue cristalina y sin color.

Todas las muestras calentadas a 130°C mostraron emulsiones significativamente mayor en comparación con 65°C, y comprendieron apenas 1/3 del volumen de muestra total. La fase acuosa de las muestras tratadas con cáustica mostró incremento de color oscuro, con el incremento de la concentración de KOH.

En las muestras tratadas con ácido, la fase acuosa permaneció cristalina en todas las muestras. En contraste a los materiales tratados con base, el sedimento prácticamente desapareció para las preparaciones tratadas con ácido, lo que indica que las células y restos celulares habían separado a la emulsión y la fase acuosa. En ninguna de las 27 condiciones probadas se observó una capa visible de aceite.

(b) Lavado con agua de la emulsión

La emulsión de las muestras tratadas con cáustico fue pobremente definida y difícil de separar de la fase acuosa para preparar la emulsión por el lavado con agua. En contraste, la emulsión ácida se mantuvo firme y se separó fácilmente de la fase acuosa. En todas las condiciones de lavado de las muestras tratadas con KOH, la emulsión se convirtió en más y más delgada hasta que casi se perdió. No se observó capa de aceite visible durante cada lavado.

Las emulsiones ácidas fueron más fáciles de manipular y separar de la fase acuosa. El tamaño de la emulsión se mantuvo esencialmente sin cambios durante ambos lavados. Sin embargo, no se observó capa de aceite visible durante cada lavado.

(c) Sonicación

Los resultados de la sonicación se registran más abajo en la Tabla 25.

Tabla 25. Impacto de la sonicación comparado con las muestras pre-sonicadas correspondientes.

Tubo	[KOH] mN	[H ₂ SO ₄] mN	Temp, 30 min	Sedimento	Emulsión, mm
1	0	0	25 °C	-5	+1
9	0	160	25 °C	-5	+1
10	0	0	65 °C	-5	+1
14	160	0	65 °C	-3	+2
18	0	160	65 °C	-5	+1
19	0	0	130 °C	-2	-7
23	160	0	130 °C	-3	-7
27	0	160	130 °C	+1	-7

Las emulsiones de muestra a 25°C y 65°C aumentaron de tamaño, mientras que el sedimento disminuyó en relación con sus homólogos previamente sonicados. Para las muestras a 130°C, la emulsión disminuyó considerablemente de tamaño, y el sedimento del control y las muestras tratadas con base disminuyó, mientras el sedimento incrementó para la preparación tratada con ácido. Particularmente a elevadas temperaturas hubo cierta formación de capa de aceite para las muestras sonicadas.

(d) Reducción de la temperatura de emulsión

Los resultados de la congelación de emulsión se registran más abajo en la Tabla 26.

Tabla 26. Resultados de la congelación de la emulsión.

Muestra	Capa de aceite	
	Sonicada	visible
160 mN HCl, 130 °C por 30 min	Sí, ciclo de 30 min	Sí
120 mN HCl, 130 °C por 30 min	No	No
160 mN KOH 130 °C por 30 min	Sí, ciclo de 30 min	Sí
Sin químicos, 130 °C por 30 min	Sí, ciclo de 30 min	Sí
Sin químicos, sin calor	Sí, ciclo de 30 min	Sí

La reducción de la temperatura de la emulsión a – 20°C proporcionó ser un interruptor eficaz de la emulsión. En todas las muestras en las que se sonicó la emulsión, se observó una gran capa de aceite. Un resultado representativo se muestra en la Figura 4.

Cuando se comparó con el contenido lipídico del material celular mediante HPLC, la capa de aceite constituyó la mayoría del aceite de la emulsión.

(c) Tratamiento enzimático #1

Las muestras incubadas a 25°C y 37°C a niveles enzimáticos del 1% mostraron una mínima formación de emulsión. Las emulsiones bastante mayores se formaron a las

concentraciones enzimáticas superiores de la combinación 4%/2% Hemicelulasa/ Driselasa, con 50°C, produciendo una emulsión más grande en comparación con la muestra de 37°C. No se observó capa de aceite con el tratamiento enzimático sólo. Las emulsiones generadas por enzima sometidas a sonicación y después a – 20°C, generaron una capa de aceite.

5 (f) Tratamiento enzimático #2

El tratamiento enzimático que combina una mannasa que degrada polisacárido y una proteasa se encontró eficaz para romper las células de *Chlorella protothecoides*. Otra técnica encontrada eficaz en debilitar lo suficiente la estructura celular para facilitar la extracción del aceite fue el almacenamiento/envejecimiento de las células. La Tablas 27, 28 y 29 muestran los resultados de los tres experimentos descritos anteriormente, y la Figura 6 muestra las capas resultantes de la biomasa tratada con enzimas (tubo a la izquierda) vs sin tratar (tubo a la derecha). En el tubo sin tratar, el sedimento celular es considerablemente mayor en tamaño, ya que incluye un gran componente lipídico. En ningún caso se encontró lípido en la fase acuosa.

15 Tabla 27. Resultado del tratamiento enzimático #2, experimento 1.

	Número de la muestra			
	1	2	3	4
ALCALASE 2.4 FG	+	+	-	-
MANNAWAY 4.0 L	+	-	+	-
% de aceite en el sedimento celular	9.6%	30.7%	50.3%	97.8%
% del aceite en la emulsión	90.4%	69.3%	49.7%	2.2%

Tabla 28. Resultado del tratamiento enzimático #2, experimento 2.

	AGE	Número de la muestra			
		1	2	3	4
ALCALASE 2.4 FG		+	+	-	-
MANNAWAY 4.0 L		+	-	+	-
DCW G/L		170.0	170.0	170.0	170.0
	Days	%P/P	%P/P	%P/P	%P/P
% de aceite en el sedimento celular	0	28.8%	48.2%	67.8%	100.0%
% del aceite en la emulsión	0	71.2%	51.8%	32.2%	0%
% de aceite en el sedimento celular	7	47.8%	81.3%	80.7%	81.8%
% del aceite en la emulsión	7	52.2%	18.7%	19.3%	18.2%

Tabla 29. Resultado del tratamiento enzimático #2, experimento 3.

	Número de la muestra				
	1	2	3	4	5
ALCALASE 2.4 FG	-	-	+	-	-
MANNWAY 4.0 L	-	-	+	-	-
DCW G/L	71.7	115.1	115.1	71.7	115.1
TEMPERATURA	50	50	50	50	50
pH	7.5	7.5	7.5	7.5	7.5
	%P/P	%P/P	%P/P	%P/P	%P/P
% de aceite en el sedimento celular	81.8%	63.0%	12.0%	49.1%	42.0%
% del aceite en la emulsión	18.2%	37.0%	88.0%	50.9%	58.0%
age	0	0	0	7	7

(g) **Análisis y caracterización del aceite**

Una muestra de aceite se aisló de cada una de las condiciones químicamente tratadas y se analizó a través de TLC para determinar si la degradación del producto ocurrió a pH extremo comparado con la condición de control. El análisis de TLC se muestra en la Figura 5. El patrón mostró una generación menor de ácidos grasos libres (FFA) en condiciones de acidez en comparación con los tratados con base de y la muestra control.

Ejemplo 8: Métodos para el cultivo de *Prototheca*

Las cepas de *Prototheca* se cultivaron para lograr un alto porcentaje de aceite por peso seco de células. Las células criopreservadas se descongelaron a temperatura ambiente y 500 µl de células se añadieron a 4.5 ml de medio (4.2 g/l de K₂HPO₄, 3.1 g/l de NaH₂PO₄, 0.24 g/l de MgSO₄ • 7H₂O, 0.25 g/l de ácido cítrico monohidrato, 0.025 g/l de CaCl₂ 2H₂O, 2 g/l de extracto de levadura) más 2% de glucosa y crecieron durante 7 días a 28°C con agitación (200 rpm) en una placa de 6 pocillos. El peso seco de las células se determinó por centrifugación de 1 ml de cultivo a 14000 rpm durante 5 min en un tubo Eppendorf pre-pesado. El sobrenadante de cultivo se desechó y el sedimento celular resultante se lavó con 1 ml de agua desionizada. El cultivo se centrifugó de nuevo, se desechó el sobrenadante, y el sedimento de células se colocó a -80°C hasta que se congeló. Las muestras se liofilizaron después por 24 horas y se calculó el peso seco de las células. Para la determinación de lípidos totales en los cultivos, se retiraron 3 ml de cultivo y se sometieron a análisis mediante el uso de un sistema de Ankom (Ankom Inc., Macedon, NY), según el protocolo del fabricante. Las muestras se sometieron a extracción por solventes con un extractor Amkom XT10 según el

protocolo del fabricante. Los lípidos totales se determinaron como la diferencia de masa entre las muestras secas hidrolizadas con ácido y las muestras secas extraídas con disolvente. Las mediciones del porcentaje de aceite del peso seco de las células se muestran en la Tabla 30.

[0088] Tabla 30. Porcentaje de aceite por peso seco de las células

Especie	Cepa	% Aceite
<i>Prototheca stagnora</i>	UTEX 327	13.14
<i>Prototheca moriformis</i>	UTEX 1441	18.02
<i>Prototheca moriformis</i>	UTEX 1435	27.17

5

[0089] Se determinó el genotipo de las muestras de microalgas de las cepas que se enumeran en la Tabla 30. El ADN genómico se aisló a partir de la biomasa de algas de la siguiente manera. Las células de cultivos líquidos (aproximadamente 200 mg) se centrifugaron 5 minutos a 14,000 x g. Las células se resuspendieron después en agua destilada estéril, se centrifugaron 5 minutos a 14,000 xg y se desechó el sobrenadante. Una sola perla de vidrio de ~ 2 mm de diámetro se añadió a la biomasa y los tubos se colocaron a -80 °C durante al menos 15 minutos. Las muestras se retiraron y se añadieron 150 µl de tampón de molturación (Sarkosil 1%, sacarosa 0.25 M, NaCl 50 mM, EDTA 20 mM, Tris-HCl 100 mM, pH 8.0, ARNasa A 0.5 µg/µl). Los sedimentos se resuspendieron con una breve agitación, seguido por la adición de 40 µl de NaCl 5M. Las muestras se agitaron brevemente, seguido por la adición de 66 µl de CTAB 5% (bromuro cetil trimetilamonio) y una agitación breve final. Las muestras se incubaron a continuación a 65°C durante 10 minutos después de lo cual se centrifugaron a 14,000 x g durante 10 minutos. El sobrenadante se transfirió a un tubo nuevo y se extrajo una vez con 300 µl de fenol:cloroformo:alcohol isoamílico a 12:12:1, seguido por la centrifugación durante 5 minutos a 14,000 x g. La fase acuosa resultante se transfirió a un tubo nuevo que contenía 0.7 vol de isopropanol (~ 190 µl), se mezcló por inversión y se incubó a temperatura ambiente durante 30 minutos o toda la noche a 4 °C. El ADN se recuperó a través de la centrifugación a 14,000 x g durante 10 minutos. El precipitado resultante se lavó dos veces con 70% etanol, seguido de un lavado final con 100% etanol. Los sedimentos se secaron al aire durante 20-30 minutos a temperatura ambiente, seguido de la resuspensión en 50 µl de TrisCl 10 mM, EDTA 1 mM (pH 8.0).

15

20

25

[0090] Cinco µl del ADN total de algas, que se preparó como se describió anteriormente, se diluyeron 1:50 en Tris 10 mM, pH 8.0. Las reacciones de PCR, de volumen final de 20 µl, se establecieron de la siguiente manera. Diez µl de la mezcla maestra iProof

30

HF 2 x (BIO-RAD) se añadieron a 0.4 µl del cebador SZ02613 (5'-TGTTGAAGAATGAGCCGCGAC-3' (sec. con núm. de ident.: 5) a una concentración patrón de 10 mM). Esta secuencia del cebador corre de la posición 567-588 en el núm. de acceso L43357 del Gen Bank y está altamente conservada en los genomas de plantas superiores y plastidios de algas. Esto se siguió con la adición de 0.4 µl del cebador SZ02615 (5'-CAGTGAGCTATTACGCACTC-3' (sec. con núm. de ident.: 6) a la concentración patrón de 10 mM). Esta secuencia del cebador es complementaria a la posición de 1112-1093 en el núm. de acceso L43357 del Gen Bank y es muy conservada en los genomas de plantas superiores y plastidios de algas. A continuación, se añadieron 5 µl del ADN total diluido y 3.2 µl dH₂O. Las reacciones de PCR se llevaron a cabo de la siguiente manera: 98°C, 45", 98°C, 8", 53 °C, 12", 72°C, 20" por 35 ciclos seguidos de 72°C durante 1 min y se mantuvo a 25°C. Para la purificación de los productos de la PCR, se añadieron a cada reacción 20 µl de Tris 10 mM, pH 8.0, seguido por la extracción con 40 µl de fenol:cloroformo:alcohol isoamílico 12:12: 1, la agitación y centrifugación a 14,000 x g durante 5 minutos. Las reacciones de PCR se aplicaron a columnas S-400 (GE Healthcare) y se centrifugaron durante 2 minutos a 3,000 x g. Los productos de la PCR purificados posteriormente se clonaron en TOPO en el PCR8/GW/TOPO y los clones positivos se seleccionaron en placas LB /Spec. El ADN del plásmido purificado se secuenció en ambas direcciones mediante el uso de los cebadores M13 directo e inverso. En total, doce cepas de *Prototheca* se seleccionaron para tener las secuenciar de ADN de sus ARNr 23S y las secuencias se enumeran en el listado de secuencias. Un resumen de las cepas y números del listado de secuencias se incluye más abajo en la Tabla 31. Las secuencias se analizaron para la divergencia general de la secuencia UTEX 1435 (sec. con núm. de ident.: 17). Surgieron dos pares (UTEX 329/UTEX 1533 y UTEX 329/UTEX 1440) como los de máxima divergencia. En ambos casos, el alineamiento de pares resultó en 75.0% de identidad de secuencia por pares. El porcentaje de identidad de secuencia a UTEX 1435 se incluye, además, más abajo en la Tabla 31.

Tabla 31. Cepas de *Prototheca* genotipadas.

	<u>Especie</u> <u>ident.</u>	<u>Cepa</u>	<u>% identidad nt</u>	<u>sec. con núm. de</u>
30	<i>Prototheca kruegani</i> ident.: 34	UTEX 329	75.2	sec. con núm.
	<i>Prototheca wickerhamii</i> ident.: 35	UTEX 1440	99	sec. con núm.
35	<i>Prototheca stagnora</i> ident.: 16	UTEX 1442	75.7	sec. con núm.

	<i>Prototheca moriformis</i> ident.: 36	UTEX 288	75.4	sec. con núm.
	<i>Prototheca moriformis</i> ident.: 17	UTEX 1439; 1441;	100	sec. con núm.
5		1435; 1437		
	<i>Prototheca wikerhamii</i> ident.: 37	UTEX 1533	99.8	sec. con núm.
	<i>Prototheca moriformis</i> ident.: 38	UTEX 1434	75.9	sec. con núm.
10	<i>Prototheca zopfii</i> ident.: 39	UTEX 1438	75.7	sec. con núm.
	<i>Prototheca moriformis</i> ident.: 32	UTEX 1436	88.9	sec. con núm.

- 15 [0091] Las muestras de lípidos de un subconjunto de las cepas enumeradas anteriormente se analizaron para el perfil lipídico usando HPLC. Los resultados se muestran más abajo en la Tabla 32.

Tabla 32. Diversidad de cadenas de lípidos en especies *Prototheca*.

Cepa	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1
UTEX 327	0	12.01	0	0	50.33	17.14	0	0	0
UTEX 1441	1.41	29.44	0.70	3.05	57.72	12.37	0.97	0.33	0
UTEX 1435	1.09	25.77	0	2.75	54.01	11.90	2.44	0	0

- 20 Los péptidos de tránsito al plastidio de algas se identificaron a través del análisis de UTEX 1435 (*Prototheca moriformis*) o de UTEX 250 (*Chlorella protothecoides*). Los ADNc que codifican las proteínas potencialmente dirigidas a los plastidios, en base a los aciertos de homología de BLAST con otras proteínas conocidas dirigidas a los plastidios, se sometieron a un análisis adicional mediante los programas de software PSORT (<http://psort.ims.u-tokyo.ac.jp/form.html>), ChloroP (<http://www.cbs.dtu.dk/services/ChloroP/>) y TargetP (<http://www.cbs.dtu.dk/services/TargetP/>). Los péptidos de tránsito a plastidios candidatos que se identificaron a través de al menos uno de estos tres programas se amplificaron entonces por la PCR del ADN genómico apropiado. Más abajo, en la Tabla 33, se muestra un resumen de las secuencias de aminoácidos dirigidas a los plastidios de algas (PTS), que se
- 25
- 30 identificaron a partir de esta selección. También se incluyen las secuencias de aminoácidos de las acil grasos ACP tioesterasas de plantas que se usaron en los ejemplos de expresión heteróloga más abajo.

Tabla 33. Resumen de las secuencias de aminoácidos dirigidas a los plastidios de alga.

	<u>ADNc</u>	<u>sec. con núm. ident..</u>
	<i>P.moriformis</i> isopentenil difosfato sintasa PTS	sec. con núm. ident.: 40
5	<i>P.moriformis</i> delta 12-desaturasa de ácido graso PTS	sec. con núm. ident.: 41
	<i>P.moriformis</i> estearoil ACP desaturasa PTS	sec. con núm. ident.: 42
	<i>C.protothecoides</i> estearoil ACP desaturasa PTS	sec. con núm. ident.: 43
	<i>Cuphea hookeriana</i> acil graso-ACP tioesterasa (C8-10)	sec. con núm. ident.: 44
	<i>Umbellularia californica</i> acil graso-ACP tioesterasa (C12)	sec. con núm. ident.: 45
10	<i>Cinnamomum camphora</i> acil graso-ACP tioesterasa (C14)	sec. con núm. ident.: 46

Ejemplo 9: Métodos para la transformación de *Prototheca*

1. Método general para la transformación biolística de *Prototheca*

[0092] Los portadores de oro S550d de Seashell Technology se prepararon de acuerdo con el protocolo del fabricante. El plásmido linealizado (20 µg) se mezcló con 50 µl de tampón de unión y 60 µl (30 mg) de los portadores de oro S550d y se incubaron en hielo durante 1 min. El tampón de precipitación (100 µl) se añadió, y la mezcla se incubó en hielo durante otro minuto. Después de agitación, las partículas de ADN recubierto se sedimentaron por rotación a 10,000 rpm en una microcentrífuga Eppendorf 5415C durante 10 segundos. El precipitado de oro se lavó una vez con 500 µl de etanol frío al 100%, se sedimentó por rotación breve en la microcentrífuga, y se resuspendió con 50 µl de etanol helado. Después de una breve (1-2 seg) sonicación, 10 µl de partículas de ADN recubiertas se trasladaron de inmediato a la membrana portadora.

[0093] Las cepas de *Prototheca* se cultivaron en medio proteosa (2 g/l de extracto de levadura, NaNO3 2.94 mM, CaCl2 • 2H2O 0.17 mM, MgSO4 • 7H2O 0.3 mM, K2HPO4 0.4 mM, KH2PO4 1.28 mM, NaCl 0.43 mM) en un agitador giratorio hasta alcanzar una densidad celular de 2x10⁶ células/ml. Las células se cosecharon, se lavaron una vez con agua destilada estéril, y se resuspendieron en 50 µl de medio. Se extendieron 1 x 10⁷ células en el tercer centro de una placa con medio no selectivo de proteosa. Las células se bombardearon con el sistema biolístico de entrega de partículas PDS-1000/He (Bio-Rad). Se usaron los discos de ruptura (1100 y 1350 psi), y las placas se colocaron 9 y 12 cm por debajo del ensamblaje de la pantalla/macro-portador. A las células se les permitió recuperarse a 25°C durante 12-24 h. Tras la recuperación, las células se rasparon de las placas con una espátula

de goma, se mezclaron con 100 µl de medio y se dispersaron en las placas que contenían la selección de antibióticos adecuada. Después de 7-10 días de incubación a 25 °C, las colonias que representan las células transformadas eran visibles en las placas a partir de los discos de ruptura de 1100 y 1350 psi y a partir de 9 y 12 cm de distancia. Las colonias se recogieron y se aplicaron en forma de puntos en placas de agar selectivo para una segunda ronda de selección.

2. Transformación de *Prototheca* con el gen de resistencia al G418

[0094] La *Prototheca moriformis* y otras cepas de *Prototheca* sensibles al G418 se pueden transformar mediante el uso de los métodos que se describen más abajo. El G418 es un antibiótico aminoglucósido que inhibe la función de los ribosomas 80S y por ello, inhibe la síntesis de proteínas. El correspondiente gen de resistencia funciona a través de la fosforilación, lo que resulta en la inactivación del G418. Las cepas de *Prototheca* UTEX 1435, UTEX 1439 y UTEX 1437 se seleccionaron para la transformación. A las tres cepas de *Prototheca* se les determinó el genotipo mediante el uso de los métodos que se describieron anteriormente. Las tres cepas de *Prototheca* tuvieron secuencias genómicas idénticas del 23S ARNr (sec. con núm. de ident.: 17).

[0095] Todos los casetes de transformación se clonaron como fragmentos EcoRI-SacI en pUC19. Se usaron técnicas estándar de biología molecular en el constructo de todos los vectores de acuerdo a Sambrook y Russell, 2001. El promotor de la beta-tubulina/5'UTR de *C. reinhardtii* se obtuvo del plásmido pHyg3 (Berthold y otros, (2002) Protist: 153(4), pp 401-412) por la PCR como un fragmento EcoRI-AscI. La nitrato reductasa 3'UTR de *Chlorella vulgaris* se obtuvo a partir de ADN genómico aislado de la cepa UTEX 1803 a través de la PCR mediante el uso de los pares de cebadores siguientes:

Directo:

5' TGACCTAGGTGATTAATTAACTCGAGGCAGCAGCAGCTCGGATAGTATCG 3'
(sec. con núm. de ident.:47)

Inverso:

5' CTACGAGCTCAAGCTTTCCATTTGTGTTC CCATCCCACACTTCC 3'
(sec. con núm. de ident.:48)

[0096] El promotor/UTR de la glutamato deshidrogenasa de *Chlorella sorokiniana* se obtuvo a través de la PCR del ADN genómico aislado de la cepa UTEX 1230 a través de la PCR mediante el uso de los pares de cebadores siguientes:

Directo:5' GATCAGAATTCCGCCTGCAACGCAAGG GCAGC 3' (sec. con núm. de ident.:34)

Inverso:5' GCATACTAGTGGCGGGACGGAGAGA GGGCG 3' (sec. con núm. de ident.:35)

5 [0097] La optimización de codón se basó en los codones para *Prototheca moriformis*. La secuencia del casete de la neomicina fosfotransferasa en la que no se optimizaron los codones (nptII) se sintetizó como un fragmento AscI-XhoI y se basó en la secuencia con núm. de acceso YP_788126 del GenBank. El casete nptII con los codones optimizados se basó también en este número de acceso del Genbank.

10 [0098] Las tres cepas de *Prototheca* se transformaron usando los métodos biolísticos descritos anteriormente. En resumen, las cepas de *Prototheca* se cultivaron heterotróficamente en medio líquido que contenía 2% de glucosa hasta alcanzar la densidad celular deseada (1×10^7 células/ml a 5×10^7 células/ml). Las células se cosecharon, se lavaron una vez con agua destilada estéril y se resuspendieron a 1×10^8 células/ml. Después, 0.5 ml de
15 células se extendieron en una placa con medio sólido no selectivo y se dejó secar en una campana estéril. Las células se bombardearon con el sistema biolístico de entrega de partículas PDS-1000/He (BioRad). Las células se dejaron recuperar a 25°C durante 24 horas. Tras la recuperación, las células se retiraron mediante el lavado de las placas con 1 ml de medio estéril y se transfirieron a nuevas placas que contenían 100 mg/ml de G418. Las
20 células se dejaron secar en una campana estéril y se dejó que se formaran colonias en la placa a temperatura ambiente por hasta tres semanas. Las colonias de UTEX 1435, UTEX 1439 y UTEX 1437 se recogieron y se aplicaron en forma de puntos en placas selectivas de agar para una segunda ronda de selección.

[0099] Un subconjunto de las colonias que sobrevivieron a una segunda ronda
25 de selección descrita anteriormente, se cultivaron en pequeño volumen y se extrajo el ADN y ARN genómico por medio de técnicas de biología molecular. Se hizo la transferencia de tipo Southern en el ADN genómico extraído de las no transformadas (WT), los transformantes y el ADN del plásmido. El ADN de cada muestra se corrió en 0.8% de geles de agarosa después de los siguientes tratamientos: sin digerir (U), digerido con AvrII (A), digerido con
30 NcoI (N), digerido con SacI (S). Los ADN de estos geles se transfirieron a membranas de Nylon + (Amersham). Estas membranas se probaron con un fragmento que corresponde a toda la región de codificación del gen nptII (sonda NeoR). La Figura 7 muestra los mapas de

los casetes que se usaron en las transformaciones. La Figura 8 muestra los resultados del análisis de la transferencia de tipo Southern en tres transformantes (todos generados en la cepa UTEX 1435) (1, 2, y 3) que se transformaron ya sea con la beta-tubulina::neo::nit (sec. con núm. de ident.: 49) (los transformantes 1 y 2) o con la glutamato deshidrogenasa::neo::nit (sec. con num. de ident.: 50) (transformante 3). El plásmido de transformación glutamato deshidrogenasa::neo::nit se corrió como un control y se cortó tanto con NcoI y con SacI. La AvrII no cortó en este plásmido. El ADN genómico aislado de la cepa sin transformar UTEX 1435 no mostró hibridación de la sonda NeoR.

L[0100] os constructos de transformantes adicionales que contienen la glutamato deshidrogenasa::neo::nit con codón optimizado (sec. con num. de ident.: 51) y la β -tubulina::neo::nit con codón optimizado (sec. con num. de ident.: 52) se recogieron y se analizaron por análisis de transferencia de tipo Southern. Como era de esperar, sólo la digestión con la SacI mostró linealización del ADN de transformación. Estos eventos de transformación son consistentes con los eventos de integración que se producen en forma de oligómeros del plásmido de transformación. Sólo la digestión con las enzimas de restricción que cortan el ADN del plásmido de transformación hace en estas moléculas la disminución del tamaño del plásmido de transformación.

[0101] El análisis de transferencia de tipo Southern se realizó también en transformantes generados en la transformación de las cepas de *Prototheca* UTEX 1437 y UTEX 1439 con el casete de la glutamato deshidrogenasa::neo::nit. La transferencia se probó con la sonda NeoR y los resultados son similares a los transformantes UTEX 1435. Los resultados son indicativos de eventos de integración que se caracterizaron por la oligomerización y la integración del plásmido de transformación. Este tipo de evento de integración se sabe que ocurre con bastante frecuencia en *Dictyostelium discoideum* (véase, por ejemplo, Kuspa, A. y Loomis, W. (1992) *PNAS*, 89:8803-8807 y Morio y otros, (1995) *J. Plant Res.* 108:111-114).

[0102] Para confirmar aún más la expresión del plásmido de transformación, en los transformantes que se seleccionaron se reañizó el análisis por transferencia de tipo Northern y el análisis por RT-PCR. La extracción de ARN se realizó mediante el uso del reactivo Trizol según las instrucciones del fabricante. El análisis de transferencia de tipo Northern se llevó a cabo de acuerdo con los métodos publicados en Sambrook y Russell, 2001. El ARN total (15 μ g) aislado de cinco transformantes UTEX 1435 y UTEX 1435 sin

transformar (carriles de control) se separó en gel de agarosa-formaldehído 1% y se transfirió a membrana de nylon. La transferencia se hibridó con la sonda neo sin optimizar, específica para las secuencias transgénicas en los transformantes 1 y 3. Los otros dos ARN transformantes, expresaron la versión de los neo-transgénicos con codón optimizado y, como era de esperar, en base a la homología de secuencia entre los neo genes que se optimizaron y los que no se optimizaron, mostraron la señal de hibridación significativamente menor.

[0103] Se extrajo ARN (1 µg) de la cepa de *Prototheca* UTEX 1435 sin transformar y dos representantes transformantes de UTEX 1435 y se realizó la transcripción inversa mediante el uso de un cebador oligo dT o un cebador específico del gen. Posteriormente estos ADNc (por duplicado) se sometieron a análisis de qPCR en el termociclador ABI Veriti que usa la química SYBR-Green de qPCR mediante el uso de los siguientes cebadores (nptII):

Directo: 5' GCCGCGACTGGCTGCTGCTGG 3' (sec. con núm. de ident.: 53)

Inverso: 5' AGGTCCTCGCCGTCGGGCATG 3' (sec. con núm. de ident.: 54)

[0104] La posible contaminación de ADN genómico se descartó por una muestra de control negativo sin transcriptasa inversa. Los resultados indican que los genes NeoR que se usaron para transformar estas cepas se transcribieron activamente en los transformantes.

3. Transformación de *Prototheca* con sacarosa invertasa secretada heteróloga

[0105] Los siguientes experimentos se realizaron mediante el uso de placas de medio líquido/ agar en base al medio basal descrito en Ueno y otros, (2002) *J Bioscience and Bioengineering* 94(2):160-65, con la adición de trazas de minerales descrita en la patente de Estados Unidos núm. 5,900,370, y el cóctel de vitamina DAS 1x (solución 1000x): tricina: 9g, tiamina HCL: 0.67g, biotina: 0.01g, cianocobalamina (vitamina B12): 0.008g, pantotenato de calcio: 0.02 g y ácido p-aminobenzoico: 0.04 g).

[0106] Dos constructos de plásmido se ensamblaron mediante el uso de las técnicas estándar de ADN recombinante. Los genes de la sacarosa invertasa de levadura (uno con codones optimizados y uno con sin codones optimizados), suc2, estaban bajo el control del promotor de la beta-tubulina de *Chlorella reinhardtii*/5'UTR y tenía la nitrato reductasa de *Chlorella vulgaris* 3'UTR. Las secuencias (incluyendo las secuencias 5'UTR y 3'UTR) para el constructo con codón no optimizado (Crβ-tub::NCO-suc2::CvNitRed), sec. con núm.

de ident.: 55, y el constructo con codón optimizado (Cr β -tub::CO-suc2::CvNitRed), sec. con
 núm. de ident.: 56, se enumeran en el listado de secuencias. La optimización de codones se
 basó en *Prototheca sp* (ver Tabla 4). La Figura 9 muestra un esquema de los dos constructos
 con los sitios de restricción de clonación relevantes y las flechas indican la dirección de la
 5 transcripción. La selección se proporcionó por Neo R. La optimización de codones se basó en
 el uso del codón preferido en cepas *Prototheca* en la Tabla 4.

[0107] Preparación del microportador ADN/oro: los microportadores
 ADN/oro se prepararon inmediatamente antes de su uso y se almacenaron en hielo hasta que
 se aplicaron a los macroportadores. El ADN del plásmido (en tampón TE) se añadió a 50 μ l
 10 de tampón de unión. La saturación de las perlas de oro se alcanzó a los 15 μ g de ADN del
 plásmido para 3 mg del portador de oro. El tampón de unión y el ADN se mezclaron bien a
 través de la agitación. El ADN y el tampón de unión se deben pre-mezclar antes de la adición
 de oro para asegurar la unión uniforme del plásmido a las partículas del portador de oro. Se
 añadieron 60 μ l del portador de oro S550d (Seashell Technologies, San Diego, CA) a la
 15 mezcla ADN/tampón de unión. Para un patrón de oro a 50 mg/ml, la adición de 60 μ l resulta
 en una proporción óptima de 15 μ g ADN/3 mg del portador de oro. La mezcla del portador de
 oro/ADN se dejó incubar en hielo durante 1 minuto y entonces se añadieron 100 μ l de
 tampón de precipitación. La mezcla se dejó incubar nuevamente en hielo durante 1 minuto y
 entonces brevemente se agitó y se centrifugó a 10,000 rpm a temperatura ambiente durante
 20 10 segundos para sedimentar el portador de oro. Se retiró el sobrenadante cuidadosamente
 con una pipeta y el sedimento se lavó con 500 μ l de etanol 100% helado. Las partículas de
 oro se re-sedimentaron por centrifugación de nuevo a 10,000 rpm durante 10 segundos. El
 etanol se eliminó y se añadieron 50 μ l de etanol helado a la mezcla de oro. Inmediatamente
 antes de aplicar el oro a los macroportadores, el oro/etanol se resuspendió con un breve pulso
 25 de 1-2 segundos en el nivel 2 en un baño de ultrasonido Misonix mediante el uso de la micro
 punta. Inmediatamente después de la resuspensión, 10 μ l de la dispersión de partículas de oro
 se trasladaron al macroportador y se dejaron secar en una campana estéril.

[0108] Las dos cepas de *Prototheca moriformis* (UTEX 1435 y 1441) se
 cultivaron heterofilicamente en medio líquido que contenía 2% de glucosa de los frascos que
 30 se criopreservaron. Cada cepa se cultivó a una densidad de 10^7 células/ml. Este cultivo de
 semillas entonces se diluyó con medio fresco a una densidad de 10^5 células/ml y se dejó
 crecer durante 12-15 horas para alcanzar una densidad celular final de aproximadamente 10^6
 células/ml. Las microalgas se formaron en alícuotas en tubos cónicos de 50 ml y se

centrifugaron durante 10 minutos a 3500 rpm. Las células se lavaron con medio fresco y se centrifugaron de nuevo durante 10 minutos a 3500 rpm. Las células se resuspendieron después a una densidad de 1.25×10^8 células/ml en medio fresco.

[0109] En una campana estéril se retiraron 0.4 ml de las células que se prepararon anteriormente y se colocaron directamente en el centro de una placa de agar (sin agente de selección). La placa se agitó suavemente con un nivel de movimiento circular para distribuir uniformemente las células a un diámetro de no más de 3 cm. Las células se dejaron secar en las placas en la campana estéril durante aproximadamente 30-40 minutos y entonces se bombardearon a una presión del disco de ruptura de 1350 psi y una placa con el macroportador a 6 cm de distancia. Las placas se cubrieron y se envolvieron con parafilm y se dejaron incubar con poca luz durante 24 horas.

[0110] Después de la recuperación de 24 horas, se introdujo 1 ml de medio estéril (sin glucosa) en el césped de células. Las células se resuspendieron mediante el uso de un asa estéril, que se aplicó en un movimiento circular al césped de las células y las células resuspendidas se recogieron con una pipeta estéril. Las células se sembraron a continuación, sobre una placa de agar fresco con 2% de glucosa y 100 mg/ml de G418. La aparición de colonias se produjo 7-12 días después de que se sembraron. Las colonias individuales se recogieron y se cultivaron en medio selectivo con 2% de glucosa y 100 µg/ml de G418. El tipo salvaje (sin transformar) y las células transgénicas se analizaron para la introducción, la integración y la expresión exitosa del transgén.

[0111] El ADN genómico de *Prototheca moriformis* UTEX 1435 y 1441 transformadas y sus contrapartes de tipo salvaje (sin transformar) se aislaron mediante el uso de métodos estándar. En resumen, las células se centrifugaron durante 5 minutos a 14,000 rpm en una centrífuga de mesa estándar Eppendorf (modelo 5418) y se congelaron rápidamente antes de la extracción de ADN. Los sedimentos de células se lisaron mediante la adición de 200 µl de tampón de lisis (Tris HCl 100 mM, pH 8.0, lauril sarcosina 1%, NaCl 50 mM, EDTA 20 mM, sacarosa 0.25 M, ARNasa A 0.5 mg/ml) por cada 100 - 200 mg de células (peso húmedo) y agitación durante 30-60 segundos. El bromuro de cetil trimetil amonio (CTAB) y el NaCl se llevaron a 1% y 1 M, respectivamente, y los extractos de células se incubaron a 60-65 °C durante 10 minutos. Posteriormente, los extractos se aclararon por centrifugación a 14,000 rpm durante 10 minutos y el sobrenadante resultante se extrajo con un volumen igual de fenol/cloroformo/alcohol isoamílico (25:24:1). Las muestras

se centrifugaron durante 5 minutos a 14,000 rpm y la fase acuosa se eliminó. El ADN se precipitó con 0.7 volúmenes de isopropanol. El ADN se granuló a través de centrifugación a 14,000 rpm durante 10 minutos y se lavó dos veces con etanol al 80%, y una vez con etanol. Después del secado, el ADN se resuspendió en Tris HCl 10 mM, pH 8.0 y se determinaron las concentraciones de ADN mediante el uso del ensayo de cuantificación de la fluorescencia PicoGreen (Molecular Probes).

El ARN de *Prototheca moriformis* UTEX 1435 y 1441 transformadas y sus contrapartes de tipo salvaje (sin transformar) se aislaron mediante el uso de métodos estándar. En resumen, las células se centrifugaron durante 5 minutos a 14,000 rpm en una centrífuga de mesa estándar Eppendorf (modelo 5418) y se congelaron rápidamente antes de la extracción del ARN. Los sedimentos de células se lisaron mediante la adición de 1 ml del reactivo Trizol (Sigma) por cada 100 mg de células (peso húmedo) y mediante agitación durante 1-2 minutos. Las muestras se incubaron a temperatura ambiente durante 5 minutos y posteriormente ajustaron con 200 µl de cloroformo por 1 ml del reactivo Trizol. Después de amplia agitación, las células se incubaron a temperatura ambiente durante 15 minutos y entonces se sometieron a centrifugación a 14,000 rpm durante 15 minutos en una microcentrífuga de mesa refrigerada. La partición del ARN en la fase acuosa superior se retiró y se precipitó por la adición de isopropanol (500 µl por 1 ml del reactivo Trizol). El ARN se recogió por centrifugación durante 10 minutos y el sedimento resultante se lavó dos veces con 1 ml de etanol al 80%, se secó, y se resuspendió en agua libre de ARNasa. La concentración de ARN se estimó mediante el ensayo de cuantificación de fluorescencia RiboGreen (Molecular Probes).

[0112] La expresión del gen de la neomicina fosfotransferasa que confiere resistencia al antibiótico G418 y la levadura invertasa se probaron en las *Prototheca moriformis* no transformadas UTEX 1435 y 1441 y los transformantes T98 (UTEX transformante 1435) y T97 (UTEX 1441 transformante) mediante el uso del análisis por PCR cuantitativa en transcripción inversa (RT-PCR cuantitativa). 20 ng de ARN total (aislado como se describió anteriormente) se sometieron a una etapa de análisis de RT-PCR cuantitativa mediante el uso del kit de RT-PCR iScript SYBR Green (BioRad Laboratories) y pares de cebadores dirigidos al gen de resistencia a la neomicina (cebador directo 5'CCGCCGTGCTGGACGTGGTG 3' y el cebador inverso 5'GGTGGCGGGGTCCAGGGTGT 3', sec. con núm. de ident.: 57 y 58, respectivamente) y los transcritos de *suc2* de invertasa (cebador directo 5' CGGCCGGCGGCTCCTTCAAC 3' y

cebador inverso 5' GGCGCTCCCGTAGGTCGGGT 3', sec. con num. de ident.: 59 y 60, respectivamente). Los transcritos de la beta tubulina endógena sirvieron como control interno positivo para la amplificación por PCR y como referencia de normalización para estimar los niveles relativos de transcripción.

5 [0113] Ambos constructos con codón optimizado y sin codón optimizado se transformaron en las células de *Prototheca moriformis* UTEX 1435 y 1441 como se describió anteriormente. Inicialmente, los transformantes se obtuvieron con los dos constructos y la presencia del transgén se comprobó mediante el análisis de transferencia de tipo Southern
10 seguido por RT-PCR para confirmar la presencia del ADN y el ARNm del transgén. Para el análisis de transferencia de tipo Southern, el ADN genómico aislado como se describió anteriormente se sometió a electroforesis en geles de agarosa al 0.7% en tampón TAE 1x. Los geles se procesaron como se describe en Sambrook y otros (Molecular Cloning; A Laboratory Manual, 2^{da} edición. Cold Spring Harbor Laboratory Press, 1989). Las sondas se prepararon por cebado aleatorio e hibridaciones que se realizaron como se describe en Sambrook y otros.
15 Los transformantes de los constructos ya sea con codón optimizado y sin codón optimizado, mostraron la presencia del casete de la invertasa, mientras que el control no transformado fue negativo. El ARNm de la invertasa se detectó también en transformantes con los constructos ya sea con codón optimizado y sin codón optimizado.

[0114] Para confirmar que los transformantes expresaban una proteína
20 invertasa activa, los transformantes se sembraron en placas de sacarosa. Los transformantes que contenían el casete de codón no optimizado no crecieron en las placas que contenían sacarosa, lo que indicaba que, si bien los genes y el ARNm que codifica la proteína suc2 estaban presentes, la proteína o bien (1) no se tradujo, o (2) se tradujo, pero no se acumuló a niveles suficientes para permitir el crecimiento en la sacarosa como única fuente de carbono.
25 Los transformantes con el casete de codón optimizado crecieron en las placas que contenían sacarosa. Para evaluar los niveles de la invertasa que expresaban estos transformantes, dos clones (T98 y T97) se sometieron a un ensayo de la invertasa de todas las células raspadas del medio sólido y el muestreo directo y la cuantificación de azúcares en los sobrenadantes de cultivo después de 48 horas de crecimiento en el medio líquido que contenía 2% de sacarosa
30 como única fuente de carbono.

[0115] Para el ensayo de la invertasa, las células (T98 y T97) se cultivaron en placas que contenían 2% de sacarosa, se rasparon y se probó la actividad de la invertasa. Se

mezclaron 10 μ l de las células tomadas por raspado con 40 μ l de NaOAc 50mM pH 5.1. Se añadieron a la mezcla de la células 12.5 μ l de sacarosa 0.5 M y se incubaron a 37°C durante 10-30 minutos. Para detener la reacción se añadieron 75 μ l de K_2HPO_4 0.2 M. Para el ensayo de la glucosa liberada, se añadieron 500 μ l del reactivo reconstituido (glucosa oxidasa/peroxidasa + o-dianisidina) de Sigma (kit de prueba GAGO-20) a cada tubo y se incubaron a 37°C durante 30 minutos. Una curva de glucosa estándar se creó también en este momento (intervalo: 25 μ l a 0.3 μ l de glucosa). Después de la incubación, se añadieron 500 μ l de HCl 6N a detener la reacción y desarrollar el color. Las muestras se leyeron a 540 nm. La cantidad de glucosa liberada se calculó a partir de la curva estándar de glucosa mediante la fórmula $y = mx + c$, donde y es la lectura de 540 nm, y x son los μ l de glucosa. El peso de glucosa se convirtió en moles de glucosa, y dada la relación equimolar entre moles de sacarosa hidrolizada a los moles de glucosa generada, los datos se expresaron como nmoles de sacarosa hidrolizada por unidad de tiempo. El ensayo mostró que tanto los clones T98 y T97 hidrolizaron la sacarosa, lo que indicó que las células produjeron y secretaron una invertasa de sacarosa funcional.

[0116] Para el análisis de azúcar en los medios de cultivo líquido después de 48 horas de crecimiento de las algas, las células T97 y T98 se cultivaron en medio que contenía 2% de sacarosa durante 48 horas y los medios de cultivo se procesaron para el análisis de azúcar. Los caldos de cultivo de cada transformante (y el control negativo de células sin transformar) se centrifugaron a 14,000 rpm durante 5 minutos. El sobrenadante resultante se retiró y se sometió a HPLC/ELSD (detección de dispersión de la luz por evaporación). La cantidad de azúcar en cada muestra se determinó mediante el uso de estándares externos y el análisis de regresión de línea. Los niveles de sacarosa en el medio de cultivo de los transformantes fueron muy bajos (menos de 1.2 g/l, y en la mayoría de los casos 0 g/l). En los controles negativos, los niveles de sacarosa se mantuvieron altos, aproximadamente a 19 g/l después de 48 horas de crecimiento.

[0117] Estos resultados fueron consistentes con los resultados de la actividad de la invertasa, y en conjunto, indicaron que los transformantes con codones optimizados, T97 y T98, secretaron una sacarosa invertasa activa que les permitió a las microalgas usar la sacarosa como única fuente de carbono, en contraste con (1) los transformantes sin codones optimizados y (2) las microalgas no transformadas de tipo salvaje, las cuales no pudieron usar sacarosa como única fuente de carbono en el medio de cultivo.

[0118] Las cepas de *Prototheca moriformis*, T98 y T97, que expresaban un transgén de sacarosa invertasa segregada (*suc2*), funcional se probaron para el crecimiento y la producción de lípidos mediante el uso de la sacarosa como única fuente de carbono.

[0119] Las cepas T98 y T97 de tipo salvaje (no transformadas), se cultivaron en medio de crecimiento (como se describió anteriormente) conteniendo 4% de glucosa o 4% de sacarosa como la única fuente de carbono bajo condiciones heterótrofas por aproximadamente 6 días. El crecimiento, según lo determinado por las lecturas de densidad óptica A750 que se tomaron de las cuatro muestras cada 24 horas y el peso seco de las células de los cultivos y los perfiles de lípidos se determinaron después de 6 días de crecimiento. Las lecturas de densidad óptica de las cepas transgénicas que crecieron tanto en las condiciones de glucosa y de sacarosa fueron comparables al crecimiento de las cepas de tipo salvaje en las condiciones de glucosa. Estos resultados indican que las cepas transgénicas crecieron en glucosa o sacarosa como única fuente de carbono a una velocidad igual a las cepas del tipo salvaje en condiciones de glucosa. Las cepas de tipo salvaje, sin transformar, no crecieron en la condición de sacarosa sola.

[0120] Se analizó el perfil de lípidos en la biomasa de la cepa de tipo salvaje cultivada en glucosa y de la cepa T98 cultivada en sacarosa. Las muestras de lípidos se prepararon a partir de la biomasa seca (lío-filizada) usando un sistema de hidrólisis ácida (Ankom Technology, NY) de acuerdo con las instrucciones del fabricante. Las determinaciones del perfil lipídico se realizaron como se describió en el Ejemplo 6. El perfil lipídico de la cepa de *Prototheca moriformis* sin transformar UTEX cepa 1435, cultivada en glucosa como única fuente de carbono y dos cepas clonal T98 (UTEX 1435 transformada con un transgén de sacarosa invertasa), cultivadas en sacarosa como única fuente de carbono, se describen en la Tabla 34 (tipo salvaje UTEX 1435 y clon 8 y el clon 11 de T98 más abajo). El lípido C:19:0 se usó como un control de calibración interna.

[0121] Tabla 34. Perfil lipídico de UTEX 1435 tipo salvaje y los clones de UTEX 1435 con el transgén *suc2*.

Nombre	tipo salvaje (% área - ISTD)	T98 clon 11 (% área - ISTD)	T98 clon 8 (% área - ISTD)
C 12:0	0.05	0.05	0.05
C 14:0	1.66	1.51	1.48
C 14:1	0.04	ND	ND
C 15:0	0.05	0.05	0.04
C 16:0	27.27	26.39	26.50
C 16:1	0.86	0.80	0.84

C 17:0	0.15	0.18	0.14
C 17:1	0.05	0.07	0.05
C 18:0	3.35	4.37	4.50
C 18:1	53.05	54.48	54.50
C 18:2	11.79	10.33	10.24
C 19:0 (ISTD)	--	--	--
C 18:3 alfa	0.90	0.84	0.81
C 20:0	0.32	0.40	0.38
C 20:1	0.10	0.13	0.12
C 20:1	0.04	0.05	0.04
C 22:0	0.12	0.16	0.12
C 20:3	0.07	0.08	0.07
C 24:0	0.12	0.11	0.10
nd - denota ninguno detectado			

El aceite extraído de *Prototheca moriformis* UTEX 1435 de tipo salvaje (a través de extracción con solventes o usando una prensa expulsora) se analizó para los carotenoides, la clorofila, los tocoferoles, otros esterole y tocotrienoles. Los resultados se resumen más abajo en la Tabla 35.

[0122] Tabla 35. Análisis de carotinoide, clorofila, tocoferol/esteroles y tocotrienol en aceite extraído de *Prototheca moriformis* (UTEX 1435).

	Aceite prensado (mcg/ml)	Aceite extraído del solvente (mcg/ml)
cis-Luteína	0.041	0.042
trans-Luteína	0.140	0.112
trans-Zeaxantina	0.045	0.039
cis-Zeaxantina	0.007	0.013
t-alfa-Critoxantina	0.007	0.010
t-beta-Critoxantina	0.009	0.010
t-alfa-Caroteno	0.003	0.001
c-alfa-Caroteno	ninguno detectado	ninguno detectado
t-beta-Caroteno	0.010	0.009
9-cis-beta-Caroteno	0.004	0.002
Licopeno	ninguno detectado	ninguno detectado
Total de Carotenoides	0.267	0.238
Clorofila	<0.01 mg/kg	<0.01 mg/kg
Tocoferoles y Esteroles		
	Aceite prensado (mg/100g)	Aceite extraído del solvente (mg/100g)
gamma Tocoferol	0.49	0.49
Campesterol	6.09	6.05
Estigmasterol	47.6	47.8
Beta-sitosterol	11.6	11.5
Otros esteroleess	445	446
Tocotrienoles		
	Aceite prensado (mg/g)	Aceite extraído del solvente (mg/g)

alfa Tocotrienol	0.26	0.26
beta Tocotrienol	< 0.01	< 0.01
gamma Tocotrienol	0.10	0.10
delta Tocotrienol	< 0.01	< 0.01
Total de Tocotrienoles	0.36	0.36

[0123] La capacidad de usar sacarosa como única fuente de carbono como el factor de selección de los clones que contenían el constructo del transgén *suc2* en lugar del G418 (u otro antibiótico) se evaluó mediante el uso de transformantes positivo para el gen *suc2*. Un subconjunto de los transformantes positivos cultivados durante 24 doblajes en placas que contenían sacarosa como única fuente de carbono y sin selección de antibiótico. Los clones entonces se retaron, con placas que contenían glucosa como única fuente de carbono y G418. Hubo un subconjunto de los clones que no se cultivó con la condición de glucosa + G418, lo que indicó una pérdida de la expresión del transgén. Un experimento adicional se realizó mediante el uso de una placa que contenía sacarosa como única fuente de carbono y no G418 y se estió un clon que expresaba el transgén *suc2* en la mitad de la placa y la *Prototheca moriformis* de tipo salvaje en la otra mitad de la placa. El crecimiento se vio tanto con el tipo salvaje como con las células de *Prototheca moriformis* que contenían el transgen. La *Prototheca moriformis* de tipo salvaje no demostró capacidad de crecer en sacarosa, por lo tanto, este resultado muestra que a diferencia de la resistencia a los antibióticos, el uso de la selección con sacarosa/invertasa no es por células autónomas. Es muy probable que los transformantes secretaran suficiente sacarosa invertasa en la placa/medio para apoyar el crecimiento del tipo salvaje mientras que la sacarosa se hidrolizara en fructosa y glucosa.

20 **Ejemplo 10: Prototheca recombinantes con los genes TE exógenos**

[0124] Como se describió anteriormente, las cepas de *Prototheca* se pueden transformar con genes exógenos. La *Prototheca moriformis* (UTEX 1435) se transformó, usando los métodos que se describieron anteriormente, ya sea con el gene de la tioesterasa C12 de *Umbellularia californica* o con el gen de la tioesterasa C14 de *Cinnamomum camphora* (ambas con codones optimizados de acuerdo a la Tabla 4). Cada una de los constructos de la transformación contenían un promotor de glutamato deshidrogenasa de *Chlorella sorokiniana*/región 5'UTR (sec. con núm. de ident.: 61) para dirigir la expresión del transgén de la tioesterasa. Las regiones de codificación de los transgenes de las tioesterasas, la tioesterasa C12 de *Umbellularia californica* (sec. con num. de ident.: 62) o la tioesterasa C14 de *Cinnamomum camphora* (sec. con num. de ident.: 63), cada uno con la secuencia

dirigida a plastidio putativa nativa. Inmediatamente después de la secuencia codificante de la tioesterasa está la secuencia codificante para una etiqueta 3x FLAG en c-terminal (sec. con núm. de ident.: 64), seguida de la 3'UTR de la nitrato reductasa de *Chlorella vulgaris* (sec. con num. de ident.: 65).

5 [0125] La preparación del ADN, el microportador de oro, y las células *Prototheca moriformis* (UTEX 1435) se realizó usando de los métodos que se describieron anteriormente en el Ejemplo 9. Las microalgas se bombardearon con la mezcla de ADN con el microportador de oro y se sembraron en las placas de selección que contenían 2% de glucosa y 100µg/ml de G418. Las colonias se dejaron desarrollar 7 a 12 días y las colonias se
10 recogieron de cada placa de transformación para la detección de la incorporación del constructo de ADN mediante el uso de ensayos de transferencia del tipo Southern y la expresión de los constructos de tioesterasa se detectó mediante el uso de la RT-PCR.

[0126] Los clones positivos se recogieron tanto de las placas de transformación de la tioesterasa C12 y C14 y se detectó la incorporación del constructo mediante ensayos de transferencia de tipo Southern. Los ensayos de transferencia de tipo
15 Southern se llevaron a cabo mediante el uso de métodos estándar (y se describieron anteriormente en el Ejemplo 9), mediante el uso de una sonda c optimizada, que se basó en la secuencia sec. con núm. de ident.: 62 y sec. con núm. de ident.: 63. El ADN del plásmido de transformación se corrió como un control positivo. Fuera de los clones que resultaron
20 positivos para la incorporación del constructo, un subgrupo se seleccionó para el análisis por PCR cuantitativa con transcripción inversa (RT-PCR cuantitativa) para la expresión de la tioesterasa C12 y la tioesterasa C14.

[0127] El aislamiento del ARN se realizó usando los métodos que se describieron en el Ejemplo 9 anterior y se realizaron las RT-qPCR de los clones positivos mediante el uso de 20 ng de ARN total de cada clon, se usó el par de cebadores que se describen más abajo y el kit de RT-PCR iScript SYBR Green (Bio-Rad Laboratories) según el protocolo del fabricante. El ARN total de *Prototheca moriformis* de tipo salvaje (no transformado) se incluyó como control negativo. La expresión del ARNm se expresó como
25 las veces de expresión relativa (RFE) en comparación con el control negativo. Los cebadores que se usaron en las detecciones RT-qPCR de a C12 de transformación de la tioesterasa
30 fueron:

cebadores PCR de tioesterasa C12 de *U. californica*:

Directo: 5' CTGGGCGACGGCTTCGGCAC 3' (sec. con núm. ident.: 66)

Inverso: 5' AAGTCGCGGCGCATGCCGTT 3' (sec. con núm. ident.: 67)

[0128] Los cebadores que se usaron en la detección RT-qPCR de la transformación de tioesterasa C14 fueron:

5 cebadores de tioesterasa C14 de *Cinnamomum camphora*:

Directo: 5' TACCCCGCCTGGGGCGACAC 3' (sec. con núm. ident.: 68)

Inverso: 5' CTTGCTCAGGCGGCGGGTGC 3' (sec. con núm. ident.: 69)

[0129] La RT-PCR cuantitativa para la expresión de la tioesterasa C12 resultó en los clones positivos que mostraron un aumento de RFE de aproximadamente 40 veces a más de 2000 veces mayor expresión en comparación con el control negativo. Se observaron resultados similares con la expresión de la tioesterasa C14 en los clones positivos con un aumento de RFE de aproximadamente 60 veces a más de 1200 veces mayor expresión en comparación con el control negativo.

[0130] Se seleccionó un subconjunto de los clones positivos de cada transformación (como los que se detectaron en los ensayos de transferencia de tipo Southern y RT-PCR cuantitativa) y crecieron en condiciones abundantes de nitrógeno y se analizó la producción y el perfil de lípidos totales. Se prepararon muestras de los lípidos a partir de la biomasa seca de cada clon. Se resuspendieron 20-40 mg de biomasa seca de cada clon transgénico en 2 ml de H₂SO₄ 3% en MeOH, y se añadieron 200 µl de tolueno que contenía una cantidad apropiada de un patrón interno adecuado (C19:0). La mezcla se sometió a sonicación brevemente para dispersar la biomasa, entonces, se calentó a 65-70°C durante dos horas. Se añadieron 2 ml de heptano para extraer los ésteres metílicos de ácidos grasos, seguido por la adición de 2 ml de K₂CO₃ (ac) 6% para neutralizar el ácido. La mezcla se agitó vigorosamente, y una parte de la capa superior se transfirió a un frasco que contenía Na₂SO₄ (anhidro) para el análisis por cromatografía gaseosa mediante el uso del método estándar FAME GC/FID (cromatografía gaseosa de éster metílico de ácidos grasos con detección de ionización de llama). El perfil lipídico (se expresó como % del área) de los clones positivos en comparación con el control negativo de tipo salvaje se resume en las Tablas 36 y 37 más abajo. Como se muestra en la Tabla 36, las veces de incremento de la producción de C12 en los transformantes C12 estuvo en el intervalo de aproximadamente un aumento de 5 veces (clon C12-5) a un aumento de más de 11 veces (clon C12-1). Las veces

de incremento de la producción de C14 en los transformantes C14 estuvo en el intervalo de aproximadamente un aumento de 1.5 veces a aproximadamente un aumento de 2.5 veces.

[0131] Tabla 36. Sumario del perfil lipídico total de transformantes de tioesterasa C12 de *Prototheca moriformis*.

	tipo salvaje	C12-1	C12-2	C12-3	C12-4	C12-5	C12-6	C12-7	C12-8
C6:0	0.03	ND	ND	ND	ND	ND	ND	ND	ND
C8:0	0.11	0.09	ND	0.11	ND	ND	ND	ND	ND
C10:0	ND	ND	ND	0.01	0.01	ND	ND	0.01	ND
C12:0	0.09	1.04	0.27	0.72	0.71	0.50	0.67	0.61	0.92
C14:0	2.77	2.68	2.84	2.68	2.65	2.79	2.73	2.56	2.69
C14:1	0.01	ND	ND	0.02	ND	ND	ND	0.01	ND
C15:0	0.30	0.09	0.10	0.54	0.19	0.09	0.13	0.97	0.09
C15:1	0.05	ND	ND	0.02	ND	ND	ND	ND	ND
C16:0	24.13	23.12	24.06	22.91	22.85	23.61	23.14	21.90	23.18
C16:1	0.57	0.62	0.10	0.52	0.69	0.63	0.69	0.49	0.63
C17:0	0.47	0.24	0.27	1.02	0.36	0.17	0.26	2.21	0.19
C17:1	0.08	ND	0.09	0.27	0.10	0.05	0.09	0.80	0.05
C18:0	ND	ND	2.14	1.75	2.23	2.16	2.38	1.62	2.47
C18:1	22.10	23.15	24.61	21.90	23.52	19.30	22.95	20.22	22.85
C18:1	ND	0.33	0.24	ND	ND	0.09	0.09	ND	0.11
C18:2	37.16	34.71	35.29	35.44	35.24	36.29	35.54	36.01	35.31
C18:3 alfa	11.68	11.29	9.26	11.62	10.76	13.61	10.64	11.97	10.81
C20:0	0.15	0.16	0.19	0.16	0.16	0.14	0.18	0.14	0.18
C20:1	0.22	0.17	0.19	0.20	0.21	0.19	0.21	0.20	0.21
C20:2	0.05	ND	0.04	0.05	0.05	0.05	0.04	0.05	0.04
C22:0	ND	ND	ND	0.01	ND	ND	ND	0.02	ND
C22:1	ND	ND	ND	ND	ND	0.01	ND	0.01	ND
C20:3	0.05	ND	0.07	0.06	0.06	0.10	0.07	0.05	0.06
C20:4	ND	ND	ND	ND	ND	0.02	ND	ND	ND
C24:0	ND	ND	0.24	0.01	0.20	0.19	0.19	0.14	0.20

5

[0132] Tabla 37. Sumario del perfil lipídico total de transformantes de tioesterasa C14 de *Prototheca moriformis*.

	tipo salvaje	C14-1	C14-2	C14-3	C14-4	C14-5	C14-6	C14-7
C6:0	0.03	ND	ND	ND	ND	ND	ND	ND
C8:0	0.11	ND	ND	ND	ND	ND	ND	ND
C10:0	ND	0.01	ND	0.01	ND	0.01	ND	ND
C12:0	0.09	0.20	0.16	0.25	0.21	0.19	0.40	0.17
C14:0	2.77	4.31	4.76	4.94	4.66	4.30	6.75	4.02
C14:1	0.01	ND	0.01	ND	ND	0.01	ND	ND
C15:0	0.30	0.43	0.45	0.12	0.09	0.67	0.10	0.33
C15:1	0.05	ND	ND	ND	ND	ND	ND	ND
C16:0	24.13	22.85	23.20	23.83	23.84	23.48	24.04	23.34
C16:1	0.57	0.65	0.61	0.60	0.60	0.47	0.56	0.67
C17:0	0.47	0.77	0.76	0.21	0.19	1.11	0.18	0.54
C17:1	0.08	0.23	0.15	0.06	0.05	0.24	0.05	0.12
C18:0	ND	1.96	1.46	2.48	2.34	1.84	2.50	2.06
C18:1	22.10	22.25	19.92	22.36	20.57	19.50	20.63	22.03

C18:1	ND	ND	ND	ND	ND	ND	0.10	ND
C18:2	37.16	34.97	36.11	34.35	35.70	35.49	34.03	35.60
C18:3 alfa	11.68	10.71	12.00	10.15	11.03	12.08	9.98	10.47
C20:0	0.15	0.16	0.19	0.17	0.17	0.14	0.18	0.16
C20:1	0.22	0.20	0.12	0.19	0.19	0.19	0.17	0.20
C20:2	0.05	0.04	0.02	0.03	0.04	0.05	0.03	0.04
C22:0	ND	ND	ND	ND	0.02	0.01	ND	ND
C22:1	ND	0.01	ND	ND	ND	ND	ND	0.01
C20:3	0.05	0.08	0.03	0.06	0.09	0.05	0.05	0.07
C20:4	ND	0.01	ND	ND	ND	ND	0.02	ND
C24:0	ND	0.17	0.14	0.19	0.20	0.16	0.22	0.17

[0133] Los experimentos que se describieron indican la transformación exitosa de *Prototheca moriformis* (UTEX 1435) con los constructos de los transgenes de dos tioesterasas diferentes (C12 y C14), que implicaron no sólo la expresión exitosa del transgén, sino también la correcta dirección de la proteína expresada a los plastidios y un efecto funcional (el cambio esperado en el perfil de lípidos) como resultado de la transformación. El mismo experimento de transformación se realizó usando un constructo de expresión que contenía (según la Tabla 4) la región de codificación con codones optimizados de la tioesterasa C8-10 de *Cuphea hookeriana* con la secuencia dirigida a los plastidio, nativa, (sec. con num. de ident.: 70) que no produjo ningún cambio en el perfil de lípidos. Si bien la introducción del transgén C8-10 de *Cuphea hookeriana* en *Prototheca moriformis* (UTEX 1435) fue un éxito y se confirmó mediante análisis de transferencia de tipo Southern, ningún cambio en la producción de ácidos grasos C10 o C8 se detectó en los transformantes en comparación con la cepa de tipo salvaje.

Ejemplo 11: Generación de la cepa de *Prototheca moriformis* con TE de plantas exógenas con secuencias dirigidas a plastidios de algas

[0134] Con el fin de investigar si el uso de las secuencias dirigidas a los cloroplastos/plastidios de algas mejorarían la expresión de la tioesterasa de cadena mediana (C8-C14) y la posterior producción de lípidos de cadena mediana de *Prototheca moriformis* (UTEX 1435), varias supuestas secuencias dirigidas a plastidios de algas, se clonaron a partir de *Chlorella protothecoides* y *Prototheca moriformis*. los constructos de tioesterasa que se basaban en la tioesterasa C8-10 de *Cuphea hookeriana*, la tioesterasa C12 de *Umbellularia californica* y la tioesterasa C14 de *Cinnamomum camphora* se hicieron mediante el uso del promotor de la glutamato deshidrogenasa de *Chlorella sorokiniana*/5'UTR y una nitrato reductasa de *Chlorella vulgaris* 3'UTR. Las secuencias codificantes de la tioesterasa se modificaron mediante la eliminación de las secuencias dirigidas a plastidios, nativas y su

reemplazo con secuencias dirigidas a plastidios de los genomas de *Chlorella protothecoides* y de *Prototheca moriformis*. Las constructos de expresión de tioesterasa y sus correspondientes números de identificación de secuencia se enumeran más abajo. Cada plásmido de transformación contenía también un constructo de resistencia a Neo que era idéntico a los que se describieron en el Ejemplo 9 anterior. Además, otro promotor derivado de algas, el promotor de la β -tubulina de *Chlamydomonas reinhardtii*, se probó también conjuntamente con los constructos de tioesterasa. La secuencia dirigida a plastido, "nativa", se refiere a la secuencia dirigida a plastidio de tioesterasa de plantas superiores. Un resumen de los constructos usadas en estos experimentos se proporciona más abajo:

Constructo ident.. Nombre	Promotor/ 5'UTR	Plastidio sec. dirigida	Gen	3'UTR	sec. con núm.
Constructo 1	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 71
Constructo 2	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 72
Constructo 3	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> isopentenil difosfato sintasa	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 73
Constructo 4	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> isopentenil difosfato sintasa	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 74
Constructo 5	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 75
Constructo 6	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 76
Constructo 7	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Umbellularia</i> <i>californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 77
Constructo 8	<i>C. sorokiniana</i> glutamato deshidrogenasa	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Cinnamomum</i> <i>camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 78
Constructo 9	<i>Chlamydomonas</i> <i>reinhardtii</i> β -tubulina	Native	<i>Cuphea</i> <i>hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 79
Constructo 10	<i>Chlamydomonas</i> <i>reinhardtii</i>	<i>P. moriformis</i> isopentenil	<i>Cuphea</i> <i>hookeriana</i>	<i>C. vulgaris</i> nitrato	sec. con núm. ident.: 80

		β -tubulina	difosfato sintasa	C8-10 TE	reductasa	
5	Constructo 11	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 81
10	Constructo 12	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 82
15	Constructo 13	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Cuphea hookeriana</i> C8-10 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 83
20	Constructo 14	<i>Chlamydomonas reinhardtii</i> β -tubulina	Nativa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 84
25	Constructo 15	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> isopentenil difosfato	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 85
30	Constructo 16	<i>Chlamydomonas reinhardtii</i> β -tubulin	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 86
35	Constructo 17	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 87
40	Constructo 18	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Umbellularia californica</i> C12 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 88
45	Constructo 19	<i>Chlamydomonas reinhardtii</i> β -tubulina	Nativa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 89
50	Constructo 20	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> isopentenil difosfato sintasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 90
55	Constructo 21	<i>Chlamydomonas reinhardtii</i> β -tubulin	<i>P. moriformis</i> delta 12 graso ácido desaturasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 91
	Constructo 22	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>C. protothecoides</i> estearoil ACP desaturasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 92
	Constructo 23	<i>Chlamydomonas reinhardtii</i> β -tubulina	<i>P. moriformis</i> estearoil ACP desaturasa	<i>Cinnamomum camphora</i> C14 TE	<i>C. vulgaris</i> nitrato reductasa	sec. con núm. ident.: 93

[0135] Cada constructo se transformó en *Prototheca moriformis* (UTEX 1435) y la selección se realizó con G418 mediante el uso de los métodos que se describieron en el

Ejemplo 9 anterior. Varios clones positivos de cada transformación se recogieron y se examinaron en busca de la presencia de transgenes de la tioesterasa mediante el uso del análisis de transferencia de tipo Southern. La expresión del transgén de la tioesterasa se confirmó mediante la RT-PCR. Un subconjunto de los clones positivos (según se confirmó por el análisis de transferencia de tipo Southern y la RT-PCR) de cada transformación se seleccionó y se cultivó para el análisis del perfil lipídico. Las muestras de los lípidos se prepararon a partir de muestras de la biomasa seca de cada clon y el análisis del perfil lipídico se realizó mediante el uso de los métodos de hidrólisis ácida que se describieron en el Ejemplo 9. Los cambios en el porcentaje del área de los ácidos grasos correspondientes al transgén de la tioesterasa se compararon con los niveles del tipo salvaje, y con los clones transformados con una tioesterasa con la secuencia dirigida a plastidio, nativa.

[0136] Los clones transformados con el constructo de la tioesterasa C8-10 de *Cuphea hookeriana* con la secuencia dirigida a plastidio, nativa, tenían el mismo nivel de ácidos grasos C8 y C10 que el tipo salvaje. Los clones transformados con el constructo de la tioesterasa C8-10 de *Cuphea hookeriana* (Constructo 1-3) con secuencias dirigidas a plastidios de algas tenían un aumento de más de 10 veces en los ácidos grasos C10 para el Constructo 3 y un aumento de más de 40 veces en los ácidos grasos C10 de los constructos 1 y 2 (en comparación con el tipo salvaje). Los clones transformados con los constructos de la tioesterasa C12 de *Umbellularia californica* con la secuencia dirigida a los plastidio nativa, tuvo un aumento moderado de 6-8 veces en los niveles de ácidos grasos C12 en comparación con el tipo salvaje. Los clones transformados con los constructos de la tioesterasa C12 de *Umbellularia californica* con los constructos dirigidas a plastidios de algas (Constructos 4-7) tuvieron un aumento de más de 80 veces en el nivel del ácido graso C12 para el Constructo 4, aproximadamente un aumento de 20 veces en el nivel de ácido graso C12 para el Constructo 6, aproximadamente un aumento de 10 veces en el nivel de ácido graso C12 para el Constructo 7 y aproximadamente un aumento de 3 veces en el nivel de ácido graso C12 para el Constructo 5 (todos en comparación con tipo salvaje). Los clones transformados con la tioesterasa C14 de *Cinnamomum camphora*, ya sea con la secuencia dirigida a plastidio nativa o el Constructo 8 (con la secuencia dirigida a plastidio de la esteroil ACP desaturasa de *Chlorella protothecoides*) tuvieron un aumento de 2-3 veces en los niveles de ácidos grasos C14 en comparación con el tipo salvaje. En general, los clones transformados con los constructos de la tioesterasa con una secuencia dirigida a los plastidio de algas tuvieron un

aumento de veces más alto en los correspondientes niveles de longitud de cadena de los ácidos grasos, que cuando se usó la secuencia dirigida nativa de plantas superiores.

1. El promotor de la β-tubulina de *Clamydomonas reinhardtii*

[0137] La expresión adicional de los constructos de la tioesterasa heteróloga

se elaboraron mediante el uso del promotor de la β-tubulina de *Chlamydomonas reinhardtii* en lugar del promotor de la glutamato deshidrogenasa de *C. sorokinana*. Los elementos del constructo y la secuencia de los constructos de expresión se enumeraron anteriormente. Cada constructo se transformó en las células hospederas *Prototheca moriformis* UTEX 1435 mediante el uso de los métodos que se describieron anteriormente. Los perfiles de lípidos se generaron a partir de un subconjunto de los clones positivos para cada constructo con el fin de evaluar el éxito y la productividad de cada constructo. Los perfiles de lípidos comparan los niveles de ácidos grasos (expresados en % del área) con las células hospederas de tipo salvaje. La columna "media" representa el promedio numérico del subconjunto de los clones positivos. La columna "muestra" representa el mejor clon positivo que se examinó (el mejor se definió como la muestra que produce el mayor cambio en % del área de la producción del ácido graso de la correspondiente longitud de cadena). La columna "bajo-alto" representa el mas bajo % de área y el más alto % área de los ácidos grasos de los clones que se examinaron. Los resultados del perfil de lípido de los Constructos 9-12 y 14-23 se resumen más abajo.

Constructo 9. <i>Cuphea hookeriana</i> C8-10 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.05	0.30	0-0.29
C 10:0	0.01	0.63	2.19	0-2.19
C 12:0	0.03	0.06	0.10	0-0.10
C 14:0	1.40	1.50	1.41	1.36-3.59
C 16:0	24.01	24.96	24.20	
C 16:1	0.67	0.80	0.85	
C 17:0	0	0.16	0.16	
C 17:1	0	0.91	0	
C 18:0	4.15	17.52	3.19	
C 18:1	55.83	44.81	57.54	
C 18:2	10.14	7.58	8.83	
C 18:3α	0.93	0.68	0.76	
C 20:0	0.33	0.21	0.29	
C 24:0	0	0.05	0.11	

Constructo 10. <i>Cuphea hookeriana</i> C8-10 TE				
Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.01	0.02	0-0.03
C 10:0	0	0.16	0.35	0-0.35

C 12:0	0.04	0.05	0.07	0-0.07
C 14:0	1.13	1.62	1.81	0-0.05
C 14:1	0	0.04	0.04	
C 15:0	0.06	0.05	0.05	
C 16:0	19.94	26.42	28.08	
C 16:1	0.84	0.96	0.96	
C 17:0	0.19	0.14	0.13	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	3.62	3.43	
C 18:1	63.96	54.90	53.91	
C 18:2	9.62	9.83	9.11	
C18:3 γ	0	0.01	0	
C 18:3α	0.63	0.79	0.73	
C 20:0	0.26	0.35	0.33	
C 20:1	0.06	0.08	0.09	
C 20:1	0.08	0.06	0.07	
C 22:0	0	0.08	0.09	
C 24:0	0.13	0.13	0.11	

Construto 11. *Cuphea hookeriana* C8-10 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.82	1.57	0-1.87
C 10:0	0	3.86	6.76	0-6.76
C 12:0	0.04	0.13	0.20	0.03- 0.20
C 14:0	1.13	1.80	1.98	1.64- 2.05
C 14:1	0	0.04	0.04	
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	25.60	25.44	
C 16:1	0.84	1.01	1.02	
C 17:0	0.19	0.13	0.11	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	2.98	2.38	
C 18:1	63.96	51.59	48.85	
C 18:2	9.62	9.85	9.62	
C18:3 γ	0	0.01	0	
C 18:3α	0.63	0.91	0.92	
C 20:0	0.26	0.29	0.26	
C 20:1	0.06	0.06	0	
C 20:1	0.08	0.06	0.03	
C 22:0	0	0.08	0.08	
C 24:0	0.13	0.06	0	

Construto 12. *Cuphea hookeriana* C8-10 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 8:0	0	0.31	0.85	0-0.85
C 10:0	0	2.16	4.35	0.20- 4.35
C 12:0	0.04	0.10	0.15	0-0.18
C 14:0	1.13	1.96	1.82	1.66- 2.97
C 14:1	0	0.03	0.04	
C 15:0	0.06	0.07	0.07	
C 16:0	19.94	26.08	25.00	
C 16:1	0.84	1.04	0.88	
C 17:0	0.19	0.16	0.16	

C 17:1	0.10	0.05	0.07
C 18:0	2.68	3.02	3.19
C 18:1	63.96	51.08	52.15
C 18:2	9.62	11.44	9.47
C18:3 γ	0	0.01	0
C 18:3 α	0.63	0.98	0.90
C 20:0	0.26	0.30	0.28
C 20:1	0.06	0.06	0.05
C 20:1	0.08	0.04	0
C 22:0	0	0.07	0
C 24:0	0.13	0.05	0

Construto 14. *Umbellularia californica* C12 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.02	0.03	0.02- 0.03
C 12:0	0.03	2.62	3.91	0.04- 3.91
C 14:0	1.40	1.99	2.11	1.83- 2.19
C 16:0	24.01	27.64	27.01	
C 16:1	0.67	0.92	0.92	
C 18:0	4.15	2.99	2.87	
C 18:1	55.83	53.22	52.89	
C 18:2	10.14	8.68	8.41	
C 18:3 α	0.93	0.78	0.74	
C 20:0	0.33	0.29	0.27	

Construto 15. *Umbellularia californica* C12 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0	0.05	0.08	0-0.08
C 12:0	0.04	8.12	12.80	4.35- 12.80
C 13:0	0	0.02	0.03	0-0.03
C 14:0	1.13	2.67	3.02	2.18- 3.37
C 14:1	0	0.04	0.03	0.03- 0.10
C 15:0	0.06	0.07	0.06	
C 16:0	19.94	25.26	23.15	
C 16:1	0.84	0.99	0.86	
C 17:0	0.19	0.14	0.14	
C 17:1	0.10	0.05	0.05	
C 18:0	2.68	2.59	2.84	
C 18:1	63.96	46.91	44.93	
C 18:2	9.62	10.59	10.01	
C 18:3 α	0.63	0.92	0.83	
C 20:0	0.26	0.27	0.24	
C 20:1	0.06	0.06	0.06	
C 20:1	0.08	0.05	0.04	
C 22:0	0	0.07	0.09	
C 24:0	0.13	0.13	0.12	

Construto 16. *Umbellularia californica* C12 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
----------------	-----------------	----------	---------	-----------

C 10:0	0	0.03	0.04	0.02-0.04
C 12:0	0.04	2.43	5.32	0.98-5.32
C 13:0	0	0.01	0.02	0-0.02
C 14:0	1.13	1.77	1.93	1.62-1.93
C 14:1	0	0.03	0.02	0.02-0.04
C 15:0	0.06	0.06	0.05	
C 16:0	19.94	24.89	22.29	
C 16:1	0.84	0.91	0.82	
C 17:0	0.19	0.16	0.15	
C 17:1	0.10	0.06	0.06	
C 18:0	2.68	3.81	3.67	
C 18:1	63.96	53.19	52.82	
C 18:2	9.62	10.38	10.57	
C 18:3 α	0.63	0.80	0.77	
C 20:0	0.26	0.35	0.32	
C 20:1	0.06	0.06	0.07	
C 20:1	0.08	0.07	0.08	
C 22:0	0	0.08	0.07	
C 24:0	0.13	0.15	0.14	

Construeto 17. *Umbellularia californica* C12 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0	0.04	0.07	0.03-0.08
C 12:0	0.04	7.02	14.11	4.32-14.11
C 13:0	0	0.03	0.04	0.01-0.04
C 14:0	1.13	2.25	3.01	1.95-3.01
C 14:1	0 0.03	0.03	0.02-0.03	
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	23.20	21.46	
C 16:1	0.84	0.82	0.77	
C 17:0	0.19	0.15	0.14	
C 17:1	0.10	0.06	0.06	
C 18:0	2.68	3.47	2.93	
C 18:1	63.96	50.30	45.17	
C 18:2	9.62	10.33	9.98	
C 18:3 γ	0	0.01	0	
C 18:3 α	0.63	0.84	0.86	
C 20:0	0.26	0.32	0.27	
C 20:1	0.06	0.07	0.06	
C 20:1	0.08	0.06	0.06	
C 22:0	0	0.08	0.09	
C 24:0	0.13	0.14	0.13	

Construeto 18. *Umbellularia californica* C12 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0	0.03	0.05	0.01-0.05
C 12:0	0.04	5.06	7.77	0.37-

				7.77
C 13:0	0	0.02	0	0-0.03
C 14:0	1.13	2.11	2.39	1.82-2.39
C 14:1	0	0.03	0.03	0.02-0.05
C 15:0	0.06	0.06	0.06	
C 16:0	19.94	24.60	23.95	
C 16:1	0.84	0.86	0.83	
C 17:0	0.19	0.15	0.14	
C 17:1	0.10	0.06	0.05	
C 18:0	2.68	3.31	2.96	
C 18:1	63.96	51.26	49.70	
C 18:2	9.62	10.18	10.02	
C18:3 γ	0	0.01	0.02	
C 18:3α	0.63	0.86	0.86	
C 20:0	0.26	0.32	0.29	
C 20:1	0.06	0.05	0.05	
C 20:1	0.08	0.07	0.04	
C 22:0	0	0.08	0.08	
C 24:0	0.13	0.13	0.13	

Construeto 19. *Cinnamomum camphora* C14 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.02	0.01	0.01	0.01-0.02
C 12:0	0.05	0.27	0.40	0.08-0.41
C 14:0	1.52	4.47	5.81	2.10-5.81
C 16:0	25.16	28.14	28.55	
C 16:1	0.72	0.84	0.82	
C 18:0	3.70	3.17	2.87	
C 18:1	54.28	51.89	51.01	
C 18:2	12.24	9.36	8.62	
C 18:3α	0.87	0.74	0.75	
C 20:0	0.33	0.33	0.31	

Construeto 20. *Cinnamomum camphora* C14 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.01	0.02	0.01-0.02
C 12:0	0.03	0.39	0.65	0.08-0.65
C 13:0	0	0.01	0.01	0.01-0.02
C 14:0	1.40	5.61	8.4	2.1-8.4
C 14:1	0	0.03	0.03	0.02-0.03
C 15:0	0	0.06	0.07	
C 16:0	24.01	25.93	25.57	
C 16:1	0.67	0.75	0.71	
C 17:0	0	0.13	0.12	
C 17:1	0	0.05	0.05	
C 18:0	4.15	3.30	3.23	
C 18:1	55.83	51.00	48.48	
C 18:2	10.14	10.38	10.35	

C 18:3 α	0.93	0.91	0.88
C 20:0	0.33	0.35	0.32
C 20:1	0	0.08	0.08
C 20:1	0	0.07	0.07
C 22:0	0	0.08	0.08
C 24:0	0	0.14	0.13

Construto 21. *Cinnamomum camphora* C14 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.01	0.01	0-0.01
C 12:0	0.03	0.10	0.27	0.04- 0.27
C 14:0	1.40	2.28	4.40	1.47- 4.40
C 16:0	24.01	26.10	26.38	
C 16:1	0.67	0.79	0.73	
C 17:0	0	0.15	0.16	
C 17:1	0	0.06	0.06	
C 18:0	4.15	3.59	3.51	
C 18:1	55.83	53.53	50.86	
C 18:2	10.14	10.83	11.11	
C 18:3 α	0.93	0.97	0.87	
C 20:0	0.33	0.36	0.37	
C 20:1	0	0.09	0.08	
C 20:1	0	0.07	0.07	
C 22:0	0	0.09	0.09	

Construto 22. *Cinnamomum camphora* C14 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
C 10:0	0.01	0.02	0.02	0.02- 0.02
C 12:0	0.03	1.22	1.83	0.59- 1.83
C 13:0	0	0.02	0.03	0.01- 0.03
C 14:0	1.40	12.77	17.33	7.97- 17.33
C 14:1	0	0.02	0.02	0.02- 0.04
C 15:0	0	0.07	0.08	
C 16:0	24.01	24.79	24.22	
C 16:1	0.67	0.64	0.58	
C 17:0	0	0.11	0.10	
C 17:1	0	0.04	0.04	
C 18:0	4.15	2.85	2.75	
C 18:1	55.83	45.16	41.23	
C 18:2	10.14	9.96	9.65	
C 18:3 α	0.93	0.91	0.85	
C 20:0	0.33	0.30	0.30	
C 20:1	0	0.07	0.06	
C 20:1	0	0.06	0.05	
C 22:0	0	0.08	0.08	

Construto 23. *Cinnamomum camphora* C14 TE

Ácido graso	tipo salvaje	Promedio	Muestra	bajo/alto
----------------	-----------------	----------	---------	-----------

C 10:0	0.01	0.01	0.02	0-0.02
C 12:0	0.05	0.57	1.08	0.16- 1.08
C 13:0	0	0.02	0.02	0-0.02
C 14:0	1.45	7.18	11.24	2.96- 11.24
C 14:1	0.02	0.03	0.03	0.02- 0.03
C 15:0	0.06	0.07	0.07	
C 16:0	24.13	25.78	25.21	
C 16:1	0.77	0.72	0.66	
C 17:0	0.19 0.13	0.11		
C 17:1	0.08	0.05	0.04	
C 18:0	3.53	3.35	3.12	
C 18:1	56.15	49.65	46.35	
C 18:2	11.26	10.17	9.72	
C 18:3α	0.84	0.95	0.83	
C 20:0	0.32	0.34	0.32	
C 20:1	0.09	0.08	0.09	
C 20:1	0.07	0.05	0.06	
C 22:0	0.07	0.08	0.08	
C 24:0	0.13	0.13	0.12	

[0138] Los constructos 9-12 eran vectores de expresión que contenían el constructo de la tioesterasa C8-10 de *Cuphea hookeriana*. Como se puede ver en los resúmenes de los datos anteriores, los mejores resultados se vieron con el constructo 11, con la muestra de ácido graso C8 que fue 1.57 % del área (en comparación con el 0 en el tipo salvaje) y el ácido graso C10 que fue 6.76 % del área (en comparación con el 0 en el tipo salvaje). También hubo un aumento moderado de los ácidos grasos C12 (un aumento de aproximadamente 2-5 veces). Si bien la secuencia dirigida a los plastidios nativos, no produjo cambio bajo el control del promotor de la glutamato deshidrogenasa de *C.sorokinana*, el mismo constructo de expresión conducido por el promotor de la β -tubulina de *C.reinhardtii* produjo cambios significativos en los ácidos grasos C8-10 en la célula hospedera. Esta es una prueba más de la idiosincrasia de la expresión heteróloga de tioesterasas en especies de *Prototheca*. Todos los clones que contenían el constructo de la tioesterasa C8-10 con el promotor de la β -tubulina de *C.reinhardtii* tuvieron mayores aumentos en los ácidos grasos C8-10 que los clones que contenían el constructo de la tioesterasa C8-10 con el promotor de la glutamato deshidrogenasa de *C.sorokinana*. Los datos del perfil lipídico para el constructo 13 no se obtuvieron y, por tanto, no se incluyeron anteriormente.

[0139] Los constructos 14-18 eran vectores de expresión que contenían el constructo de la tioesterasa C12 de *Umbellularia californica*. Como puede verse en los resúmenes de los datos anteriores, los mejores resultados se vieron con el Constructo 15 (la

secuencia dirigida a plastidio de la isopentenil difosfato sintasa de *P.moriformis*) y la 17 (la secuencia dirigida a plastidio de la estearoil ACP desaturasa de *C.protothecoides*). El mayor cambio en la producción del ácido graso C12 se vio con el constructo 17, con niveles de ácidos grasos C12 de hasta 14.11% del área, en comparación con el 0.04% del área en el tipo salvaje. También se observaron cambios moderados (aproximadamente de 2 veces) con los niveles de ácidos grasos C14. En comparación con el mismo constructo con el promotor de la glutamato deshidrogenasa de la *C.sorokinana*, las mismas tendencias fueron ciertas con el promotor de la β -tubulina de *C.reinhardtii* – las secuencias dirigidas a plastidios de, la estearoil ACP desaturasa de *C.protothecoides* y la isopentenil difosfato sintasa de *P.moriformis* produjeron el mayor cambio en los niveles de ácidos grasos C12 con ambos promotores.

[0140] Los constructos 19-23 eran vectores de expresión que contenían el constructo de la tioesterasa C14 de *Cinnamomum camphora*. Como puede verse en los resúmenes de los datos anteriores, los mejores resultados se vieron con el Constructo 22 y el Constructo 23. El mayor cambio en la producción del ácido graso C14 se vio con el Constructo 22 y con niveles del ácido graso C14 de hasta el 17.35 % del área (cuando los valores de C140 y C141 se combinan), en comparación con el 1.40 % en el tipo salvaje. Los cambios en los ácidos grasos C12 también se observaron (5-60 veces). En comparación con el mismo constructo con el promotor de la glutamato deshidrogenasa de *C.sorokinana*, las mismas tendencias fueron ciertas con el promotor de la β -tubulina –las secuencias dirigidas a plastidios de la estearoil ACP desaturasa de *C.protothecoides* y la estearoil ACP desaturasa de *P.moriformis* produjeron el mayor cambio en los niveles de ácidos grasos C14 con ambos promotores. En consonancia con todas las constructos de expresión de tioesterasa, el constructo con el promotor de la β -tubulina de *C.reinhardtii* produjo los mayores cambios en los niveles de ácidos grasos C8-14 que el de la glutamato deshidrogenasa de *C.sorokiniana*.

[0141] Dos clones positivos del constructo de 22 se seleccionaron y se cultivaron a alta presión selectiva (50 mg/l de G418). Después de 6 días de cultivo, los clones se cosecharon y se determinó el perfil lipídico mediante el uso de los métodos que se describieron anteriormente. Los datos del perfil lipídico se resumen más abajo y se expresaron en % del área.

Constructo 22 clones + 50mg/l G418		
Ácido	Constructo	Constructo
graso	22 A	22 B
C 12:0	3.21	3.37

C 14:0	27.55	26.99
C 16:0	25.68	24.37
C 16:1	0.99	0.92
C 18:0	1.37	1.23
C 18:1	28.35	31.07
C 18:2	11.73	11.05
C 18:3α	0.92	0.81
C 20:0	0.16	0.17

[0142] Ambos clones, cuando se cultivaron a presión constante, selectiva alta, produjeron un aumento en la cantidad de los ácidos grasos C14 y C12, aproximadamente el doble de los niveles que se observaron con el Constructo 22 anterior. Estos clones produjeron 5 mas de 30% del área de los ácidos grasos C12-14, en comparación con el 1.5 % del área de los ácidos grasos C12-14 que se vio en las células de tipo salvaje.

REIVINDICACIONES:

1. Un método para extraer un lípido de un microorganismo que comprende las etapas de:

(a) lisar un microorganismo cultivado para producir un lisado, en donde el microorganismo:

- (i) no se ha sometido a una etapa de secado entre el cultivo y la lisis; y
- (ii) contiene un lípido;

(b) tratar el lisado con un solvente orgánico durante un período de tiempo suficiente para permitir a los lípidos de los microorganismos solubilizarse en el solvente orgánico; y

(c) separar el lisado en capas que comprenden una capa lípido:solvente orgánico y una capa acuosa;

a través de lo cual el lípido se extrae del microorganismo.

2. El método de la reivindicación 1, en donde la etapa de tratamiento comprende tratar el lisado con más de aproximadamente 5% v/v de un solvente orgánico.

3. El método de la reivindicación 1, en donde las capas comprenden, además, una capa de lípido:emulsión acuosa y/o un sedimento celular.

4. El método de la reivindicación 1, en donde la etapa (a) se realiza a un pH ácido.

5. El método de la reivindicación 1, en donde la etapa (a) comprende uno o más métodos para lisar el microorganismo seleccionados de la lista consistente de calentamiento, sonicación, lisis mecánica, choque osmótico, expresión de un gen de autólisis, exposición a un pH por encima de 8, exposición a pH por debajo de 6, y digestión con una enzima.

6. El método de la reivindicación 1, en donde el microorganismo produce un lípido que comprende al menos 25% p/v del lípido celular total como C18 triacilglicérol.

7. El método de la reivindicación 1, en donde el microorganismo produce un lípido que comprende al menos 25% p/v del lípido celular total como C16 triacilglicérol.

8. El método de la reivindicación 1, en donde el microorganismo produce un lípido que comprende al menos 10% p/v del lípido celular total como C14 triacilglicérol.
9. El método de la reivindicación 1, en donde el microorganismo produce un lípido que comprende al menos 10% p/v del lípido celular total como C12 triacilglicérol.
10. El método de la reivindicación 1, en donde el microorganismo produce un lípido que comprende al menos 10% p/v del lípido celular total como C10 triacilglicérol.
11. El método de la reivindicación 1, en donde el microorganismo produce un lípido con un perfil de lípido que comprende al menos 4% de C8-C14.
12. El método de la reivindicación 1, en donde el microorganismo produce un lípido con un perfil de lípido que comprende entre 4% y 40% de C10-C14.
13. El método de la reivindicación 1, en donde la etapa (a) comprende lisis ácida y calentamiento.
14. El método de la reivindicación 1, en donde el solvente orgánico comprende un aceite seleccionado del grupo que consiste en aceite de soja, colza, canola, palma, almendra de palma, coco, maíz, residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de buey, sebo porcino, microalgas, macroalgas, *Cuphea*, lino, cacahuate, grasa blanca selecta, manteca de cerdo, *Camelina sativa*, semilla de mostaza, nuez de marañón, avena, lupino, kenaf, caléndula, cáñamo, café, semilla de lino, avellana, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez, jojoba, jatrofa, macadamia, nueces del Brasil, aguacate, petróleo, o una fracción destilada de cualquiera de los aceites anteriores.
15. El método de la reivindicación 1, en donde el solvente orgánico es un surfactante.
16. El método de la reivindicación 1, en donde la etapa (c) comprende reducir la temperatura de la mezcla por debajo de 25°C.

17. El método de la reivindicación 1, en donde el microorganismo se selecciona del grupo de microorganismos que consiste en una bacteria, una cianobacteria, una microalga eucariota, una levadura oleaginosa, y un hongo.

18. El método de la reivindicación 17, en donde el microorganismo es del género *Chlorella* o *Prototheca*.

19. El método de la reivindicación 1, que comprende, además, la etapa de:
(d) eliminar la composición lípido:solvente orgánico de la(s) otra(s) capas(s).

20. El método de la reivindicación 19, en donde la etapa (d) se realiza sin separar sustancialmente el lípido de la composición de solvente orgánico.

21. Un método para producir combustible que comprende las etapas de:

(a) lisar un microorganismo que contiene lípido para producir un lisado;

(b) tratar el lisado con un solvente orgánico durante un período de tiempo suficiente para permitir a los lípidos de los microorganismos solubilizarse en el solvente orgánico;

(c) separar el lisado en una composición lípido:solvente orgánico y una composición acuosa y, opcionalmente, una composición emulsionada y/o composición de sedimento celular;

(d) eliminar la composición lípido:solvente orgánico de la composición acuosa, composición de la emulsión, o composición de sedimento celular; y

(e) someter la composición lípido:solvente orgánico a una reacción química, a través de lo cual se produce un combustible.

22. El método de la reivindicación 21, en donde la reacción química es transesterificación y el combustible es biodiesel.

23. El método de la reivindicación 22, en donde el biodiesel cumple o excede la norma de biodiesel ASTM D6751.

24. El método de la reivindicación 22, en donde el biodiesel cumple o excede la norma de biodiesel EN 14214.

25. El método de la reivindicación 21, en donde la reacción química comprende una o más reacciones seleccionadas del grupo que consiste en hidroprocesamiento, desoxigenación, e isomerización y el combustible es diesel renovable.

26. El método de la reivindicación 25, en donde el diesel renovable cumple o excede la norma ASTM D 975.

27. El método de la reivindicación 21, en donde la reacción química comprende hidrotratamiento y el combustible es combustible de reactor.

28. El método de la reivindicación 21, en donde el microorganismo no se ha sometido a una etapa de secado.

29. El método de la reivindicación 21, en donde el solvente orgánico comprende un aceite seleccionado del grupo que consiste en aceite de soja, colza, canola, palma, almendra de palma, coco, maíz, residuos vegetales, sebo de China, oliva, girasol, semilla de algodón, grasa de pollo, sebo de buey, sebo porcino, microalgas, macroalgas, *Cuphea*, flax, peanut, grasa blanca selecta, manteca de cerdo, *Camelina sativa*, semilla de mostaza, nuez de marañón, avena, lupino, kenaf, caléndula, cáñamo, café, semilla de lino, avellana, euforbia, semillas de calabaza, cilantro, camelia, ajonjolí, cártamo, arroz, árbol del tung, cacao, copra, amapola del opio, semillas de ricino, nuez, jojoba, jatrofa, macadamia, nueces del Brasil, aguacate, petróleo o una fracción destilada de cualquiera de los aceites anteriores.

30. El método de la reivindicación 21, en donde el solvente orgánico es un surfactante.

31. El método de la reivindicación 21, en donde el microorganismo se selecciona del grupo de microorganismos que consiste en una bacteria, una cianobacteria, una microalga eucariota, una levadura oleaginosa, y un hongo.

32. El método de la reivindicación 31, en donde el microorganismo es del género *Chlorella* o *Prototheca*.

33. El método de la reivindicación 21, en donde el microorganismo produce un lípido con un perfil de lípido que comprende al menos 4% C8-C14.

34. El método de la reivindicación 21, en donde el microorganismo produce un lípido con un perfil de lípido que comprende entre 4% y 40% de C10-C14.

35. El método de la reivindicación 21, en donde la etapa (e) se realiza sin separar sustancialmente el lípido del solvente orgánico.

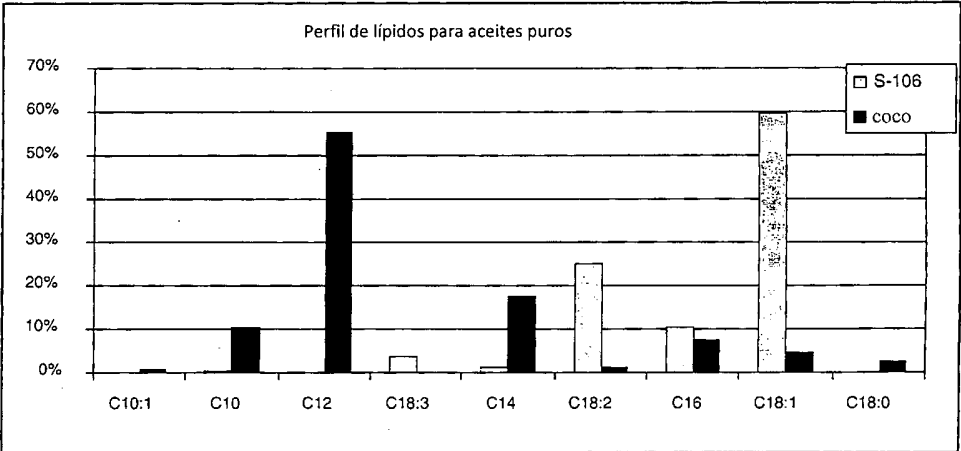
36. El método de la reivindicación 21, en donde el microorganismo tiene al menos 85% de identidad de secuencia genómica del ARNr 23s con una o más secuencias seleccionadas del grupo que consiste en sec. con núm. de ident.:7, sec. con núm. de ident.:8, sec. con núm. de ident.:9, sec. con núm. de ident.:10, sec. con núm. de ident.:11, sec. con núm. de ident.:12, sec. con núm. de ident.:13, sec. con núm. de ident.:14, sec. con núm. de ident.:15, sec. con núm. de ident.:16, sec. con núm. de ident.:17, sec. con núm. de ident.:18, sec. con núm. de ident.:19, sec. con núm. de ident.:20, sec. con núm. de ident.:21, sec. con núm. de ident.:22, sec. con núm. de ident.:23, sec. con núm. de ident.:24, sec. con núm. de ident.:25, sec. con núm. de ident.:26, sec. con núm. de ident.:27, sec. con núm. de ident.:28, sec. con núm. de ident.:29, sec. con núm. de ident.:30, sec. con núm. de ident.:31, sec. con núm. de ident.:32, sec. con núm. de ident.:33, sec. con núm. de ident.:34, sec. con núm. de ident.:35, sec. con núm. de ident.:36, sec. con núm. de ident.:37, sec. con núm. de ident.:38, y sec. con núm. de ident.:39.

RESUMEN DE LA INVENCION

La invención se refiere generalmente a la producción de composiciones de hidrocarburo, tales como un lípido, en microorganismos. En particular, la invención proporciona métodos para la extracción, recuperación, aislamiento y obtención de un lípido a partir de un microorganismo y composiciones que comprenden el lípido. La invención describe, además, métodos para producir composiciones de hidrocarburo para usar como biodiesel, diesel renovable, combustible de reactor, y otros materiales.

FIGURA 1

1/8



h/E

FIGURA 2

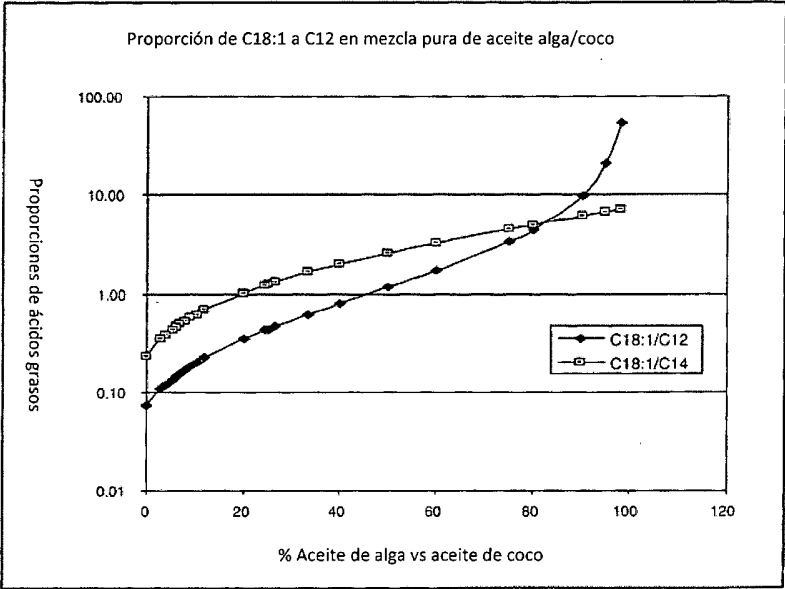


FIGURA 3

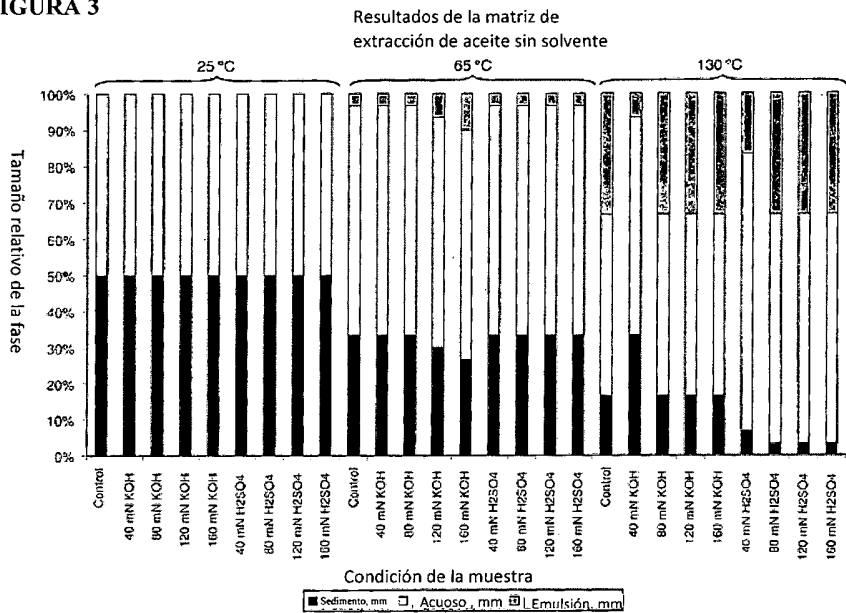


FIGURA 4

418

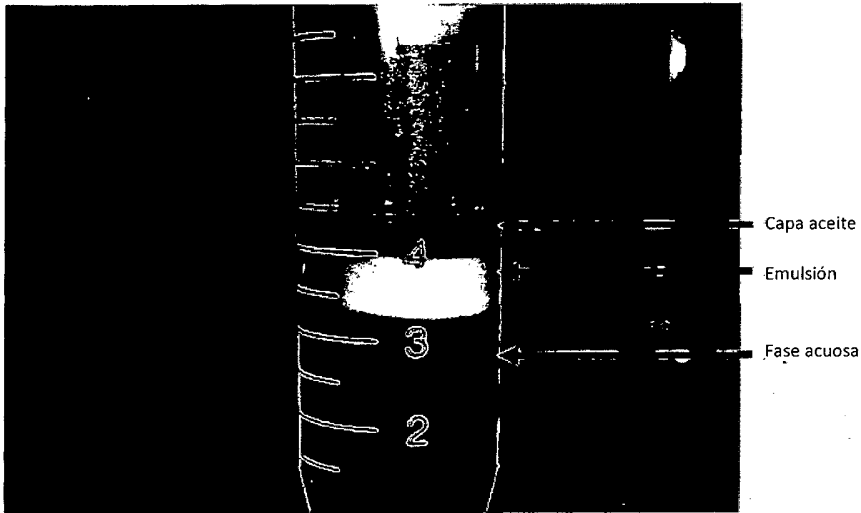


FIGURA 5

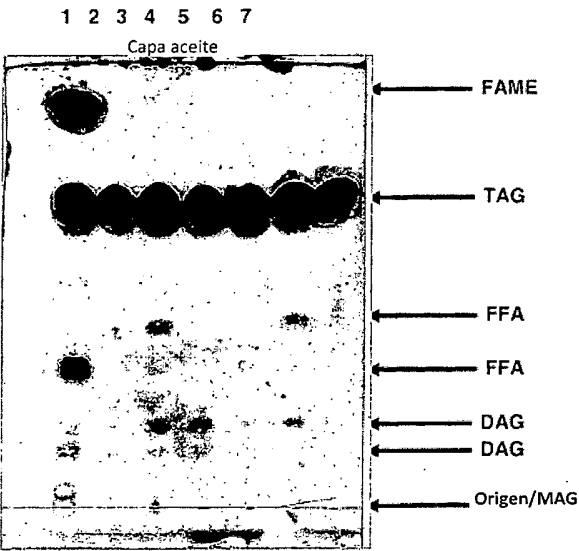


FIGURA 6

8/9

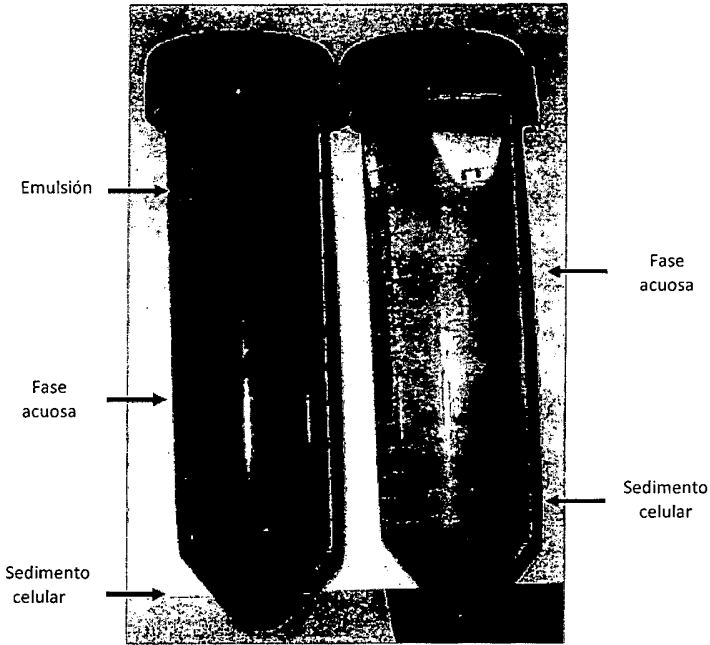
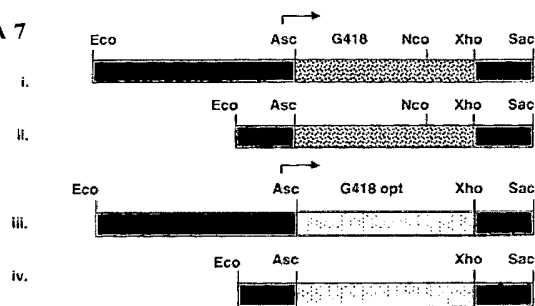
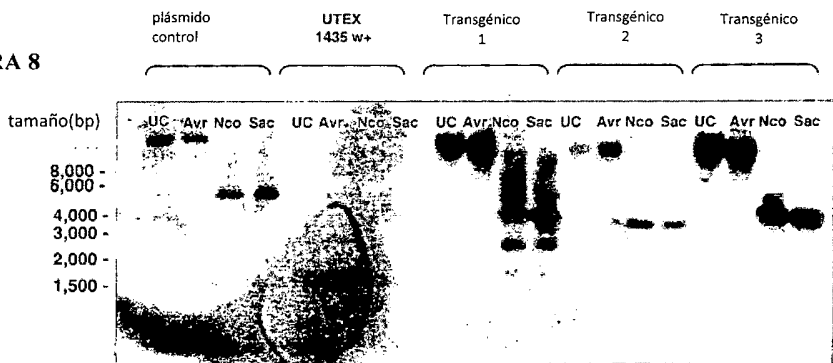


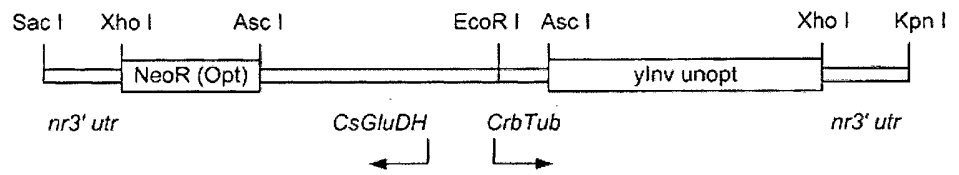
FIGURA 7



7/8

FIGURA 8





8/8

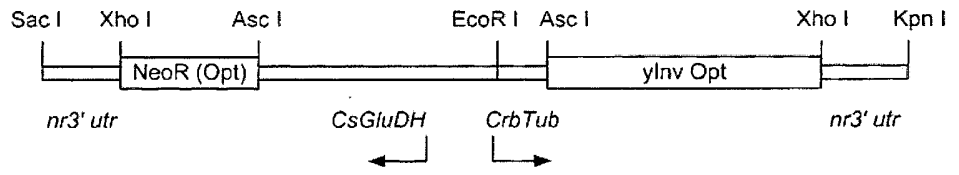


FIGURA 9

32378198_1.TXT
SEQUENCE LISTING

<110> SOLAZYME, INC.
<120> FRACTIONATION OF OIL-BEARING MICROBIAL BIOMASS
<130> 026172-002630PC
<140> PCT/US2010/036238
<141> 2010-05-26
<150> 61/181,252
<151> 2009-05-26
<160> 96
<170> PatentIn version 3.5
<210> 1
<211> 1187
<212> DNA
<213> Chlorella sp.
<400> 1
gatcagacgg gcctgacctg cgagataatc aagtgtctgt aggcaaccaa ctcagcagct 60
gcttgggtgtt ggggtctgcag gatagtgttg cagggcccca aggacagcag gggaacttac 120
accttgtccc cgacctcagt ttatggagtg cattgcctca agagcctagc cggagcgcta 180
ggctacatac ttgccgcacc ggtatgaggg gatatagtagc tcgcactgcg ctgtctagtgt 240
agatggggcag tgctgccccat aaacaactgg ctgtctagcc atttgttggc ggaccattct 300
ggggggggcca gcaatgcctg actttcgggt aggggtgaaa ctgaacaaag actacaaaa 360
cagaatttct tcctccttgg aggttaagcgc aggccggccc gcctgcgccc acatggcgct 420
ccgaacacct ccatagctgt aagggcgcaa acatggccgg actgttgtca gcactcttc 480
atggccatac aaggtcatgt cgagattagt gctgagtaag acactatcac cccatgttcg 540
attgaagccg tgacttcatg ccaacctgcc cctgggcgta gcagacgtat gccatcatga 600
ccactagccg acatgcgctg tcttttgcca ccaaaacaac tggtagaccg ctcgaagtcg 660
tgccgcacac ctccgggagt gagtccggcg actcctcccc ggccggccgc ggccctacct 720
gggtagggtc gccatacgcc cagcacaaa cgacgcagga ggggattggg gtaggggaatc 780
ccaaccagcc taaccaagac ggcacctata ataatagggt gggggactaa cagccctata 840
tcgcaagctt tgggtgccta tcttgagaag cagcagttgg agtggctgtg tacggtcgac 900
cctaagggtg gtgtgccga gcctgaaaca aagcgtctag cagctgcttc tataatgtgt 960
cagccgttgt gtttcagtta tattgtatgc tattgtttgt tcgtgctagg gtggcgagg 1020
cccacctact gtggcgggccc attggttggg gcttgaattg cctcaccatc taaggctctga 1080
acgctcactc aaacgccttt gtacaactgc agaactttcc ttggcgctgc aactacagtg 1140
tgcaaaccag cacatagcac tcccttacat caccagcag tacaaca 1187

<210> 2
<211> 1414
<212> DNA
<213> Chlorella ellipsoidea
<400> 2
cgctgcgcac cagggccgcc agctcgctga tgctgctcca aatgcgggtcc cccgattttt 60
tggtcttcat cttctccacc ttggtggcct tcttgccag ggccttcagc tgcattgcga 120

32378198_1.TXT

cagaccgttg agctcctgat cagcatcctc aggaggccct ttgacaagca agccccctgtg	180
caagcccatt cacgggggtac cagtgggtgct gaggtagatg ggtttgaaaa ggattgctcg	240
gtcgattgct gctcatggaa ttggcatgtg catgcatgtt cacaatatgc caccaggctt	300
tggagcaaga gagcatgaat gccttcaggc aggttgaaaag ttcctggggg tgaagaggca	360
gggccgagga ttggaggagg aaagcatcaa gtcgtcgtc atgctcatgt tttcagtcag	420
agtttgccaa gctcacagga gcagagacaa gactggctgc tcagggtgtg catcgtgtgt	480
gtggtggggg gggggggggt aatacggtag gaaatgcact tggaattccc acctcatgcc	540
agcggaccca catgcttgaa ttcgaggcct gtgggggtgag aaatgctcac tctgccctcg	600
ttgctgaggt acttcaggcc gctgagctca aagtcgatgc cctgctcgtc tatcagggcc	660
tgcacctctg ggctgaccgg ctcagcctcc ttcgcgggca tggagtaggc gccggcagcg	720
ttcatgtccg ggcccagggc agcggtggtg ccataaatgt cggtgatggt ggggaggggg	780
gccgtcgcca caccattgcc gttgctggct gacgcatgca catgtggcct ggctggcacc	840
ggcagcactg gtctccagcc agccagcaag tggctgttca ggaaagcggc catgttgttg	900
gtccctgctc atgtaattcc ccagatcaaa ggagggaaaca gcttggattt gatgtagtgc	960
ccaaccggac tgaatgtgcg atggcaggtc cctttgagtc tcccgaatta ctagcagggc	1020
actgtgacct aacgcagcat gccaacgcga aaaaaatgat tgacagaaaa tgaagcgggtg	1080
tgtcaatatt tgctgtatth attcgthth atcagcaacc aagttcgaaa cgcaactatc	1140
gtggtgatca agtgaacctc atcagactta cctcgthcg caaggaaacg gaggcaccaa	1200
attccaatth gatattatcg cthgccaagc tagagctgat cthtgggaaa ccaactgcca	1260
gacagtggac tgtgatggag tgccccgagt ggtggagcct cthcgattcg gttagtcatt	1320
actaacgtga accctcagtg aagggaccat cagaccagaa agaccagatc tcctcctcga	1380
caccgagaga gtgttgccgc agtaggacga caag	1414

<210> 3
 <211> 1113
 <212> DNA
 <213> Botryococcus braunii

<400> 3	
aattggaaac cccgcgcaag accgggttgt ttggccgcct gaccggaaaag ggggggcctg	60
tcccgaaggg ggtctatctc ttgggggatg tcgggcgcgg aaagtcgatg ttgatggacc	120
tcttcttcga ccatgtcggg gtcgaggcca agagccgcgt ccatttcgcc gagttcatga	180
tggaggtgaa tgaccgcac gccaccgaac gcgccaagaa gcgggcgacc gatcgcccc	240
gtcgtgcag cccttgccga ggaagtccgg ctgctggcgt tcgacgagat gatggtgacg	300
aacagcccgg acgcgatgat cctgtcgcgg ctgttcaccg cgctgatcga ggcgggggtg	360
acgatcgtca ccacctcaa ccggccgcc agggatctct ataagaacgg gctcaaccgc	420
gagcatttcc tgcccttcat cgcgtgatc gaggcgcggc tggacgtgct ggcgtgaac	480
ggccccgaccg actatcggcg cgaccggctg gggcggtg acacgtgggt ggtgcccaat	540
ggcccaagg cgacgattac cthgtcggcg gcgttcttcc gcctgaccga ctatccggtc	600
gaggatgccg cgcatgtgcc ctctgaggac ctgaagggtg gcgggcgcgt gctgaatgtc	660
cccaaggcgc tgaaggcgct cgcggtcttc tcgttcaagc ggttggtgcg cgaagcgcgg	720

32378198_1.TXT
ggggcgggcgg actatctggc ggtcgcgcgg ggcttccaca ccgtcatcct ggtcgggaatc 780
cccaagctgg gggcggagaa ccgcaacgag gcggggcgct tcgtccagct gatcgacgcg 840
ctctacgaac ataagggtcaa gctgctcgcc gcagccgatg ccagcccgcc gaactctatg 900
aaaccggcga cggccgggttc gagtttgagc gcagatcagc cggttggaag agatgcgctc 960
cgaggattat ctggcccaag gccatggctc ggaggggcct tgatcaggcc ttaatgcact 1020
tcgcaaccat tatcgtttaa aatcttaaac tctgtggaat aacggttccc cgacgccgca 1080
atacacgtac gtccactacg gagtaggatt gga 1113

<210> 4
<211> 253
<212> DNA
<213> Chlamydomonas reinhardtii

<400> 4
cgcttagaag atttcgataa ggcgccagaa ggagcgcagc caaaccagga tgatgtttga 60
tggggtattt gagcacttgc aacccttata cggaagcccc ctggcccaca aaggctaggc 120
gccaatgcaa gcagttcgca tgcagcccct ggagcgggtgc cctcctgata aaccggccag 180
ggggcctatg ttctttactt ttttacaaga gaagtcactc aacatcttaa acggtcttaa 240
gaagtctatc cgg 253

<210> 5
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 5
tgttgaagaa tgagccggcg ac 22

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 6
cagtgaagta ttacgcactc 20

<210> 7
<211> 565
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 7
tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60
cttagcgaaa gcgagtctga atagggcgat caaatatttt aatatttaca atttagtcat 120
tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg 240

32378198_1.TXT

aaataccagt cgaacccgga gctagctggt tctccccgaa atgcggtgag gcgcagcagt	300
acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaaa acggtaccaa	360
atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag	420
ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta	480
gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggtt ggcttagaag cagccatcct	540
ttaaagagtg cgtaatagct cactg	565

<210> 8

<211> 546

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 8

tgttgaagaa tgagccggcg acttagaaaa cgtggcaagg ttaaggaaac gtatccggag	60
ccgaagcgaa agcaagtctg aacagggcga ttaagtcatt ttttctagac ccgaacccgg	120
gtgatctaac catgaccagg atgaagcttg ggtgacacca agtgaaggtc cgaaccgacc	180
gatgttgaaa aatcggcgga tgagttgtgg ttagcgggtga aataccagtc gaactcggag	240
ctagctggtt ctccccgaaa tgcggtgagg cgcagcgggtt cataaggctg tctaggggta	300
aagcactgtt tcggtgcggg ctgcgaaagc ggtaccaa atgcggtgaa tctgaatact	360
agatatgcta tttatggggc agtgagacgg tgggggataa gcttcacgtg cgagagggaa	420
acagcccaga tcactagcta agggccccaa atgatcgta agtgacaaag gaggtgagaa	480
tgcagaaaca accaggaggt ttgcttagaa gcagccaccc ttaaagagtg gcgtaatagc	540
tcactg	546

<210> 9

<211> 565

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 9

tgttgaagaa tgagccggcg acttagaaaa agtggcggtg ttaaggaaaa attccgaagc	60
cttagcgaaa gcgagtctga atagggcgat caaatatattt aatatttaca atttagtcat	120
tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc	180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg	240
aaataccagt cgaacccgga gctagctggt tctccccgaa atgcggtgag gcgcagcagt	300
acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaaa acggtaccaa	360
atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag	420
ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta	480
gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggtt ggcttagaag cagccatcct	540
ttaaagagtg cgtaatagct cactg	565

32378198_1.TXT

<210> 10
<211> 565
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 10
tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60
cttagcgaaa gcgagtctga atagggcgat caaatatttt aatatttaca atttagtcat 120
tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg 240
aaataccagt cgaacccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300
acatctagtc tatctagggg taaagcactg tttcggtgcg ggctgtgaaa acggtaccaa 360
atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420
ctccattgtc aagaggggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480
gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggtt ggcttagaag cagccatcct 540
ttaaagagtg cgtaatagct cactg 565

<210> 11
<211> 548
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 11
tgttgaagaa tgagccggcg acttagaaga agtggccttg ttaaggataa ctatccggag 60
ccagagcgaa agcaagtctg aatagggcgc ttaaagggtca ctttttctag acccgaaccc 120
gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga 180
ccgatgttga aaaatcggcg gatgagttgt ggtagcgggt gaaataccaa tcgaactcgg 240
agctagctgg ttctccccga aatgcgttga ggcgagcggg tttatgaggc tgtctagggg 300
taaagcactg tttcggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360
ctagatatgc tattcatgag ccagtgagac ggtgggggat aagcttcac gtcaagaggg 420
aaacagccca gatcaccagc taaggcccca aaatggtcgt taagtggcaa aggaggtgag 480
aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540
gctcactg 548

<210> 12
<211> 548
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 12
tgttgaagaa tgagccggcg acttagaaga agtggccttg ttaaggataa ctatccggag 60

32378198_1.TXT

```

ccagagcgaa agcaagtctg aatagggcgc ttaaagggtca ctttttctag acccgaaccc 120
gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga 180
ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240
agctagctgg ttctccccga aatgcgttga ggcgcagcgg tttatgaggc tgtctagggg 300
taaagcactg tttcgggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360
ctagatatgc tattcatgag ccagtgaagc ggtgggggat aagcttcac gtcaagaggg 420
aaacagccca gatcaccagc taaggcccca aaatggtcgt taagtggcaa aggaggtgag 480
aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540
gctcactg 548

```

<210> 13

<211> 548

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 13

```

tggtgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag 60
ccagagcgaa agcaagtctg aatagggcgc ttaaagggtca ctttttctag acccgaaccc 120
gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga 180
ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240
agctagctgg ttctccccga aatgcgttga ggcgcagcgg tttatgaggc tgtctagggg 300
taaagcactg tttcgggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360
ctagatatgc tattcatgag ccagtgaagc ggtgggggat aagcttcac gtcaagaggg 420
aaacagccca gatcaccagc taaggcccca aaatggtcgt taagtggcaa aggaggtgag 480
aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540
gctcactg 548

```

<210> 14

<211> 548

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 14

```

tggtgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag 60
ccagagcgaa agcaagtctg aatagggcgc ttaaagggtca ctttttctag acccgaaccc 120
gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga 180
ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240
agctagctgg ttctccccga aatgcgttga ggcgcagcgg tttatgaggc tgtctagggg 300
taaagcactg tttcgggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360
ctagatatgc tattcatgag ccagtgaagc ggtgggggat aagcttcac gtcaagaggg 420
aaacagccca gatcaccagc taaggcccca aaatggtcgt taagtggcaa aggaggtgag 480

```

32378198_1.TXT

aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540
gctcactg 548

<210> 15
<211> 565
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 15
tggtgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60
cttagcgaaa gcgagtctga atagggcgat caaatatttt aatatttaca atttagtcac 120
tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg 240
aaataccagt cgaacccgga gctagctggg tctccccgaa atgcgttgag gcgcagcagt 300
acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaaa acggtaccaa 360
atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420
ctccattgtc aagaggggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480
gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggtt ggcttagaag cagccatcct 540
ttaaagagtg cgtaatagct cactg 565

<210> 16
<211> 541
<212> DNA
<213> Prototheca stagnora

<400> 16
tggtgaagaa tgagccggcg agttaaaaa aatggcatgg ttaaagatat ttctctgaag 60
ccatagcgaa agcaagtttt acaagctata gtcatttttt ttagaccgga aaccgagtga 120
tctacccatg atcaggggtga agtgttgggtc aaataacatg gaggcccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt ttgtgggctt cgaaagtgtt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagttag accttggggg ataagctcct tgggtcaaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggacgt gagtatgtca 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

<210> 17
<211> 573
<212> DNA
<213> Prototheca moriformis

<400> 17
tggtgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga 60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct 120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180

32378198_1.TXT

gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt	240
tagtggtgaa ataccagtcg aactcagagc tagctgggttc tccccgaaat gcgttgaggc	300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat	360
ggtaccaa at cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg	420
gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg	480
ataatgaagt ggtaaaggag gtgaaaaatgc aaatacaacc aggaggttgg cttagaagca	540
gccatccttt aaagagtgcg taatagctca ctg	573

<210> 18
<211> 573
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 18	
tggtgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga	60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaacatta aataaaatct	120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg	180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt	240
tagtggtgaa ataccagtcg aactcagagc tagctgggttc tccccgaaat gcgttgaggc	300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat	360
ggtaccaa at cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg	420
gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg	480
ataatgaagt ggtaaaggag gtgaaaaatgc aaatacaacc aggaggttgg cttagaagca	540
gccatccttt aaagagtgcg taatagctca ctg	573

<210> 19
<211> 565
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 19	
tggtgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc	60
cttagcga aa gcgagtctga atagggcgat caaatatttt aatatttaca atttagtcat	120
tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc	180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg	240
aaataccagt cgaacccgga gctagctggg tctccccgaa atgcgttgag gcgcagcagt	300
acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaaa acggtaccaa	360
atcgtggcaa actctgaata ctagaaatga cggtgtagta gtgagactgt gggggataag	420
ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta	480
gtgacaaagg aggtgaaaat gcaaacacaa ccaggaggtt ggcttagaag cagccatcct	540

32378198_1.TXT

ttaaagagtg cgtaatagct cactg

565

<210> 20

<211> 565

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 20

tggtgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60

cttagcgaaa gcgagtctga atagggcgat caaatatattt aatatttaca atttagtcat 120

tttttctaga cccgaaccgg ggtgatctaa ccatgaccag gatgaaactt ggggtgatacc 180

aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg 240

aaataccagt cgaaccggga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300

acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaaa acggtaccaa 360

atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420

ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggtt ggcttagaag cagccatcct 540

ttaaagagtg cgtaatagct cactg 565

<210> 21

<211> 546

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 21

tggtgaagaa tgagccggcg acttagaaaa cgtggcaagg ttaaggacat gtatccggag 60

ccgaagcgaa agcaagtctg aatagggcgc ctaagtcatt ttttctagac ccgaaccggg 120

gtgatctaac catgaccagg atgaagcttg ggtgacacca agtgaagggtc cgaaccgacc 180

gatgttgaaa aatcggcgga tgagttgtgg ttagcgggtga aataccagtc gaactcggag 240

ctagctgggtt ctccccgaaa tgcgttgagg cgcagcgggtt cataaggctg tctaggggta 300

aagcactgtt tcgggtgcggg ctgcgaaagc ggtaccaaact cgtggcaaac tctgaatact 360

agatatgcta tttatgagcc agtgagacgg tgggggataa gcttcatcgt cgagagggaa 420

acagcccaga tcactagcta agggcccctaa atgatcgta agtgacaaag gaggtgagaa 480

tgcagaaaca accaggaggt ttgcttagaa gcagccaccc tttaaagagt gcgtaatagc 540

tcactg 546

<210> 22

<211> 550

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 22

32378198_1.TXT

tggtgaagaa tgagccggcg acttatagga agtggcaggg ttaaggaaga atctccggag	60
cccaagcgaa agcgagtctg aaaagggcga tttggctact tcttatggac ccgaacctgg	120
atgatctaata catggccaag ttgaagcatg ggtaacacta tgctcaggac tgaaccacc	180
gatgttgaaa aatcggggga tgagctgtga ttagcgggtga aattccaatc gaattcagag	240
ctagctggat ctccccgaaa tgcgttgagg cgcagcggcg acgatgtcct gtctaagggt	300
agagcgactg tttcggtgcg ggctgcgaaa gcggtaccaa gtcgtggcaa actccgaata	360
ttaggcaaaag gattccgtga gccagtgaaga ctgtggggga taagcttcac agtcaagagg	420
gaaacagccc agaccatcag ctaaggcccc taaatggctg ctaagtggaa aaggatgtga	480
gaatgctgaa acaaccagga ggctcgctta gaagcagcta ttccttgaaa gagtgcgtaa	540
tagctcactg	550

<210> 23

<211> 548

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 23

tggtgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag	60
ccagagcgaa agcaagtctg aatagggcgc ttaaagggtca ctttttctag acccgaaccc	120
gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga	180
ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg	240
agctagctgg ttctccccga aatgcgttga ggcgcagcgg tttatgaggc tgtctagggg	300
taaagcactg tttcggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata	360
ctagatatgc tattcatgag ccagtgaagc ggtgggggat aagcttcacg gtcaagaggg	420
aaacagccca gatcaccagc taaggcccca aaatggctgt taagtggcaa aggaggtgag	480
aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata	540
gctcactg	548

<210> 24

<211> 556

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 24

tggtgaagaa tgagccggcg acttataggg ggtggcgtgg ttaaggaagt aatccgaagc	60
caaagcgaaa gcaagttttc aatagagcga ttttgtcacc ctttatggac ccgaaccggg	120
gtgatctaac cttgaccagg atgaagcttg ggtaacacca agtgaagggtc cgaactcatc	180
gatcttgaaa aatcgtggga tgagttgggg ttagttgggt aaatgctaata cgaactcggg	240
gctagctggg tctccccgaa atgtgttgag gcgcagcgat taacgaaata ttttgtacgg	300
tttaggggta aagcactgtt tcggtgcggg ctgcgaaagc ggtaccaaata cgtggcaaac	360
tctgaatact aagcctgtat accgttagtc agtgagagta taggggataa gctctatact	420

32378198_1.TXT

caagaggggaa acagcccaga tcaccagcta agggcccaaa atgacagcta agtggcaaag 480
gaggtgaaaag tgcagaaaca accaggaggt tcgcttagaa gcagcaaccc tttaaagagt 540
gcgtaatagc tcaactg 556

<210> 25
<211> 548
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 25
tgttgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag 60
ccagagcgaa agcaagtctg aatagggcgc ttaaagggtca ctttttctag acccgaaccc 120
gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga 180
ccgatgttga aaaatcggcg gatgagttgt ggttagcggg gaaataccaa tcgaactcgg 240
agctagctgg ttctccccga aatgcgttga ggcgcagcgg tttatgaggc tgtctagggg 300
taaagcactg tttcgggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360
ctagatatgc tattcatgag ccagtgaagc ggtgggggat aagcttcac gtcaagaggg 420
aaacagccca gatcaccagc taaggcccca aaatggtcgt taagtggcaa aggaggtgag 480
aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540
gctcactg 548

<210> 26
<211> 565
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 26
tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60
cttagcga aa gcgagtctga atagggcgat caaatat tttt aatatttaca atttagtcat 120
tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcg atgagttgtg gttagcgggtg 240
aaataccagt cgaacccgga gctagctggg tctccccgaa atgcgttgag gcgcagcagt 300
acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaaa acggtaccaa 360
atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420
ctccattgtc aagaggggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480
gtgacaaagg aggtgaaaat gcaaacacaa ccaggaggtt ggcttagaag cagccatcct 540
ttaaagagtgcgtaatagct cactg 565

<210> 27
<211> 573
<212> DNA
<213> Artificial Sequence

32378198_1.TXT

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 27

tggtgaagaa tgagccggcg acttataggg ggtggcttgg ttaaggacta caatccgaag	60
cccaagcgaa agcaagtttg aagtgtacac acattgtgtg tctagagcga ttttgtcact	120
ccttatggac ccgaacccgg gtgatctatt catggccagg atgaagcttg ggtaacacca	180
agtgaaggtc cgaactcatc gatgttga aaatcgtggga tgagttgtga ataggggtga	240
aatgccaatc gaactcggag ctagctggtt ctccccgaaa tgtgttgagg cgcagcgatt	300
cacgatctaa agtacggttt aggggtaaag cactgtttcg gtgcgggctg ttaacgcggt	360
accaaactcg ggcaaactaa gaatactaaa cttgtatgcc gtgaatcagt gagactaaga	420
gggataagct tcttagtcaa gagggaaaca gccagatca ccagctaagg ccccaaatg	480
acagctaagt ggcaaaggag gtgagagtgc agaaacaacc aggaggtttg cttagaagca	540
gccatccttt aaagagtgcg taatagctca ctg	573

<210> 28

<211> 573

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 28

tggtgaagaa tgagccggcg acttataggg ggtggcttgg ttaaggacta caatccgaag	60
cccaagcgaa agcaagtttg aagtgtacac acgttgtgtg tctagagcga ttttgtcact	120
ccttatggac ccgaacccgg gtgatctatt catggccagg atgaagcttg ggtaacacca	180
agtgaaggtc cgaactcatc gatgttga aaatcgtggga tgagttgtga ataggggtga	240
aatgccaatc gaactcggag ctagctggtt ctccccgaaa tgtgttgagg cgcagcgatt	300
cacgatctaa agtacggttt aggggtaaag cactgtttcg gtgcgggctg ttaacgcggt	360
accaaactcg ggcaaactaa gaatactaaa cttgtatgcc gtgaatcagt gagactaaga	420
gggataagct tcttagtcaa gagggaaaca gccagatca ccagctaagg ccccaaatg	480
acagctaagt ggcaaaggag gtgagagtgc agaaacaacc aggaggtttg cttagaagca	540
gccatccttt aaagagtgcg taatagctca ctg	573

<210> 29

<211> 573

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 29

tggtgaagaa tgagccggcg acttataggg ggtggcttgg ttaaggacta caatccgaag	60
cccaagcgaa agcaagtttg aagtgtacac acattgtgtg tctagagcga ttttgtcact	120
ccttatggac ccgaacccgg gtgatctatt catggccagg atgaagcttg ggtaacacca	180
agtgaaggtc cgaactcatc gatgttga aaatcgtggga tgagttgtga ataggggtga	240

32378198_1.TXT

aatgccaatc gaactcggag ctagctggtt ctccccgaaa tgtgttgagg cgcagcgatt 300
cacgatctaa agtacgggtt aggggtaaa cactgtttcg gtgcgggctg ttaacgcggt 360
accaaactcg ggcaactaa gaatactaaa cttgtatgcc gtgaatcagt gagactaaga 420
gggataagct tcttagtcaa gagggaaaca gcccagatca ccagctaagg ccccaaatg 480
acagctaagt ggcaaaggag gtgagagtgc agaaacaacc aggaggtttg cttagaagca 540
gccatccttt aaagagtgcg taatagctca ctg 573

<210> 30

<211> 546

<212> DNA

<213> Chlorella protothecoides

<400> 30

tggtgaagaa tgagccggcg acttagaaaa cgtggcaagg ttaaggaaac gtatccggag 60
ccgaagcgaa agcaagtctg aacagggcga ttaagtcatt ttttctagac ccgaaccg 120
gtgatctaac catgaccagg atgaagcttg ggtgacacca agtgaaggct cgaaccgacc 180
gatgttgaaa aatcggcgga tgagttgttg ttagcggatga aataccagtc gaactcggag 240
ctagctggtt ctccccgaaa tgcgttgagg cgcagcggtt cataaggctg tctaggggta 300
aagcactgtt tcggtgcggg ctgcgaaagc ggtaccaaact cgtggcaaac tctgaatact 360
agatatgcta tttatgggcc agtgagacgg tgggggataa gcttcatcgt cgagagggaa 420
acagcccaga ttagtagcta agggcccaaa atgacgttta agtgacaaag gaggtgagaa 480
tgcagaaaca accaggaggt ttgcttagaa gcagccaccc tttaaagagt gcgtaatagc 540
ttagtg 546

<210> 31

<211> 565

<212> DNA

<213> Chlorella protothecoides

<400> 31

tggtgaagaa tgagccggcg acttagaaaa agtggcgtag ttaaggaaaa attccgaagc 60
cttagcgaaa gcgagtctga atagggcgat caaatatttt aatatttaca atttagtc 120
tttttctaga cccgaaccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180
aagtgaaggc ccgaaccgac cgatgttgaa aaatcggcg atgagttgtg gttagcgg 240
aaataccagt cgaaccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300
acatctagtc tatctagggg taaagcactg tttcgggtgc ggctgtgaaa acggtaccaa 360
atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420
ctccattgtc aagaggggaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480
gtgacaaagg aggtgaaaat gcaaacacaa ccaggaggtt ggcttagaag cagccatcct 540
ttaaagagtg cgtaatagct cactg 565

<210> 32

<211> 565

<212> DNA

<213> Prototheca moriformis

<400> 32

tggtgaagaa tgagccggcg acttagaaaa ggtggcatgg ttaaggaaat attccgaagc 60

32378198_1.TXT

cgtagcaaaa gcgagtctga atagggcgat aaaatatatt aatatttaga atctagtc	120
tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaagctt gggtgatacc	180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg	240
aaataccagt cgaacccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt	300
acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaga acggtaccaa	360
atcgtggcaa actctgaata ctagaaatga cgatgtagta gtgagactgt gggggataag	420
ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta	480
gtgacaaagg aggtgaaaat gcaaatacaa ccaggagggtt ggcttagaag cagccatcct	540
ttaaagagtg cgtaatagct cactg	565

<210> 33
 <211> 573
 <212> DNA
 <213> Prototheca moriformis

<400> 33	
tggtgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga	60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaacatta aataaaatct	120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg	180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt	240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc	300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat	360
ggtaccaaata cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg	420
gggataagct ccatagtcga gagggaaaca gccagacca ccagttaagg ccccaaatg	480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca	540
gccatccttt aaagagtgcg taatagctca ctg	573

<210> 34
 <211> 541
 <212> DNA
 <213> Prototheca kruegani

<400> 34	
tggtgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa tactctggag	60
ccatagcgaa agcaagttta gtaagcttag gtcattcttt ttagacccga aaccgagtga	120
tctacccatg atcaggggtga agtgtagta aaataacatg gagggccgaa ccgactaatg	180
ttgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag	240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcag taaaaataga ggggtaaagc	300
actgtttctt ttgtgggctt cgaaagtgtg acctcaaagt ggcaaactct gaatactcta	360
tttagatatc tactagttag accttggggg ataagctcct tgggtcaaaag ggaaacagcc	420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggatatgtca	480
aaacctccag caggtttagct tagaagcagc aatcctttca agagtgcgta atagctcact	540
g	541

<210> 35
 <211> 573

32378198_1.TXT

<212> DNA

<213> Prototheca wickerhamii

<400> 35

```

tggtgaagaa tgagccggcg acttaaaata aatggcaggc taagagattt aataactcga      60
aacctaagcg aaagcaagtc ttaatagggc gtcaatttaa caaaacttta aataaattat      120
aaagtcattt atttttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg      180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggg      240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc      300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat      360
ggtagcaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg      420
gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg      480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca      540
gccatccttt aaagagtgcg taatagctca ctg                                     573

```

<210> 36

<211> 541

<212> DNA

<213> Prototheca moriformis

<400> 36

```

tggtgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagataa ttctctggag      60
ccatagcgaa agcaagttaa acaagctaaa gtcacccttt ttagaccgga aaccgagtga      120
tctacccatg atcaggggtga agtggttgga aaataacatg gagggccgaa ccgactaatg      180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag      240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc      300
actgtttctt ttgtgggctt cgaaagtgtg acctcaaagt ggcaaactct gaatactcta      360
tttagatata tactagttag accttggggg ataagctcct tggcaaaaag ggaaacagcc      420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggatatgta      480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact      540
g                                     541

```

<210> 37

<211> 573

<212> DNA

<213> Prototheca wickerhamii

<400> 37

```

tggtgaagaa tgagccgtcg acttaaaata aatggcaggc taagagaatt aataactcga      60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct      120
aaagtcattt atttttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg      180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggg      240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc      300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat      360
ggtagcaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg      420
gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg      480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca      540

```

32378198_1.TXT

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 38
<211> 541
<212> DNA
<213> Prototheca moriformis

<400> 38
tggtgaagaa tgagccggcg agttaaaaag agtggcgtgg ttaaagaaaa ttctctggaa 60
ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagaccgga aaccgagtga 120
tctacccatg atcaggggtga agtggttgga aaataacatg gagggccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt ttgtgggctc cgaaagtgt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagttag accttggggg ataagctcct tggtcgaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

<210> 39
<211> 541
<212> DNA
<213> Prototheca zopfii

<400> 39
tggtgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa ttctctggag 60
ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagaccgga aaccgagtga 120
tctacccatg atcaggggtga agtggttgga aaataacatg gagggccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt tcgtgggctt cgaaagtgt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagttag accttggggg ataagctcct tggtcgaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

<210> 40
<211> 39
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 40
Met Thr Phe Gly Val Ala Leu Pro Ala Met Gly Arg Gly Val Ser Leu
1 5 10 15
Pro Arg Pro Arg Val Ala Val Arg Ala Gln Ser Ala Ser Gln Val Leu
20 25 30

32378198_1.TXT

Glu Ser Gly Arg Ala Gln Leu
35

<210> 41
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 41
Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr
1 5 10 15

Ile Gly Thr Leu Arg Lys Ala Ile Pro Ala His Cys Phe Glu Arg Ser
20 25 30

Ala Leu Arg Gly Arg Ala Gln Leu
35 40

<210> 42
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 42
Met Ala Ser Ala Ala Phe Thr Met Ser Ala Cys Pro Ala Met Thr Gly
1 5 10 15

Arg Ala Pro Gly Ala Arg Arg Ser Gly Arg Pro Val Ala Thr Arg Leu
20 25 30

Arg Gly Arg Ala
35

<210> 43
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 43
Met Ala Thr Ala Ser Thr Phe Ser Ala Phe Asn Ala Arg Cys Gly Asp
1 5 10 15

Leu Arg Arg Ser Ala Gly Ser Gly Pro Arg Arg Pro Ala Arg Pro Leu
20 25 30

Pro Val Arg Gly Arg Ala Gln Leu
35 40

<210> 44
<211> 87
<212> PRT
<213> Artificial Sequence

32378198_1.TXT

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 44

Met Val Ala Ala Ala Ser Ser Ala Phe Phe Pro Val Pro Ala Pro
1 5 10 15

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
20 25 30

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val
35 40 45

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser
50 55 60

Leu Lys Ser Gly Ser Leu Asn Thr Gln Glu Asp Thr Ser Ser Ser Pro
65 70 75 80

Pro Pro Arg Thr Phe Leu His
85

<210> 45

<211> 60

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 45

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

Gln Leu Arg Ala Gly Asn Ala Pro Thr Ser Leu Lys Met Ile Asn Gly
35 40 45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Arg Leu
50 55 60

<210> 46

<211> 60

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 46

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

Gln Leu Arg Ala Gly Asn Ala Gln Thr Ser Leu Lys Met Ile Asn Gly

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Lys Leu
50 55 60

<210> 47
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 47
tgacctagggt gattaattaa ctcgaggcag cagcagctcg gatagtatcg 50

<210> 48
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

● <400> 48
ctacgagctc aagctttcca ttgtgttcc catcccacta ctcc 45

<210> 49
<211> 1568
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

● <400> 49
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt 300
cagtcacaac ccgcaaacgg cgcgccatat caatgattga acaagatgga ttgcacgcag 360
gttctccggc cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg 420
gctgctctga tgccgccgtg ttccggctgt cagcgcaggg ggcgccggtt ctttttgtca 480
agaccgacct gtccggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc 540
tgggcacgac gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg 600
actggctgct attgggcgaa gtgccggggc aggatctcct gtcactcac cttgctcctg 660
ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta 720
cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag 780
ccggtcttgt cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac 840
tgttcgccag gctcaaggcg cgcattgccc acggcgagga tctcgtcgtg acccatggcg 900
atgcctgctt gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg 960

32378198_1.TXT

```

gccggctggg tgtggcggac cgctatcagg acatagcgtt ggctacccgt gatattgctg 1020
aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct ttacgggtatc gccgctcccg 1080
attcgcagcg catcgccctt tatcgccctt ttgacgagtt cttctaagat ctgtcgatcg 1140
acaagtgact cgaggcagca gcagctcgga tagtatcgac acactctgga cgctggtcgt 1200
gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct gccgctttta 1260
tcaaacagcc tcagtgtgtt tgatcttggt tgtacgcgct tttgcgagtt gctagctgct 1320
tgtgctattt gcgaatacca cccccagcat ccccttccct cgtttcatat cgcttgcatc 1380
ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg ctctgctca 1440
ctgccccctg cacagccttg gtttgggctc cgctgtatt ctctggtac tgcaacctgt 1500
aaaccagcac tgcaatgctg atgcacggga agtagtggga tgggaacaca aatggaaagc 1560
ttgagctc 1568

```

<210> 50
 <211> 2571
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
 <221> modified_base
 <222> (997)..(999)
 <223> a, c, t, g, unknown or other

```

<400> 50
gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagaggct gggcagtgc 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgc 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc 240
tccagggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcg acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaactaac catagctgat caacactgca atcatcggcg gctgatgcaa 540
gcatcctgca agacacatgc tgtgcgatgc tgcgctgctg cctgctgcgc acgccgttga 600
gttggcagca gctcagccat gactggatc aggctgggct gccactgcaa tgtggtggat 660
aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc 720
accatcagca ggagcgcag cgaagggact ggccccatgc acgcatgcc aaaccggagc 780
gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaaccatg ctcatggacg 840
catgtagcgc tgacgtccct tgacggcgct cctctcgggt gtgggaaacg caatgcagca 900
caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcggggggcc acccttcttg 960
cggggtcgcg cccagccag cggatgatgc ctgatcnnc caaacgagtt cacattcatt 1020
tgcagcctgg agaagcgagg ctggggcctt tgggctggtg cagcccgcaa tggaatgcgg 1080

```

32378198_1.TXT

gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcac 1140
tggactgcat ttgggtgggg cggccggctg tttctttcgt gttgcaaaac gcgccacgtc 1200
agcaacctgt cccgtgggtc ccccggtgcc atgaaatcgt gtgcacgccg atcagctgat 1260
tgccccggctc gcgaagtagg cgccctcttt ctgctcgccc tctctccgtc ccgccactag 1320
tggcgcgcca tatcaatgat tgaacaagat ggattgcacg caggttctcc ggccgcttgg 1380
gtggagaggc tattcggtc tgactgggca caacagacaa tcggctgctc tgatgccgcc 1440
gtgttccggc tgtcagcgca ggggcgcccc gttctttttg tcaagaccga cctgtccggt 1500
gccctgaatg aactgcagga cgaggcagcg cggctatcgt ggctggccac gacgggcgtt 1560
ccttgcgag ctgtgctcga cgttgctact gaagcgggaa gggactggct gctattgggc 1620
gaagtgccgg ggcaggatct cctgtcatct caccttgctc ctgccgagaa agtatccatc 1680
atggctgatg caatgcggcg gctgcatacg cttgatccgg ctacctgcc attcgaccac 1740
caagcgaaac atcgcatcga gcgagcacgt actcggatgg aagccgggtct tgtcgatcag 1800
gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc caggctcaag 1860
gcgcgcatgc ccgacggcga ggatctcgtc gtgacccatg gcgatgcctg cttgccgaat 1920
atcatggtgg aaaatggccg cttttctgga ttcacgact gtggccggct ggggtgtggcg 1980
gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa 2040
tgggctgacc gcttcctcgt gctttacggc atcgccgctc ccgattcgca gcgcacgcc 2100
ttctatcgcc ttcttgacga gttcttctaa gatctgtcga tcgacaagtg actcgaggca 2160
gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 2220
ccacacttgc tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt 2280
gtttgatctt gtgtgtacgc gcttttgca gttgctagct gcttgtgcta tttgcaata 2340
ccacccccag catccccctc cctcgtttca tatcgcttgc atcccaaccg caacttatct 2400
acgctgtcct gctatccctc agcgtgctc ctgctcctgc tctactgccc tcgcacagcc 2460
ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg 2520
ctgatgcacg ggaagtagtg ggatgggaac acaaattggaa agcttgagct c 2571

<210> 51
<211> 2550
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>
<221> modified_base
<222> (997)..(999)
<223> a, c, t, g, unknown or other

<400> 51
gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgccctgccgc atgcttgtgc tgggtaggct gggcagtgtc 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180
gccgggcaat gcgcattggt gcaagagggc ggcagcactt gctggacgtg ccgcgggtgcc 240

32378198_1.TXT

tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
 cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
 gcaggagtca tccaactaac catagctgat caacactgca atcatcggcg gctgatgcaa 540
 gcatcctgca agacacatgc tgtgcatgc tgcgctgctg cctgctgcgc acgccgttga 600
 gttggcagca gctcagccat gcaactggatc aggctgggct gccactgcaa tgtggtggat 660
 aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc 720
 accatcagca ggagcgcgtg cgaagggact ggccccatgc acgccatgcc aaaccggagc 780
 gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaaccatg ctcatggagc 840
 catgtagcgc tgacgtccct tgacggcgct cctctcgggt gtgggaaacg caatgcagca 900
 caggcagcag aggcggcgcc agcagagcgg cggcagcagc ggcgggggcc acccttcttg 960
 cggggtcgcg cccagccag cggatgagc ctgactnnnc caaacgagtt cacattcatt 1020
 tgcagcctgg agaagcgagg ctggggcctt tgggctggtg cagcccgcaa tggaatgcgg 1080
 gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcatt 1140
 tggactgcat ttgggtgggg cggccggctg tttctttcgt gttgaaaaac gcgccacgtc 1200
 agcaacctgt cccgtgggtc ccccgctgcc atgaaatcgt gtgcacgccg atcagctgat 1260
 tgccccgctc gcgaagtagg cgccctcttt ctgctcgccc tctctccgtc ccgccactag 1320
 tggcgcgcca tatcaatgat cgagcaggac ggccctccac ccggctcccc cgccgcctgg 1380
 gtggagcgcc tgttcggcta cgactgggccc cagcagacca tcggctgctc cgacgccgcc 1440
 gtgttccgcc tgtccgcca gggccgcccc gtgctgttcg tgaagaccga cctgtccggc 1500
 gccctgaacg agctgcagga cgaggccgcc cgctgtctct ggctggccac caccggcgtg 1560
 ccctgcgccg ccgtgctgga cgtggtgacc gaggccggcc gcgactggct gctgctgggc 1620
 gaggtgcccc gccaggacct gctgtcctcc cacctggccc ccgccgagaa ggtgtccatc 1680
 atggccgacg ccatgcgccg cctgcacacc ctggaccccc ccacctgccc ctctgaccac 1740
 caggccaagc accgcatcga gcgcgccccg acccgcatgg aggcggcct ggtggaccag 1800
 gacgacctgg acgaggagca ccagggcctg gccccgcgc agctgttcgc ccgcctgaag 1860
 gcccgcgtgc ccgacggcga ggacctggtg gtgacccacg gcgacgcctg cctgccccaac 1920
 atcatggtgg agaacggcgc cttctccggc ttcactgact gcggccgcct gggcgtggcc 1980
 gaccgctacc aggacatcgc cctggccacc cgcgacatcg ccgaggagct gggcggcgag 2040
 tgggccgacc gcttctggtg gctgtacggc atcgccgccc ccgactccca gcgcatcgcc 2100
 ttctaccgcc tgctggacga gttcttctga ctcgaggcag cagcagctcg gatagtatcg 2160
 acacactctg gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct 2220
 gtgaatatcc ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg 2280
 cttttgcgag ttgctagctg cttgtgctat ttgcgaatac cccccccagc atccccctcc 2340
 ctcgtttcat atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca 2400
 gcgctgctcc tgctcctgct cactgcccct cgcacagcct tggtttgggc tccgcctgta 2460
 ttctcctggt actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg 2520

32378198_1.TXT

gatgggaaca caaatggaaa gcttgagctc 2550

<210> 52
 <211> 1547
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 52
 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 300
 cagtcacaac ccgcaaacgg cgcgccatat caatgatcga gcaggacggc ctccacgccg 360
 gctccccgc gcctgggtg gagcgctgt tcggctacga ctgggcccag cagaccatcg 420
 gctgctccga cgccgccgtg ttccgcctgt ccgcccaggg ccgccccgtg ctgttcgtga 480
 agaccgacct gtccggcgcc ctgaacgagc tgcaggacga ggccgcccgc ctgtcctggc 540
 tggccaccac cggcgtgccc tgcgccgccg tgctggacgt ggtgaccgag gccggccgcg 600
 actggctgct gctgggcgag gtgcccggcc aggacctgct gtccctccac ctggcccccg 660
 ccgagaaggt gtccatcatg gccgacgcca tgcgccgcct gcacaccctg gaccccgcca 720
 cctgcccctt cgaccaccag gccaaagcacc gcatcgagcg cgcgcgcacc cgcattggagg 780
 ccggcctggt ggaccaggac gacctggacg aggagcacca gggcctggcc cccgcccagc 840
 tgctcgcccc cctgaaggcc cgcattgccc acggcgagga cctgggtggtg acccacggcg 900
 acgcctgcct gcccaacatc atggtggaga acggccgctt ctccggcttc atcgactgcg 960
 gccgcctggg cgtggccgac cgctaccagg acatcgccct ggccacccgc gacatcgccg 1020
 aggagctggg cggcgagtgg gccgaccgct tcttgggtgt gtacggcatc gccgcccccg 1080
 actcccagcg catcgcttc taccgcctgc tggacgagtt cttctgactc gaggcagcag 1140
 cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac 1200
 acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt 1260
 gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac 1320
 ccccgacatc ccctccctc gtttcataac gcttgcatcc caaccgcaac ttatctacgc 1380
 tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 1440
 tttgggctcc gcctgtattc tcctgttact gcaacctgta aaccagcact gcaatgctga 1500
 tgcacgggaa gtagtgggat gggaacacaa atggaaagct tgagctc 1547

<210> 53
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

32378198_1.TXT

<400> 53
gccgcgactg gctgctgctg g 21

<210> 54
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 54
aggtcctcgc cgtcgggcat g 21

<210> 55
<211> 2357
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 55
ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc agggccccga ttgcaaagac 180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240
ccactcgagc ttgtgatcgc actccgctaa gggggcgccct cttcctcttc gtttcagtca 300
caaccgcgaa acggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg 360
gttttgctgc caagatcagc gcctctatga cgaacgaaac ctcggataga ccacttgctg 420
actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag 480
atgccaagtg gcatctgtac tttcaataca acccgaaacga tactgtctgg gggacgcat 540
tgttttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta 600
tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtggtt gactacaaca 660
atacttccgg ctttttcaac gataccattg acccgagaca acgctgcgtg gccaratgga 720
cttacaacac accggagtcc gaggagcagt acatctcgta tagcctggac ggtggataca 780
cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc 840
cgaaggtctt ttggtacgag ccctcgaga agtggatcat gacagcggca aagtcacagg 900
actacaagat cgaaatttac tcgtctgacg accttaaate ctggaagctc gaateccgct 960
tcgcaaacga gggctttctc ggctaccaat acgaatgcc aggcctgata gaggtcccaa 1020
cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac 1080
cggcaggagg ttcttttaat cagtacttcg tcggaagctt taacggaact catttcgagg 1140
catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt 1200
tcttcaatac tgacccgacc tatgggagcg ctcttggcat tgcgtgggct tctaactggg 1260
agtattccgc attcgttcct acaaaccctt ggaggtcctc catgtcgtc gtgaggaaat 1320
tctctctcaa cactgagtac caggccaacc cggaaccga actcataaac ctgaaagccg 1380
aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt 1440

32378198_1.TXT

tgacgaaagc caacagctac aacgtcgatc tttcgaatag caccggtaca cttgaatttg	1500
aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct	1560
ccctctgggtt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt	1620
ctgcgtcctc cttcttcctt gatcgcgga acagcaaagt aaaatttggt aaggagaacc	1680
catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt	1740
cgtactacaa agtgtatggt ttgcttgatc aaaatatcct ggaactctac ttcaacgatg	1800
gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga	1860
acatgacgac ggggtgtggat aacctgttct acatcgacaa attccagggtg agggaaagtca	1920
agtgagatct gtcgatcgac aagctcgagg cagcagcagc tcggatagta tcgacacact	1980
ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata	2040
tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc	2100
gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatccct tccctcgttt	2160
catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc	2220
tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct	2280
ggctactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga	2340
acacaaatgg aaagctt	2357

<210> 56

<211> 2335

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 56

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg	60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc	120
gctccgatgc cgctccaggg cgagcgctgt taaatagcc agggccccga ttgcaaagac	180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg	240
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct ctctctcttc gtttcagtca	300
caaccgcgaa acggcgcgcc atgctgctgc aggccttctt gtctctgctg gccggcttcg	360
ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg gtgcacttca	420
cccccaacaa gggctggatg aacgaccca acggcctgtg gtacgacgag aaggacgcca	480
agtggcacct gtacttccag tacaaccga acgacaccgt ctgggggacg cccttgttct	540
ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc gccatcgccc	600
cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac aacaacacct	660
ccggcttctt caacgacacc atcgaccgc gccagcgctg cgtggccatc tggacctaca	720
acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc tacaccttca	780
ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc gacccgaagg	840
tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc caggactaca	900
agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc gcgttcgcca	960

32378198_1.TXT

```

acgagggcctt cctcggctac cagtacgagt gccccggcct gatcgaggtc cccaccgagc      1020
aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc gccccggccg      1080
gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc gaggccttcg      1140
acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag accttcttca      1200
acaccgaccc gacctacggg agcgccctgg gcatcgcggtg ggccctcaac tgggagtact      1260
ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc aagttctccc      1320
tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag gccgagccga      1380
tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc acgttgacga      1440
aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag ttcgagctgg      1500
tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac ctctccctct      1560
ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag gtgtccgct      1620
cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag aaccctact      1680
tcaccaaccg catgagcgtg aacaaccagc cttcaagag cgagaacgac ctgtcctact      1740
acaagggtga cggttgctg gaccagaaca tcctggagct gtacttcaac gacggcgacg      1800
tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc gtgaacatga      1860
cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag gtcaagtgat      1920
taattaactc gaggcagcag cagctcggat agtatcgaca cactctggac gctggtcgtg      1980
tgatggactg ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat      2040
caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt      2100
gtgctatttg cgaataccac ccccagcatc ccctccctc gtttcataac gcttgcatcc      2160
caaccgcaac ttatctacgc tgcctgcta tccctcagcg ctgctcctgc tcctgctcac      2220
tgcccctcgc acagccttgg tttgggctcc gcctgtattc tcctgggtact gcaacctgta      2280
aaccagcact gcaatgctga tgcacgggaa gtagtgggat gggaaacaaa atgga      2335

```

<210> 57

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 57

ccgccgtgct ggacgtggtg

20

<210> 58

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 58

ggtggcgggg tccagggtgt

20

<210> 59

32378198_1.TXT

```

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
        primer

<400> 59
cggccggcgg ctccttcaac                                20

<210> 60
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
        primer

• <400> 60
ggcgctcccc taggtcgggt                                20

<210> 61
<211> 1335
<212> DNA
<213> Chlorella sorokiniana

<400> 61
cgccctgcaac gcaagggcag ccacagccgc tcccacccgc cgctgaaccg acacgtgctt      60
gggcgcctgc cgctgcctg ccgcatgctt gtgctggtga ggctgggcag tgctgccatg      120
ctgattgagg cttggttcat cgggtggaag cttatgtgtg tgctgggctt gcatgccggg      180
caatgcgcat ggtggcaaga gggcggcagc acttgctgga gctgccgcgg tgcctccagg      240
tggttcaatc gcggcagcca gagggatttc agatgatcgc gcgtacaggt tgagcagcag      300
tgtcagcaaa ggtagcagtt tgccagaatg atcggttcag ctgttaatca atgccagcaa      360
gagaaggggt caagtgcaaa cacgggcatg ccacagcacg ggcaccgggg agtggaatgg      420
caccaccaag tgtgtgcgag ccagcatcgc cgctggctg tttcagctac aacggcagga      480
gtcatccaac gtaaccatga gctgatcaac actgcaatca tcgggcgggc gtgatgcaag      540
catgcctggc gaagacacat ggtgtgcgga tgctgccggc tgctgcctgc tgcgcacgcc      600
gttgagttgg cagcaggctc agccatgcac tggatggcag ctgggctgcc actgcaatgt      660
ggtggatagg atgcaagtgg agcgaatacc aaaccctctg gctgcttgct gggttgcatg      720
gcatgcgacc atcagcagga gcgcatgcga agggactggc cccatgcacg ccatgccaaa      780
ccggagcgca ccgagtgtcc aactgtcac caggcccgcga agctttgcag aaccatgctc      840
atggacgcat gtagcgctga cgtccettga cggcgctcct ctcgggtgtg ggaaacgcaa      900
tgcagcacag gcagcagagg cggcggcagc agagcggcgg cagcagcggc gggggccacc      960
cttcttgctg ggtcgcgccc cagccagcgg tgatgcgctg atcccaaacg agttcacatt     1020
catttgcatg cctggagaag cgaggctggg gcctttgggc tggatgcagc cgcaatggaa     1080
tgcgggaccg ccaggctagc agcaaaggcg cctcccctac tccgcatcga tgttccatag     1140
tgcattggac tgcatttggg tggggcggcc ggctgtttct ttcgtgttgc aaaacgcgcc     1200
agctcagcaa cctgtcccgt gggccccccg tgccgatgaa atcgtgtgca cgccgatcag     1260
ctgattgccc ggctcgcgaa gtaggcgccc tcctttctgc tcgccctctc tccgtcccgc     1320

```

32378198_1.TXT

cactagtggc gcgcc

1335

<210> 62

<211> 1146

<212> DNA

<213> Umbellularia californica

<400> 62

atggccacca ccagcctggc ctccgccttc tgctccatga aggccgtgat gctggcccg	60
gacggcccg gcatgaagcc ccgcagctcc gacctgcagc tgcgcgccgg caacgcccc	120
acctccctga agatgatcaa cggcaccaag ttcagctaca ccgagagcct gaagcgctg	180
cccgactggc ccatgctgtt cgccgtgatc accaccatct tcagcgccgc cgagaagcag	240
tggaccaacc tggagtggaa gcccagccc aagctgcccc agctgctgga cgaccattc	300
ggcctgcacg gcctgggtgtt ccgccgcacc ttcgccatcc gctcctacga ggtgggcccc	360
gaccgcagca cctccatcct ggccgtgatg aaccacatgc aggaggccac cctgaaccac	420
gccaagagcg tgggcatcct gggcgacggc ttcggcacca ccctggagat gtccaagcg	480
gacctgatgt ggggtggtgcg ccgcacccac gtggccgtgg agcgctaccc cacctggggc	540
gacaccgtgg aggtggagtg ctggatcggc gccagcggca acaacggcat gcgccgcgac	600
ttcctggtgc gcgactgcaa gaccggcgag atcctgacct gctgcacctc cctgagcgtg	660
ctgatgaaca cccgcacccg ccgcctgagc accatccccg acgaggtgcg cggcgagatc	720
ggccccgcct tcacgacaa cgtggccgtg aaggacgacg agatcaagaa gctgcagaag	780
ctgaacgact ccaccgccga ctacatccag ggccggcctga cccccgctg gaacgacctg	840
gacgtgaacc agcacgtgaa caacctgaag tacgtggcct ggggtgttcga gaccgtgccc	900
gacagcatct tcgagtccca ccacatcagc tccttcaccc tggagtaccg ccgcgagtgc	960
acccgcgact ccgtgctgcg cagcctgacc accgtgagcg gcggcagctc cgaggccggc	1020
ctggtgtgcg accacctgct gcagctggag ggccgcagcg aggtgctgcg cggccgcacc	1080
gagtggcgcc ccaagctgac cgactcctc cgccggcatca gcgtgatccc cgccgagccc	1140
cgcgctg	1146

<210> 63

<211> 1146

<212> DNA

<213> Cinnamomum camphora

<400> 63

atggccacca cctccctggc ctccgccttc tgcagcatga aggccgtgat gctggcccg	60
gacggcccg gcatgaagcc ccgctccagc gacctgcagc tgcgcgccgg caacgcccag	120
acctccctga agatgatcaa cggcaccaag ttctcctaca ccgagagcct gaagaagctg	180
cccgactggc ccatgctgtt cgccgtgatc accaccatct tctccgccgc cgagaagcag	240
tggaccaacc tggagtggaa gcccagccc aacccccccc agctgctgga cgaccattc	300
ggccccacg gcctgggtgtt ccgccgcacc ttcgccatcc gcagctacga ggtgggcccc	360
gaccgctcca ccagcatcgt ggccgtgatg aaccacctgc aggaggccgc cctgaaccac	420
gccaagtccg tgggcatcct gggcgacggc ttcggcacca ccctggagat gtccaagcg	480
gacctgatct ggggtggtgaa gcgcacccac gtggccgtgg agcgctaccc cgcctggggc	540
gacaccgtgg aggtggagtg ctgggtgggc gcctccggca acaacggccg ccgccacgac	600

32378198_1.TXT

```

ttcctggtgc gcgactgcaa gaccggcgag atcctgaccc gctgcacctc cctgagcgtg      660
atgatgaaca cccgcacccg ccgcctgagc aagatccccg aggaggtgcg cggcgagatc      720
ggccccgcct tcatcgacaa cgtggccgtg aaggacgagg agatcaagaa gccccagaag      780
ctgaacgact ccaccgccga ctacatccag ggcggcctga ccccccgctg gaacgacctg      840
gacatcaacc agcacgtgaa caacatcaag tacgtggact ggatcctgga gaccgtgccc      900
gacagcatct tcgagagcca ccacatctcc tccttcacca tcgagtaccg ccgcgagtgc      960
accatggaca gcgtgctgca gtccctgacc accgtgagcg gcggctcctc cgaggccggc     1020
ctggtgtgcg agcacctgct gcagctggag ggcggcagcg aggtgctgcg cgccaagacc     1080
gagtggcgcc ccaagctgac cgactccttc cgcggcacatca gcgtgatccc cgccgagtcc     1140
agcgtg                                           1146

```

<210> 64

<211> 72

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 64

```

atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac      60
gacgacaagt ga                                           72

```

<210> 65

<211> 408

<212> DNA

<213> Chlorella vulgaris

<400> 65

```

ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga      60
ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag     120
cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat     180
ttgcgaatac cccccccagc atcccccttc ctcgtttcat atcgcttgca tcccaaccgc     240
aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgcccct     300
cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaaccagc     360
actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatgga      408

```

<210> 66

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial sequence: Synthetic primer

<400> 66

```

ctgggcgacg gcttcggcac                                           20

```

<210> 67

<211> 20

<212> DNA

<213> Artificial Sequence

32378198_1.TXT

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 67
aagtcgcggc gcatgccgtt 20

<210> 68
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 68
taccccgctt gggcgacac 20

<210> 69
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 69
cttgctcagg cggcgggtgc 20

<210> 70
<211> 1317
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 70
atgggtggccg ccgcccctc cagcgccttc ttccccgtgc ccgcccccg cgcctcccc 60
aagcccggca agttcggcaa ctggccctcc agcctgagcc cctccttcaa gccaagtcc 120
atccccaacg gcggcttcca ggtgaaggcc aacgacagcg cccaccccaa ggccaacggc 180
tccgccgtga gcctgaagag cggcagcctg aacacccagg aggacacctc ctccagcccc 240
cccccccgca ccttcctgca ccagctgccc gactggagcc gcctgctgac cgccatcacc 300
accgtgttcg tgaagtccaa gcgccccgac atgcacgacc gcaagtccaa gcgccccgac 360
atgctgggtg acagcttcgg cctggagtcc accgtgcagg acggcctggt gttccgccag 420
tccttctcca tccgtccta cgagatcggc accgaccgca ccgccagcat cgagaccctg 480
atgaaccacc tgcaggagac ctccctgaac cactgcaaga gcaccggcat cctgctggac 540
ggcttcggcc gcaccctgga gatgtgcaag cgcgacctga tctgggtggt gatcaagatg 600
cagatcaagg tgaaccgcta ccccgccctg ggcgacaccg tggagatcaa caccgcttc 660
agccgcctgg gcaagatcgg catgggccgc gactggctga tctccgactg caacaccggc 720
gagatcctgg tgcgcgccac cagcgcctac gccatgatga accagaagac ccgcccctg 780
tccaagctgc cctacgaggt gcaccaggag atcgtgcccc tgttcgtgga cagccccgtg 840
atcgaggact ccgacctgaa ggtgcacaag ttcaaggatga agaccggcga cagcatccag 900

32378198_1.TXT

```

aagggcctga cccccggctg gaacgacctg gacgtgaacc agcacgtgtc caacgtgaag      960
tacatcggct ggatcctgga gagcatgccc accgaggtgc tggagacca ggagctgtgc      1020
tccctggccc tggagtaccg ccgcgagtgc ggccgcgact ccgtgctgga gagcgtgacc      1080
gccatggacc ccagcaaggt gggcgtgctc tcccagtacc agcacctgct gcgcctggag      1140
gacggcaccg ccatcgtgaa cggcgccacc gagtggcgcc ccaagaacgc cggcgccaac      1200
ggcgccatct ccaccggcaa gaccagcaac ggcaactccg tgtccatgga ctacaaggac      1260
cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtga      1317

```

<210> 71

<211> 1170

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 71

```

atggccaccg catccacttt ctcggcgctc aatgcccgtc gcggcgacct gcgtcgctcg      60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcccagctg      120
cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc      180
gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag      240
tccaccgtgc aggacggcct ggtgttccgc cagtccttct ccatccgctc ctacgagatc      300
ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg      360
aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc      420
aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc      480
tggggcgaca ccgtggagat caacaccgca ttcagccgcc tgggcaagat cggcatgggc      540
cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tgggtgcgcgc caccagcgcc      600
tacgccatga tgaaccagaa gaccgcgcgc ctgtccaagc tgccctacga ggtgcaccag      660
gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac      720
aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac      780
ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg      840
cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag      900
tgcgggccgc actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg      960
cgctcccagt accagcacct gctgcgcctg gaggacggca ccgccatcgt gaacggcgcc      1020
accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc      1080
aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac      1140
atcgactaca aggacgacga cgacaagtga      1170

```

<210> 72

<211> 1170

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

32378198_1.TXT

<400> 72
atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg 60
cgcaaggcca tccccgcgca ctgttttcgag cgctcggcgc ttcgtgggcg cgcccagctg 120
cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc 180
gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag 240
tccaccgtgc aggacggcct ggtgttccgc cagtccttct ccatccgctc ctacgagatc 300
ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg 360
aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc 420
aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc 480
tggggcgaca ccgtggagat caacaccgcg ttcagccgcc tgggcaagat cggcatgggc 540
cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tggtgcgcg caccagcgcc 600
tacgccatga tgaaccagaa gaccgcgcg ctgtccaagc tgccctacga ggtgcaccag 660
gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac 720
aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac 780
ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg 840
cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag 900
tgcggccgcy actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggctg 960
cgctcccagt accagcacct gctgcgcctg gaggacggca ccgccatcgt gaacggcgcc 1020
accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc 1080
aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac 1140
atcgactaca aggacgacga cgacaagtga 1170

<210> 73
<211> 1167
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 73
atgacgttcg gggtcgcctt cccggccatg ggccgcggtg tctcccttcc ccggcccagg 60
gtcgcggtgc gcgcccagtc ggcgagtcag gttttggaga gcgggcgcgc ccagctgccc 120
gactggagcc gcctgctgac cgccatcacc accgtgttcg tgaagtccaa gcgccccgac 180
atgcacgacc gcaagtccaa gcgccccgac atgctggtgg acagcttcgg cctggagtcc 240
accgtgcagg acggcctggg gttccgccag tccttctcca tccgctccta cgagatcggc 300
accgaccgca ccgccagcat cgagaccctg atgaaccacc tgcaggagac ctccctgaac 360
cactgcaaga gcaccggcat cctgctggac ggcttcggcc gcaccctgga gatgtgcaag 420
cgcgacctga tctgggtggt gatcaagatg cagatcaagg tgaaccgcta cccgcctgg 480
ggcgacaccg tggagatcaa caccgcctt agccgcctgg gcaagatcgg catgggcccgc 540
gactggctga tctccgactg caacaccggc gagatcctgg tgcgcgccac cagcgcctac 600
gccatgatga accagaagac ccgccgcctg tccaagctgc cctacgaggt gcaccaggag 660
atcgtgcccc tgttcgtgga cagccccgtg atcaggagact ccgacctgaa ggtgcacaag 720

32378198_1.TXT

ttcaaggtga agaccggcga cagcatccag aagggcctga cccccggctg gaacgacctg	780
gacgtgaacc agcacgtgtc caacgtgaag tacatcggct ggatcctgga gagcatgccc	840
accgaggtgc tggagaccca ggagctgtgc tccctggccc tggagtaccg ccgcgagtgc	900
ggccgcgact ccgtgctgga gagcgtgacc gccatggacc ccagcaaggt gggcgtgcgc	960
tcccagtacc agcacctgct gcgcctggag gacggcaccg ccacgtgaa cggcgccacc	1020
gagtggcgcc ccaagaacgc cggcgccaac ggcgccatct ccaccggcaa gaccagcaac	1080
ggcaactccg tgtccatgga ctacaaggac cacgacggcg actacaagga ccacgacatc	1140
gactacaagg acgacgacga caagtga	1167

<210> 74

<211> 1149

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 74

atgacgttcg gggtcgccct cccggccatg ggccgcggtg tctcccttcc ccggcccagg	60
gtcgcggtgc gcgcccagtc ggcgagtcag gttttggaga gcgggcgcg ccccgactgg	120
tccatgctgt tcgccgtgat caccaccatc ttcagcgccg ccgagaagca gtggaccaac	180
ctggagtgga agcccaagcc caagctgccc cagctgctgg acgaccactt cggcctgcac	240
ggcctgggtg tccgccgcac cttcgccatc cgctcctacg aggtgggccc cgaccgcagc	300
acctccatcc tggccgtgat gaaccacatg caggaggcca ccctgaacca cgccaagagc	360
gtgggcatcc tgggcgacgg cttcggcacc accctggaga tgtccaagcg cgacctgatg	420
tgggtggtgc gccgcacca cgtggccgtg gagcgctacc ccacctgggg cgacaccgtg	480
gaggtggagt gctggatcgg cgccagcggc aacaacggca tgcgccgcga cttcctggtg	540
cgcgactgca agaccggcga gatcctgacc cgctgcacct ccctgagcgt gctgatgaac	600
acccgcaccc gccgcctgag caccatcccc gacgaggtgc gcggcgagat cggccccgcc	660
ttcatcgaca acgtggccgt gaaggacgac gagatcaaga agctgcagaa gctgaacgac	720
tccaccgccg actacatcca gggcggcctg accccccgct ggaacgacct ggacgtgaac	780
cagcacgtga acaacctgaa gtacgtggcc tgggtgttcg agaccgtgcc cgacagcatc	840
ttcgagtccc accacatcag ctcttcacc ctggagtacc gccgcgagtg caccgcgac	900
tccgtgctgc gcagcctgac caccgtgagc ggccggcagct ccgaggccgg cctggtgtgc	960
gaccacctgc tgcagctgga gggcggcagc gaggtgctgc gcgcccgcac cgagtggcgc	1020
cccaagctga ccgactcctt ccgcggcatc agcgtgatcc ccgccgagcc ccgcgtgatg	1080
gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa ggacgacgac	1140
gacaagtga	1149

<210> 75

<211> 1146

<212> DNA

<213> Artificial Sequence

<220>

32378198_1.TXT

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 75

atggcttccg cggcattcac catgtcggcg tgccccgcga tgactggcag ggccccctggg	60
gcacgtcgct ccggacggcc agtcgccacc cgcctgaggg ggcgcgcccc cgactgggtcc	120
atgctgttcg ccgtgatcac caccatcttc agcgccgccg agaagcagtg gaccaacctg	180
gagtgggaagc ccaagcccaa gctgccccag ctgctggacg accacttcgg cctgcacggc	240
ctggtgttcc gccgcacctt cgccatccgc tcctacgagg tgggccccga ccgcagcacc	300
tccatcctgg ccgtgatgaa ccacatgcag gaggccaccc tgaaccacgc caagagcgtg	360
ggcatcctgg gcgacggctt cggcaccacc ctggagatgt ccaagcgcga cctgatgtgg	420
gtggtgcgcc gcacccacgt ggccgtggag cgctacccca cctggggcga caccgtggag	480
gtggagtgtc ggatcggcgc cagcggcaac aacggcatgc gccgcgactt cctggtgcgc	540
gactgcaaga ccggcgagat cctgacccgc tgcacctccc tgagcgtgct gatgaacacc	600
cgcacccgcc gcctgagcac catccccgac gagggtgcgc gcgagatcgg ccccgcttc	660
atcgacaacg tggccgtgaa ggacgacgag atcaagaagc tgcagaagct gaacgactcc	720
accgccgact acatccaggg cggcctgacc ccccgctgga acgacctgga cgtgaaccag	780
cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga ccgtgcccga cagcatcttc	840
gagtcccacc acatcagctc cttcaccctg gagtaccgcc gcgagtgcac ccgcgactcc	900
gtgctgcgca gcctgaccac cgtgagcggc ggcagctccg aggccggcct ggtgtgcgac	960
cacctgtgc agctggaggg cggcagcgag gtgctgcgcg cccgcaccga gtggcgcccc	1020
aagctgaccg actccttccg cggcatcagc gtgatccccg ccgagccccg cgtgatggac	1080
tacaaggacc acgacggcga ctacaaggac cagacatcg actacaagga cgacgacgac	1140
aagtga	1146

<210> 76

<211> 1155

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 76

atggccaccg catccacttt ctcggcgttc aatgcccgtc gcggcgacct gcgtcgctcg	60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccccgac	120
tggatcatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc	180
aacctggagt ggaagcccaa gcccagctg cccagctgc tggacgacca cttcggcctg	240
cacggcctgg tgttccgccg caccttcgcc atccgctcct acgaggtggg ccccgaccgc	300
agcacctcca tcctggccgt gatgaaccac atgcaggagg ccacctgaa ccacgccaag	360
agcgtgggca tcctgggcga cggttcggc accaccctgg agatgtccaa gcgcgacctg	420
atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgt accccacctg gggcgacacc	480
gtggaggtgg agtctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttcctg	540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgctgatg	600

32378198_1.TXT

```

aacacccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc 660
gccttcatcg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac 720
gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacgtg 780
aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgcagc 840
atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcacccgc 900
gactccgtgc tgcgcagcct gaccaccgtg agcggcgcca gctccgaggc cggcctggtg 960
tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgccc caccgagtgg 1020
cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg 1080
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140
gacgacaagt gatga 1155

```

<210> 77

<211> 1152

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 77

```

atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg 60
cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg ccccccgac 120
tggtccatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc 180
aacctggagt ggaagcccaa gcccaagctg cccagctgc tggacgacca cttcggcctg 240
cacggcctgg tgttccgccc caccttcgcc atccgctcct acgaggtggg cccgcaccgc 300
agcacctcca tcctggccgt gatgaaccac atgcaggagg ccaccctgaa ccacgccaag 360
agcgtgggca tcctgggcga cggcttcggc accaccctgg agatgtccaa gcgcgacctg 420
atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgtc accccacctg gggcgacacc 480
gtggagggtg agtgctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttcctg 540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgctgatg 600
aacacccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc 660
gccttcatcg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac 720
gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacgtg 780
aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgcagc 840
atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcacccgc 900
gactccgtgc tgcgcagcct gaccaccgtg agcggcgcca gctccgaggc cggcctggtg 960
tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgccc caccgagtgg 1020
cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg 1080
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140
gacgacaagt ga 1152

```

<210> 78

<211> 1155

<212> DNA

32378198_1.TXT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 78

atggccaccg catccacttt ctcggcggtc aatgcccgtc gcggcgacct gcgtcgctcg	60
gcggggtccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcctccgac	120
tggtccatgc tggtcgccgt gatcaccacc atcttctccg ccgcccagaa gcagtggacc	180
aacctggagt ggaagcccaa gcccaccccc cccagctgc tggacgacca cttcggtccc	240
cacggcctgg tggtccgccc caccttcgcc atccgcagct acgaggtggg ccccgaccgc	300
tccaccagca tcgtggccgt gatgaaccac ctgcaggagg ccgccctgaa ccacgccaag	360
tccgtgggca tcctgggcca cggcttcggc accaccctgg agatgtccaa gcgcgacctg	420
atctgggtgg tgaagcgac ccacgtggcc gtggagcgt accccgcctg ggcgcacacc	480
gtggaggtgg agtgctgggt gggcgccctc ggcaacaac gccgccgcca cgacttcctg	540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgatgatg	600
aacacccgca cccgccgcct gagcaagatc cccgaggagg tgcgcggcga gatcggtccc	660
gccttcacgc acaacgtggc cgtgaaggac gaggagatca agaagcccca gaagctgaac	720
gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacatc	780
aaccagcacg tgaacaacat caagtacgtg gactggatcc tggagaccgt gcccgcacgc	840
atcttcgaga gccaccacat ctctctctc accatcgagt accgccgcga gtgcaccatg	900
gacagcgtgc tgcagtcctt gaccaccgtg agcggcggt cctccgaggc cggcctggtg	960
tgcgagcacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcca gaccgagtgg	1020
cgcaccaagc tgaccgactc ctccgcggc atcagcgtga tccccgccga gtccagcgtg	1080
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac	1140
gacgacaagt gatga	1155

<210> 79

<211> 4817

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 79

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt	120
gccatgtgta ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acaggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaagatgc gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcagccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540

32378198_1.TXT

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctggt gcagcccga	1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccgctgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcccctc ttctgctcgc cctctctccg	1320
tcccgctctt agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc	1380
ctgggtggag cgctgtttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgagggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgcccgcct	1860
gaaggccccgc atgcccgcgc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggcgg	2040
cgagtgggccc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgcac	2100
cgccttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgctcc tgcctactgc ccctgcaca gccttggttt gggctccgcc	2460
tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760

32378198_1.TXT

gaatttccttt cttgcgctat gacacttcca gcaaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacgg cgcgccatgg tggccgccgc gcctccagc gccttcttcc	3120
ccgtgccccg ccccgggcgcc tcccccaagc ccggcaagtt cggcaactgg ccctccagcc	3180
tgagcccctc cttcaagccc aagtccatcc ccaacggcgg cttccagggtg aaggccaacg	3240
acagcgccca ccccaaggcc aacggctccg ccgtgagcct gaagagcggc agcctgaaca	3300
cccaggagga cacctcctcc agcccccccc ccgcacactt cctgcaccag ctgcccgact	3360
ggagccgcct gctgaccgcc atcaccaccg tgttcgtgaa gtccaagcgc cccgacatgc	3420
acgaccgcaa gtccaagcgc cccgacatgc tgggtggacag cttcggcctg gagtccaccg	3480
tgcaggacgg cctgggtgtc cgccagtcct tctccatccg ctctacgag atcggcaccg	3540
accgcaccgc cagcatcgag accctgatga accacctgca ggagacctcc ctgaaccact	3600
gcaagagcac cggcatcctg ctggacggct tcggccgcac cctggagatg tgcaagcgcg	3660
acctgatctg ggtgggtgatc aagatgcaga tcaagggtgaa ccgctacccc gcctggggcg	3720
acaccgtgga gatcaacacc cgcttcagcc gcctgggcaa gatcggcatg ggccgcgact	3780
ggctgatctc cgactgcaac accggcgaga tcctgggtgcg cgccaccagc gcctacgcca	3840
tgatgaacca gaagacccgc cgctgtcca agctgcccta cgaggtgcac caggagatcg	3900
tgcccctgtt cgtggacagc cccgtgatcg aggactccga cctgaagggtg cacaagtta	3960
aggtgaagac cggcgacagc atccagaagg gcctgacccc cggctggaac gacctggacg	4020
tgaaccagca cgtgtccaac gtgaagtaca tcggctggat cctggagagc atgcccaccg	4080
aggtgctgga gaccaggag ctgtgctccc tggccctgga gtaccgccgc gagtgcggcc	4140
gcgactccgt gctggagagc gtgaccgcca tggaccccag caagggtggc gtgcgctccc	4200
agtaccagca cctgctgcgc ctggaggacg gcaccgcat cgtgaacggc gccaccgagt	4260
ggcgcccaa gaacgccggc gccaacggcg ccactctcac cggcaagacc agcaacggca	4320
actccgtgtc catggactac aaggaccacg acggcgacta caaggaccac gacatcgact	4380
acaaggacga cgacgacaag tgactcgagg cagcagcagc tcggatagta tcgacacact	4440
ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata	4500
tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc	4560
gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct tccctcgttt	4620
catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc	4680
tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct	4740
ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga	4800
acacaaatgg aaagctt	4817

<210> 80

<211> 4665

<212> DNA

<213> Artificial Sequence

32378198_1.TXT

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 80

ggtagcccgcc	tgcaacgcaa	gggcagccac	agccgctccc	acccgccgct	gaaccgacac	60
gtgcttgggc	gcctgccgcc	tgctgccgc	atgcttgtgc	tggtgaggct	gggcagtgt	120
gccatgtga	ttgaggcttg	gttcacggg	tggaagctta	tgtgtgtgt	gggcttgc	180
gccgggcaat	gcgcagtgtg	gcaagagggc	ggcagcactt	gctggagctg	ccgcggtgcc	240
tccagggtgt	tcaatcgcg	cagccagagg	gatttcagat	gatcgcgct	acaggttgag	300
cagcagtgtc	agcaaaggta	gcagtttgcc	agaatgatcg	gttcagctgt	taatcaatgc	360
cagcaagaga	aggggtcaag	tgcaaacacg	ggcatgccac	agcacgggca	ccggggagt	420
gaatggcacc	accaagtgtg	tgcgagccag	catcgccgcc	tggctgtttc	agctacaacg	480
gcaggagtca	tccaacgtaa	ccatgagctg	atcaacactg	caatcatcgg	gcgggcgtga	540
tgcaagcatg	cctggcggaag	acacatgggtg	tgcggtatgct	gccggctgct	gcctgtgcg	600
cacgccgttg	agttggcagc	aggctcagcc	atgcactgga	tggcagctgg	gctgccactg	660
caatgtgggtg	gataggatgc	aagtggagcg	aataccaaac	cctctggctg	cttgcgtgggt	720
tgcatggcat	cgcaccatca	gcaggagcgc	atgcgaaggg	actggcccca	tgcacgcat	780
gccaaaccgg	agcgcaccga	gtgtccacac	tgccaccagg	cccgcaagct	ttgcagaacc	840
atgctcatgg	acgcatgtag	cgctgacgtc	ccttgacggc	gctcctctcg	ggtgtgggaa	900
acgcaatgca	gcacaggcag	cagaggcggc	ggcagcagag	cggcggcagc	agcggcgggg	960
gccacccttc	ttgcggggtc	gcgccccagc	cagcgggtgat	gcgctgatcc	caaacgagtt	1020
cacattcatt	tgcatgcctg	gagaagcgag	gctggggcct	ttgggctggt	gcagcccgc	1080
atggaatgcg	ggaccgccag	gctagcagca	aaggcgccctc	ccctactccg	catcgatgtt	1140
ccatagtga	ttggactgca	tttgggtggg	gcggccggct	gtttctttcg	tgttgcaaaa	1200
cgcgccagct	cagcaacctg	tcccgtgggt	cccccgctgc	gatgaaatcg	tgtgcacgcc	1260
gatcagctga	ttgcccggct	cgcgaaagtag	gcgcccctct	ttctgctcgc	cctctctccg	1320
tcccgctct	agaatatcaa	tgatcgagca	ggacggcctc	cacgccgggt	ccccgcgcg	1380
ctgggtggag	cgctgttcg	gctacgactg	ggcccagcag	accatcggct	gctccgacgc	1440
cgccgtgttc	cgctgtccg	cccagggccg	ccccgtgctg	ttcgtgaaga	ccgacctgtc	1500
cggcgcctg	aacgagctgc	aggacgaggc	cgcccgctg	tcctggctgg	ccaccaccgg	1560
cgtgccctgc	gccgccgtgc	tggacgtgg	gaccgaggcc	ggccgcgact	ggctgctgct	1620
gggcgagggtg	cccggccagg	acctgctgtc	ctccacactg	gccccgcg	agaagggtgc	1680
catcatggcc	gacgccatgc	gccgcctgca	caccctggac	cccgccacct	gccccctcga	1740
ccaccaggcc	aagcaccgca	tcgagcgcg	ccgcacccgc	atggaggccg	gcctggtgga	1800
ccaggacgac	ctggacgagg	agcaccagg	cctggcccc	gccgagctgt	tcgcccgcct	1860
gaaggccgc	atgcccagc	gcgaggacct	ggtggtgacc	cacggcgacg	cctgcctgcc	1920
caacatcatg	gtggagaacg	gccgcttctc	cggttctatc	gactgcggcc	gcctgggcgt	1980
ggccgaccgc	taccaggaca	tcgccctggc	caccgcgac	atcgccgagg	agctgggcgg	2040
cgagtgggcc	gaccgcttcc	tgggtgctga	cggcacgcgc	gccccgact	cccagcgc	2100

32378198_1.TXT

cgcccttctac cgccctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
 cgcgcttttg cgagttgcta gctgcttggt ctatttgcca ataccacccc cagcatcccc 2340
 ttccctcggt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc 2400
 ctcagcgctg ctctgctcc tgctcactgc ccctcgaca gccttggttt gggctccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
 gaaggctcgc cctctgctgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttggtt cttcgtccat tagcgaagcg tccgggtcac acacgtgcca cgttggcgag 2700
 gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatc 2760
 gaattccttt cttgcgctat gacacttcca gcaaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
 cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 3060
 cagtcacaac ccgcaaacac tagtatgacg ttcggggctc ccctcccggc catgggcccgc 3120
 ggtgtctccc ttccccggcc cagggtcgcg gtgcgcgccc agtcggcgag tcaggttttg 3180
 gagagcgggc gcgcccagct gcccgactgg agccgcctgc tgaccgccat caccaccgtg 3240
 ttcgtgaagt ccaagcgcgc cgacatgcac gaccgcaagt ccaagcgcgc cgacatgctg 3300
 gtggacagct tcggcctgga gtccaccgtg caggacggcc tgggtgtccg ccagtccttc 3360
 tccatccgct cctacgagat cggcaccgac cgcaccgcca gcatcgagac cctgatgaac 3420
 cacctgcagg agacctccct gaaccactgc aagagcaccg gcatcctgct ggacggcttc 3480
 ggccgcaccc tggagatgtg caagcgcgac ctgatctggg tggatgatcaa gatgcagatc 3540
 aagggtgaacc gctaccccg cttggggcgac accgtggaga tcaacacccg cttcagccgc 3600
 ctgggcaaga tcggcatggg ccgcgactgg ctgatctccg actgcaacac cggcgagatc 3660
 ctggtgcgcg ccaccagcgc ctacgccatg atgaaccaga agaccgcgcg cctgtccaag 3720
 ctgccctacg aggtgcacca ggagatcgtg cccctgttcg tggacagccc cgtgatcgag 3780
 gactccgacc tgaagggtgca caagttcaag gtgaagaccg gcgacagcat ccagaagggc 3840
 ctgacccccg gctggaacga cctggacgtg aaccagcacg tgtccaacgt gaagtacatc 3900
 ggctggatcc tggagagcat gccaccgag gtgctggaga cccaggagct gtgctccctg 3960
 gccctggagt accgccgcga gtgcggccgc gactccgtgc tggagagcgt gaccgccatg 4020
 gacccagca aggtgggcgt gcgctcccag taccagcacc tgctgcgcct ggaggacggc 4080
 accgccatcg tgaacggcgc caccgagtgg cgccecaaga acgcccggcg caacggcgcc 4140
 atctccaccg gcaagaccag caacggcaac tccgtgtcca tggactaaa ggaccacgac 4200
 ggcgactaca aggaccacga catcgactac aaggacgacg acgacaagtg actcgaggca 4260
 gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 4320

32378198_1.TXT

ccacacttgc tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt	4380
gtttgatctt gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata	4440
ccacccccag catccccctt cctcgtttca tatcgcttgc atcccaaccg caacttatct	4500
acgctgtcct gctatccctc agcgtgctc ctgctcctgc tcttgcccc tcgcacagcc	4560
ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg	4620
ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agctt	4665

<210> 81

<211> 4668

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 81

ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagggt gggcagtgt	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
gccgggcaat gcgcatgggt gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgct acaggttag	300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcagccac agcacgggca ccggggagt	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttgc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaag acacatgggt tgcggatgct gccggctgct gcctgctgcg	600
cagccggttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
gccaaaccgg agcgaccgga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc cttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcggg	960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccga	1080
atggaatgcg ggaccgccag gctagcagca aaggcgctc ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccgctgcc gatgaaatcg tgtgcacgc	1260
gatcagctga ttgccggct cgcgaaagtag gcgcctcct ttctgctcgc cctctctccg	1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc	1380
ctgggtggag cgctgttcg gctacgactg ggccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgccgcctg tcctggctgg ccaccaccg	1560

32378198_1.TXT

cgTgccctgc gccgccgtgc tggacgtggt gaccgagggcc ggccgcgact ggctgctgct	1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgccccct	1860
gaaggccgc atgcccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggttctatc gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gccccgact ccagcgcacat	2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggaatgt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgctcc tgctcactgc ccctcgaca gccttggtt gggctccgcc	2460
tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag	3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgactgttt cgagcgtcg	3180
gcgcttcgtg ggcgcgcca gctgcccgc tggagccgcc tgctgaccgc catcaccacc	3240
gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg	3300
ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtgtt ccgccagtcc	3360
ttctccatcc gtcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg	3420
aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcctcct gctggacggc	3480
ttcggccgca ccctggagat gtgcaagcgc gacctgatct ggggtggtgat caagatgcag	3540
atcaagggtga accgctaccc cgcctggggc gacaccgtgg agatcaacac ccgcttcagc	3600
cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag	3660
atcctggtgc gcgccaccag cgcctacgcc atgatgaacc agaagacccg ccgcctgtcc	3720
aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgatgc	3780
gaggactccg acctgaaggt gcacaagttc aaggtgaaga ccggcgacag catccagaag	3840

32378198_1.TXT

ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac	3900
atcggtctgga tcctggagag catgcccacc gaggtgctgg agacccagga gctgtgctcc	3960
ctggcccttg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc	4020
atggacccca gcaagggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac	4080
ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc	4140
gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac	4200
gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgactcgag	4260
gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg	4320
ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag	4380
tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttggt ctatttgcca	4440
ataccacccc cagcatcccc ttccctcggt tcatatcgct tgcaccccaa ccgcaactta	4500
tctacgctgt cctgctatcc ctcagcgctg ctctgctcc tgctcactgc ccctcgacac	4560
gccttggttt gggctccgcc tgtattctcc tggactgca acctgtaaac cagcactgca	4620
atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt	4668

<210> 82
 <211> 4668
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 82	
ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtagggt gggcagtgt	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgttgtgtgt gggcttgcac	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccagggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcggt acagggttag	300
cagcagtgtc agcaaaggta gcagtttgcc agaattgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctgta	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc cttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcgggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccga	1080

32378198_1.TXT

atggaatgcg	ggaccgccag	gctagcagca	aaggcgccctc	ccctactccg	catcgatggt	1140
ccatagtgca	ttggactgca	tttgggtggg	gcggccggct	gtttctttcg	tgttgcaaaa	1200
cgcgccagct	cagcaacctg	tcccgtgggt	cccccgtgcc	gatgaaatcg	tgtgcacgcc	1260
gatcagctga	ttgcccggct	cgcgaagtag	gcgccctcct	ttctgctcgc	cctctctccg	1320
tcccgcctct	agaatatcaa	tgatcgagca	ggacggccctc	cacgccggct	cccccgccgc	1380
ctgggtggag	cgcctgttcg	gctacgactg	ggcccagcag	accatcggct	gctccgacgc	1440
cgccgtgttc	cgcctgtccg	cccagggccg	ccccgtgctg	ttcgtgaaga	ccgacctgtc	1500
cggcgccctg	aacgagctgc	aggacgaggc	cgcccgccctg	tcctggctgg	ccaccaccgg	1560
cgtgccctgc	gccgccgtgc	tggacgtggt	gaccgaggcc	ggccgcgact	ggctgctgct	1620
gggcgagggtg	cccggccagg	acctgctgtc	ctcccacctg	gcccccgccg	agaagggtgtc	1680
catcatggcc	gacgccatgc	gccgcctgca	caccctggac	cccgccacct	gccccttcga	1740
ccaccaggcc	aagcaccgca	tcgagcgcgc	ccgcacccgc	atggaggccg	gcctggtgga	1800
ccaggacgac	ctggacgagg	agcaccaggg	cctggccccc	gccgagctgt	tcgcccgcct	1860
gaaggcccg	atgcccagcg	gcgaggacct	ggtggtgacc	cacggcgacg	cctgcctgcc	1920
caacatcatg	gtggagaacg	gccgcttctc	cggcttcacg	gactgcggcc	gcctgggcgt	1980
ggccgaccgc	taccaggaca	tcgccctggc	caccgcgac	atcgccgagg	agctgggcgg	2040
cgagtgggcc	gaccgcttcc	tgggtctgta	cggcatcgcc	gcccccgact	cccagcgc	2100
cgcttctac	cgctgctgg	acgagttctt	ctgacaattg	gcagcagcag	ctcggatagt	2160
atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	ccgccacact	tgctgccttg	2220
acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	tgtgtttgat	cttgtgtgta	2280
cgcgcttttg	cgagttgcta	gctgcttggt	ctatttgcca	ataccacccc	cagcatcccc	2340
ttccctcggt	tcatatcgct	tgcaccccaa	ccgcaactta	tctacgctgt	cctgctatcc	2400
ctcagcgcgt	ctcctgctcc	tgtcactgc	ccctcgca	gccttggttt	gggctccgcc	2460
tgtattctcc	tgggtactgca	acctgtaaac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaatg	gaggatcccc	cgtctcgaac	agagcgcgca	gaggaacgct	2580
gaaggctctg	cctctgtcgc	acctcagcgc	ggcatacacc	acaataacca	cctgacgaat	2640
gcgcttgggt	cttcgtccat	tagcgaagcg	tccggttcac	acacgtgcca	cgttggcgag	2700
gtggcagggtg	acaatgatcg	gtggagctga	tggtcgaaac	gttcacagcc	tagggatatt	2760
gaattccttt	cttgcgctat	gacacttcca	gcaaaaaggta	gggcgggctg	cgagacggct	2820
tcccggcgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgctgtttaa	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaagccat	attcaaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgactc	cgctaagggg	gcgcctcttc	ctcttcgttt	3060
cagtcacaac	ccgcaaacac	tagtatggcc	accgcatcca	ctttctcggc	gttcaatgcc	3120
cgctgcggcg	acctgcgtcg	ctcggcgggc	tccgggcccc	ggcggccagc	gaggccctc	3180
cccgtgcgcg	ggcgcgcca	gctgcccgc	tggagccgcc	tgctgaccgc	catcaccacc	3240
gtgttcgtga	agtccaagcg	ccccgacatg	cacgaccgca	agtccaagcg	ccccgacatg	3300

32378198_1.TXT

```

ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtgtt ccgccagtcc 3360
ttctccatcc gctcctacga gatcgggcacc gaccgcaccg ccagcatcga gaccctgatg 3420
aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480
ttcggccgca ccctggagat gtgcaagcgc gacctgatct ggggtggtgat caagatgcag 3540
atcaaggtga accgctaccc cgcctggggc gacaccgtgg agatcaacac ccgcttcagc 3600
cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag 3660
atcctggtgc gcgccaccag cgcctacgcc atgatgaacc agaagacccg ccgcctgtcc 3720
aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgatgac 3780
gaggactccg acctgaaggt gcacaagttc aagggtgaaga ccggcgacag catccagaag 3840
ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac 3900
atcggctgga tcctggagag catgcccacc gaggtgctgg agaccagga gctgtgctcc 3960
ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020
atggacccca gcaaggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac 4080
ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc 4140
gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200
gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgaaa gtgactcgag 4260
gcagcagcag ctcgatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg 4320
ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag 4380
tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttggt ctatttgca 4440
ataccacccc cagcatcccc ttccctcgtt tcatatcgct tgcacccaa ccgcaactta 4500
tctacgctgt cctgctatcc ctacgcgtg ctctgctcc tgctactgc ccctgcaca 4560
gccttggttt gggctccgcc tgtattctcc tggactgca acctgtaaac cagcactgca 4620
atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 4668

```

<210> 83

<211> 4656

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 83

```

ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt 120
gccatgctga ttgaggcttg gttcatcggg tggaaactta tgtgtgtgct gggcttgcac 180
gccgggcaat gcgcatgggt gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgc agaagatgc gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagt 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

```

32378198_1.TXT

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcagag gctggggcct ttgggctggt gcagcccgca	1080
atggaatgcg ggaccgccag gctagcagca aaggcgccct ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccggtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcccctct ttctgctcgc cctctctccg	1320
tcccgctctt agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc	1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgcctgttc cgcctgtccg ccagggccg ccccgctgctg ttctgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgctg tcttggtgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgagggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaagggtgc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctgggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggcccg atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgcat	2100
cgccttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgcctgcctg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgctcc tgcctactgc ccctcgaca gccttggttt gggctccgcc	2460
tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820

32378198_1.TXT

tcccggcgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atggggcgctc	cgatgccgct	ccagggcgag	cgctgtttta	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaagccat	attcaaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgcactc	cgctaagggg	gcgcctcttc	ctcttcgttt	3060
cagtcacaac	ccgcaaacac	tagtatggct	tccgcggcat	tcaccatgtc	ggcgtgcccc	3120
gcgatgactg	gcagggcccc	tggggcacgt	cgctccggac	ggccagtcgc	caccgcctg	3180
agggggcgcg	cccccgactg	gagccgcctg	ctgaccgcca	tcaccaccgt	gttcgtgaag	3240
tccaagcgcc	ccgacatgca	cgaccgcaag	tccaagcgcc	ccgacatgct	ggtggacagc	3300
ttcggcctgg	agtcaccagt	gcaggacggc	ctggtgttcc	gccagtcctt	ctccatccgc	3360
tcctacgaga	tcggcaccga	ccgcaccgcc	agcatcgaga	ccctgatgaa	ccacctgcag	3420
gagacctccc	tgaaccactg	caagagcacc	ggcatcctgc	tggacggctt	cggccgcacc	3480
ctggagatgt	gcaagcgcg	cctgatctgg	gtggtgatca	agatgcagat	caaggtgaac	3540
cgctaccccc	cctggggcg	caccgtggag	atcaacaccc	gcttcagccg	cctgggcaag	3600
atcggcatgg	gccgcgactg	gctgatctcc	gactgcaaca	ccggcgagat	cctggtgcgc	3660
gccaccagcg	cctacgccat	gatgaaccag	aagacccgcc	gcctgtccaa	gctgccctac	3720
gaggtgcacc	aggagatcgt	gccccgtgtc	gtggacagcc	ccgtgatcga	ggactccgac	3780
ctgaagggtg	acaagttcaa	ggtgaagacc	ggcgacagca	tccagaaggg	cctgaccccc	3840
ggctggaacg	acctggacgt	gaaccagcac	gtgtccaacg	tgaagtacat	cggctggatc	3900
ctggagagca	tgcccaccga	ggtgctggag	accaggagc	tgtgctccct	ggccctggag	3960
taccgccgcg	agtgcggccg	cgactccgtg	ctggagagcg	tgaccgccat	ggacccccagc	4020
aagggtggcg	tgcgctccca	gtaccagcac	ctgctgcgcc	tggaggacgg	caccgccatc	4080
gtgaacggcg	ccaccgagtg	gcgccccaa	aacgccggcg	ccaacggcgc	catctccacc	4140
ggcaagacca	gcaacggcaa	ctccgtgtcc	atggactaca	aggaccacga	cggcgactac	4200
aaggaccacg	acatcgacta	caaggacgac	gacgacaagt	gactcgaggc	agcagcagct	4260
cggatagtat	cgacacactc	tggacgctgg	tcgtgtgatg	gactgttgcc	gccacacttg	4320
ctgccttgac	ctgtgaatat	ccctgccgct	tttatcaaac	agcctcagtg	tgtttgatct	4380
tgtgtgtacg	cgctttttcg	agttgctagc	tgcttgtgct	atttgcgaa	accaccccc	4440
gcacccccct	ccctcgtttc	atatcgcttg	catcccaacc	gcaacttatc	tacgctgtcc	4500
tgctatccct	cagcgctgct	cctgctcctg	ctcactgccc	ctcgcacagc	cttggtttgg	4560
gctccgcctg	tattctcctg	gtactgcaac	ctgtaaacca	gactgcaat	gctgatgcac	4620
gggaagtagt	gggatgggaa	cacaaatgga	aagctt			4656

<210> 84

<211> 4721

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 84

ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac

60

32378198_1.TXT

gtgcttgggc	gcctgccgcc	tgcttccgc	atgcttgtgc	tggtgaggct	gggcagtgc	120
gccatgctga	ttgaggcttg	gttcatcggg	tggaagctta	tgtgtgtgct	gggcttgc	180
gccgggcaat	gcgcatggtg	gcaagagggc	ggcagcactt	gctggagctg	ccgcggtgcc	240
tccagggtgt	tcaatcgcgg	cagccagagg	gatttcagat	gatcgcgcgt	acaggttgag	300
cagcagtgtc	agcaaaggta	gcagtttgcc	agaatgatcg	gttcagctgt	taatcaatgc	360
cagcaagaga	aggggtcaag	tgcaaacacg	ggcatgccac	agcacgggca	ccggggagtg	420
gaatggcacc	accaagtgtg	tgcgagccag	catcgccgcc	tggctgtttc	agctacaacg	480
gcaggagtca	tccaacgtaa	ccatgagctg	atcaacactg	caatcatcgg	gcgggcgtga	540
tgcaagcatg	cctggcgaag	acacatggtg	tgcggtgctg	gccggctgct	gcctgctgcg	600
cacgccgttg	agttggcagc	aggctcagcc	atgcaactga	tggcagctgg	gctgccactg	660
caatgtggtg	gataggatgc	aagtggagcg	aataccaaac	cctctggctg	cttgctgggt	720
tgcattggcat	cgcaccatca	gcaggagcgc	atgcgaaggg	actggcccca	tgacgccc	780
gccaaccggg	agcgaccga	gtgtccacac	tgaccaggg	cccgaagct	ttgcagaacc	840
atgctcatgg	acgcatgtag	cgctgacgtc	ccttgacggc	gctcctctcg	ggtgtgggaa	900
acgcaatgca	gcacaggcag	cagaggcggc	ggcagcagag	cgccggcagc	agcggcgggg	960
gccacccttc	ttgcggggtc	gcgccccagc	cagcgggtgat	gcgctgatcc	caaacgagtt	1020
cacattcatt	tgcatgcctg	gagaagcgag	gctggggcct	ttgggctggt	gcagcccgc	1080
atggaatgcg	ggaccgccag	gctagcagca	aaggcgcctc	ccctactccg	catcgatgtt	1140
ccatagtgc	ttggactgca	tttgggtggg	gcggccggct	gtttctttcg	tggtgcaaaa	1200
cgccgcagct	cagcaacctg	tcccgtgggt	ccccgtgcc	gatgaaatcg	tgtgcacgcc	1260
gatcagctga	ttgcccggct	cgcaagtag	gcgcccctct	ttctgctcgc	cctctctccg	1320
tcccgcctct	agaatatcaa	tgatcgagca	ggacggcctc	cacgccggct	ccccgcgcgc	1380
ctgggtggag	cgctgttctg	gctacgactg	ggcccagcag	accatcggct	gctccgacgc	1440
cgccgtgttc	cgctgtccg	cccagggccg	ccccgtgctg	ttcgtgaaga	ccgacctgtc	1500
cgccgccttg	aacgagctgc	aggacgaggc	cgccgccttg	tcctggctgg	ccaccaccgg	1560
cgtgccctgc	gccgccgtgc	tggacgtggt	gaccgaggcc	ggccgcgact	ggctgctgct	1620
gggcgagggtg	cccggccagg	acctgctgtc	ctcccacctg	gccccgcgcg	agaagggtgc	1680
catcatggcc	gacgccatgc	gccgcctgca	caccctggac	cccgccacct	gccccttcga	1740
ccaccaggcc	aagcaccgca	tcgagcgcgc	ccgcacccgc	atggaggccg	gcctggtgga	1800
ccaggacgac	ctggacgagg	agcaccaggg	cctggcccc	gccgagctgt	tcgcccgcct	1860
gaaggccgc	atgcccagc	gcgaggacct	ggtggtgacc	cacggcgacg	cctgcctgcc	1920
caacatcatg	gtggagaacg	gccgcttctc	cggttcatc	gactgcggcc	gcctgggcgt	1980
ggccgaccgc	taccaggaca	tcgccctggc	cacccgcgac	atcgccgagg	agctgggcgg	2040
cgagtgggccc	gaccgcttcc	tggtgctgta	cgccatcgcc	gccccgact	cccagcgc	2100
cgcccttctac	cgctgctgg	acgagttctt	ctgacaattg	gcagcagcag	ctcgatag	2160
atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	ccgccacact	tgctgccttg	2220
acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	tgtgtttgat	cttgtgtgta	2280

32378198_1.TXT

cgcgcttttg	cgagttgcta	gctgcttggtg	ctatttgcga	ataccacccc	cagcatcccc	2340
ttccctcggtt	tcatatcgct	tgcatcccaa	ccgcaactta	tctacgctgt	cctgctatcc	2400
ctcagcgctg	ctcctgctcc	tgctcaactgc	ccctcgacaca	gccttggttt	gggctccgcc	2460
tgtattctcc	tggtactgca	acctgtaaac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaatg	gaggatcccc	cgctctgaac	agagcgcgca	gaggaacgct	2580
gaaggctctg	cctctgtcgc	acctcagcgc	ggcatacacc	acaataacca	cctgacgaat	2640
gcgcttggtt	cttcgtccat	tagcgaagcg	tccggttcac	acacgtgcca	cgttggcgag	2700
gtggcagggtg	acaatgatcg	gtggagctga	tggtcgaaac	gttcacagcc	tagggatatc	2760
gaattccttt	cttgcgctat	gacacttcca	gcaaaaaggta	gggcgggctg	cgagacggct	2820
tcccggcgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgctgtttaa	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaaagccat	attcaaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgactc	cgctaagggg	gcgcctcttc	ctcttcgttt	3060
cagtcacaac	ccgcaaacgg	cgcgccatgg	ccaccaccag	cctggcctcc	gccttctgct	3120
ccatgaaggc	cgtgatgctg	gcccgcgacg	gccgcggcat	gaagccccgc	agctccgacc	3180
tgcagctgcg	cgccggcaac	gccccacct	ccctgaagat	gatcaacggc	accaagttca	3240
gctacaccga	gagcctgaag	cgctgccccg	actggtccat	gctgttcgcc	gtgatcacca	3300
ccatcttcag	cgccgccgag	aagcagtgga	ccaacctgga	gtggaagccc	aagcccaagc	3360
tgccccagct	gctggacgac	cacttcggcc	tgcacggcct	ggtgttcgcg	cgcaccttcg	3420
ccatccgctc	ctacgagggtg	ggccccgacc	gcagcacctc	catcctggcc	gtgatgaacc	3480
acatgcagga	ggccaccctg	aaccacgcca	agagcgtggg	catcctgggc	gacggcttcg	3540
gcaccaccct	ggagatgtcc	aagcgcgacc	tgatgtgggt	ggtgcgccgc	accacgtgg	3600
ccgtggagcg	ctaccccacc	tggggcgaca	ccgtggaggt	ggagtgtctg	atcggcgcca	3660
gcggcaacaa	cggcatgcgc	cgcgacttcc	tggtgcgcga	ctgcaagacc	ggcgagatcc	3720
tgacccgctg	cacctccctg	agcgtgctga	tgaacacccg	cacccgccgc	ctgagcacca	3780
tccccgacga	ggtgcgcggc	gagatcggcc	ccgccttcat	cgacaacgtg	gccgtgaagg	3840
acgacgagat	caagaagctg	cagaagctga	acgactccac	cgccgactac	atccagggcg	3900
gcctgacccc	ccgctggaac	gacctggacg	tgaaccagca	cgtgaacaac	ctgaagtacg	3960
tggcctgggt	gttcgagacc	gtgcccgcga	gcatcttcga	gtcccaccac	atcagctcct	4020
tcaccctgga	gtaccgccgc	gagtgacccc	gcgactccgt	gctgcgcagc	ctgaccaccg	4080
tgagcggcgg	cagctccgag	gccggcctgg	tgtgcgacca	cctgctgcag	ctggagggcg	4140
gcagcgaggt	gctgcgcgcc	cgcaccgagt	ggcgccccaa	gctgaccgac	tccttccgcg	4200
gcatcagcgt	gatccccgcc	gagccccgcg	tgatggacta	caaggaccac	gacggcgact	4260
acaaggacca	cgacatcgac	tacaaggacg	acgacgacaa	gtgatgactc	gaggcagcag	4320
cagctcggat	agtatcgaca	cactctggac	gctggtcgtg	tgatggactg	ttgccgccac	4380
acttgctgcc	ttgacctgtg	aatatccctg	ccgcttttat	caaacagcct	cagtgtgttt	4440
gatcttgtgt	gtacgcgctt	ttgcgagttg	ctagctgctt	gtgctatttg	cgaataccac	4500
ccccagcatc	cccttccctc	gtttcatatc	gcttgcatcc	caaccgcaac	ttatctacgc	4560

32378198_1.TXT

tgctctgcta tccctcagcg ctgctcctgc tcctgctcac tgccccctcgc acagccttgg	4620
tttgggctcc gcctgtattc tcctgggtact gcaacctgta aaccagcact gcaatgctga	4680
tgcacgggaa gtagtgggat gggaacacaa atggaaagct t	4721

<210> 85

<211> 4650

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 85

ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtagggct gggcagtgtc	120
gccatgctga ttgaggcttg gttcatcggg tggaaagctta tgtgtgtgct gggcttgc	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagt	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
gccaaaccgg agcgcaccga gtgtccacac tgtaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcggg	960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccga	1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcccctct ttctgctcgc cctctctccg	1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgcccggct cccccgccgc	1380
ctgggtggag cgctgttctg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaaggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740

32378198_1.TXT

ccaccaggcc aagcaccgca tcgagcgcg	ccgcacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccagg	cctggccccc gccgagctgt tcgcccgcct	1860
gaaggcccg atgcccgacg gcgaggacct	ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc	cggcttcacg gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc	cacccgcgac atcgccgagg agctgggcgg	2040
cgagtgggccc gaccgcttcc tgggtgctgta	cggcatcgcc gccccgact cccagcgcat	2100
cgccttctac cgcctgctgg acgagttctt	ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga	tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa	acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggtg	ctatttgcca ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcacccaa	ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgctg ctctgctcc tgctcactgc	ccctcgaca gccttggttt gggctccgcc	2460
tgtattctcc tgggtactgca acctgtaa	aac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatccc	gcgtctgaac agagcgcgca gaggaacgct	2580
gaaggctctg cctctgctgc acctcagcg	cggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg	tccggttcac acacgtgcca cgttggcgag	2700
gtggcagggtg acaatgatcg gtggagctga	tggtcgaaac gttcacagcc tagggatatt	2760
gaattccttt cttgcgctat gacacttcca	gcaaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc	ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag	cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat	attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc	cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatgacg	ttcggggctc ccctcccggc catgggccgc	3120
ggtgtctccc ttccccggcc cagggtcgcg	gtgcgcgccc agtcggcgag tcaggttttg	3180
gagagcgggc gcgccccga ctggtccatg	ctgttcgccg tgatcaccac catcttcagc	3240
gccgccgaga agcagtggac caacctggag	tgggaagcca agcccaagct gcccagctg	3300
ctggacgacc acttcggcct gcacggcctg	gtgttcgcc gcaccttcgc catccgctcc	3360
tacgagggtg gccccgaccg cagcacctcc	atcctggccg tgatgaacca catgcaggag	3420
gccaccctga accacgcca gagcgtgggc	atcctgggcg acggcttcgg caccaccctg	3480
gagatgtcca agcgcgacct gatgtgggtg	gtgcgccgca cccacgtggc cgtggagcgc	3540
tacccacct ggggcgacac cgtggagggtg	gagtgtgga tcggcgccag cggcaacaac	3600
ggcatgcgcc gcgacttctt ggtgcgcgac	tgcaagaccg gcgagatcct gaccgcgtgc	3660
acctccctga gcgtgctgat gaacacccgc	acccgccgcc tgagcaccat ccccgacgag	3720
gtgcgcggcg agatcgcccc cgccttcac	gacaacgtgg ccgtgaagga cgacgagatc	3780
aagaagctgc agaagctgaa cgactccacc	gccgactaca tccagggcgg cctgaccccc	3840
cgttggaacg acctggacgt gaaccagcac	gtgaacaacc tgaagtacgt ggcctgggtg	3900
ttcgagaccg tgccccgacg catcttcgag	tcccaccaca tcagctcctt caccctggag	3960

32378198_1.TXT
taccgccgcg agtgcacccg cgactccgtg ctgcgcagcc tgaccaccgt gagcggcggc 4020
agctccgagg ccggcctggt gtgcgaccac ctgctgcagc tggagggcgg cagcgagggtg 4080
ctgcgcgccc gcaccgagtg gcgccccaaag ctgaccgact ccttccgcgg catcagcgtg 4140
atccccgccg agccccgcgt gatggactac aaggaccacg acggcgacta caaggaccac 4200
gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata 4260
gtatcgacac actctggacg ctggctcgtgt gatggactgt tgccgccaca cttgctgcct 4320
tgacctgtga atatccctgc cgcttttata aaacagcctc agtgtgtttg atcttgtgtg 4380
tacgcgcttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc 4440
ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgct gtcctgctat 4500
ccctcagcgc tgctcctgct cctgctcact gcccctcgca cagccttgggt ttgggctccg 4560
cctgtattct cctggactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620
tagtgggatg ggaacacaaa tggaaagctt 4650

<210> 86
<211> 4653
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 86
ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtagggt gggcagtgt 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgttgtgtgct gggcttgcac 180
gccgggcaat gcgcatgggt gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcagccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggtgct gccggtgct gcctgctgcg 600
cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcg atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc cttgacggc gtcctctctg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcattgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgc 1080
atggaatgcg ggaccgccag gctagcagca aaggcgctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200

32378198_1.TXT

cgcgccagct	cagcaacctg	tcccgtgggt	ccccctgccc	gatgaaatcg	tgtgcacgcc	1260
gatcagctga	ttgcccggct	cgcgaagtag	gcgcccctct	ttctgctcgc	cctctctccg	1320
tcccgcctct	agaatatcaa	tgatcgagca	ggacggcctc	cacgccggct	cccccgccgc	1380
ctgggtggag	cgctgtttcg	gctacgactg	ggcccagcag	accatcggct	gctccgacgc	1440
cgccgtgttc	cgctgttccg	cccagggccg	ccccgtgctg	ttcgtgaaga	ccgacctgtc	1500
ggcgccctg	aacgagctgc	aggacgaggc	cgcccgcctg	tcctggctgg	ccaccaccgg	1560
cgtgccctgc	gccgccgtgc	tggacgtggg	gaccgaggcc	ggccgcgact	ggctgctgct	1620
gggcgagggtg	cccggccagg	acctgctgtc	ctcccacctg	gcccccgccg	agaagggtgc	1680
catcatggcc	gacgccatgc	gccgcctgca	caccctggac	cccgccacct	gccccttcga	1740
ccaccaggcc	aagcaccgca	tcgagcgcg	ccgcacccgc	atggaggccg	gcctggtgga	1800
ccaggacgac	ctggacgagg	agcaccaggg	cctggccccc	gccgagctgt	tcgcccgcct	1860
gaaggccccg	atgcccgacg	gcgaggacct	ggtggtgacc	cacggcgacg	cctgcctgcc	1920
caacatcatg	gtggagaacg	gccgcttctc	cggtcttcac	gactgcggcc	gcctgggcgt	1980
ggccgaccgc	taccaggaca	tcgccctggc	caccgcgac	atcgccgagg	agctgggcgg	2040
cgagtggggc	gaccgcttcc	tgggtgctga	cggcatcgcc	gccccgact	cccagcgcat	2100
cgcttcttac	cgctgctgg	acgagttctt	ctgacaattg	gcagcagcag	ctcggatagt	2160
atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	ccgccacact	tgctgccttg	2220
acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	tgtgtttgat	cttgtgtgta	2280
cgcgcttttg	cgagttgcta	gctgcttggt	ctatttgcca	ataccacccc	cagcatcccc	2340
ttccctcggt	tcatatcgct	tgcatcccaa	ccgcaactta	tctacgctgt	cctgctatcc	2400
ctcagcgctg	ctcctgctcc	tgctcactgc	ccctcgca	gccttggttt	gggctccgcc	2460
tgtattctcc	tggtagtgca	acctgtaaac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaaatg	gaggatcccc	cgtctcgaac	agagcgcgca	gaggaacgct	2580
gaaggctctg	cctctgtcgc	acctcagcgc	ggcatacacc	acaataacca	cctgacgaat	2640
gcgcttggtt	cttcgtccat	tagcgaagcg	tccggttcac	acacgtgcca	cgttggcgag	2700
gtggcagggtg	acaatgatcg	gtggagctga	tggtcgaaac	gttcacagcc	tagggatatc	2760
gaattccttt	cttgcgctat	gacacttcca	gcaaaaggta	gggcgggctg	cgagacggct	2820
tcccggcgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgctgtttaa	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaagccat	attcaaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgactc	cgctaagggg	gcgcctcttc	ctcttcgttt	3060
cagtcacaac	ccgcaaacac	tagtatgggt	atcaagacga	acaggcagcc	tgtggagaag	3120
cctccgttca	cgatcgggac	gctgcgcaag	gccatccccg	cgcactgttt	cgagcgctcg	3180
gcgcttcgtg	ggcgcgcccc	cgactgggtc	atgctgttcg	ccgtgatcac	caccatcttc	3240
agcgccgccg	agaagcagtg	gaccaacctg	gagtgggaagc	ccaagcccaa	gctgccccag	3300
ctgctggacg	accacttcgg	cctgcacggc	ctggtgttcc	gccgcacctt	cgccatccgc	3360
tcctacgagg	tgggccccga	ccgcagcacc	tccatcctgg	ccgtgatgaa	ccacatgcag	3420
gaggccaccc	tgaaccacgc	caagagcgtg	ggcatcctgg	gcgacggctt	cggcaccacc	3480

32378198_1.TXT

```

ctggagatgt ccaagcgcgga cctgatgtgg gtggtgccc gcacccacgt ggccgtggag      3540
cgctacccca cctggggcgga caccgtggag gtggagtgtt ggatcggcgc cagcggcaac      3600
aacggcatgc gccgcgactt cctggtgccc gactgcaaga ccggcgagat cctgacccgc      3660
tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac      3720
gaggtgcccg gcgagatcgg ccccgccctt atcgacaacg tggccgtgaa ggacgacgag      3780
atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc      3840
ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg      3900
gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcacctg      3960
gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcccga gcctgaccac cgtgagcggc      4020
ggcagctccg aggccggcct ggtgtgcccg caccgtgtgc agctggaggg cggcagcag      4080
gtgctgcccg cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc      4140
gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac      4200
cacgacatcg actacaagga cgacgacgac aagtgtgac tcgaggcagc agcagctcgg      4260
atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg      4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctgagtgtgt ttgatcttgt      4380
gtgtacgccc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca      4440
tccccctccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc      4500
tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct      4560
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg      4620
aagtagtggg atgggaacac aaatggaaag ctt                                     4653

```

<210> 87

<211> 4653

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 87

```

ggtagccgcc tgcaacgcaa gggcagccac agccgtccc acccgccgct gaaccgacac      60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagggct gggcagtgt      120
gccatgctga ttgaggcttg gttcatcggg tgggaagctta tgtgtgtgct gggcttgcac      180
gccgggcaat gcgcatgggt gcaagagggc ggcagcactt gctggagctg ccgcgggtgcc      240
tccagggtgg tcaatcgccg cagccagagg gatttcagat gatcgcgcgt acaggttgag      300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc      360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg      420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg      480
gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggctgga      540
tgcaagcatg cctggcgaag acacatgggt tgcggatgct gccggctgct gcctgctgccc      600
cacgccgttg agttggcagc aggtcagcc atgactgga tggcagctgg gctgccactg      660
caatgtgggt gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt      720

```

32378198_1.TXT

tgcatggcat	cgcacccatca	gcaggagcgc	atgcgaaggg	actggcccca	tgacgccat	780
gccaaaccgg	agcgcaccga	gtgtccacac	tgccaccagg	cccgaagct	ttgcagaacc	840
atgctcatgg	acgcatgtag	cgctgacgtc	ccttgacggc	gctcctctcg	ggtgtgggaa	900
acgcaatgca	gcacaggcag	cagaggcggc	ggcagcagag	cggcggcagc	agcggcgggg	960
gccacccttc	ttgcggggtc	gcgccccagc	cagcggtgat	gcgctgatcc	caaacgagtt	1020
cacattcatt	tgcatgcctg	gagaagcgag	gctggggcct	ttgggctggt	gcagcccgc	1080
atggaatgcg	ggaccgccag	gctagcagca	aaggcgcctc	ccctactccg	catcgatgtt	1140
ccatagtgc	ttggactgca	tttgggtggg	gcggccggct	gtttctttcg	tgttgcaaaa	1200
cgcgccagct	cagcaacctg	tcccgtgggt	ccccgtgcc	gatgaaatcg	tgtgcacgcc	1260
gatcagctga	ttgcccggt	cgcgaagtag	gcgcccctct	ttctgctcgc	cctctctccg	1320
tcccgcctct	agaatatcaa	tgatcgagca	ggacggcctc	cacgccggct	ccccgccgc	1380
ctgggtggag	cgctgttcg	gctacgactg	ggcccagcag	accatcggct	gctccgacgc	1440
cgccgtgttc	cgctgtccg	cccagggccg	ccccgtgctg	ttcgtgaaga	ccgacctgtc	1500
cggcgccctg	aacgagctgc	aggacgaggc	cgcgcgcctg	tcctggctgg	ccaccaccgg	1560
cgtgccctgc	gccgccgtgc	tggacgtggt	gaccgaggcc	ggccgcgact	ggctgctgct	1620
gggcgagggtg	cccggccagg	acctgctgtc	ctcccacctg	gccccgcgcg	agaagggtgc	1680
catcatggcc	gacgccatgc	gccgcctgca	cacctggac	cccgccacct	gccccttcga	1740
ccaccaggcc	aagcaccgca	tcgagcgcgc	ccgcacccgc	atggaggccg	gcctggtgga	1800
ccaggacgac	ctggacgagg	agcaccaggg	cctggccccc	gccgagctgt	tcgcccgcct	1860
gaaggcccg	atgcccagc	gcgaggacct	ggtggtgacc	cacggcgacg	cctgcctgcc	1920
caacatcatg	gtggagaacg	gccgcttctc	cggcttcctc	gactgcggcc	gcctgggcgt	1980
ggccgaccgc	taccaggaca	tcgccctggc	caccgcgcac	atgcccgagg	agctgggcgg	2040
cgagtgggcc	gaccgcttcc	tgggtgctga	cggcatcgcc	gccccgact	cccagcgc	2100
cgccttctac	cgctgctgg	acgagttctt	ctgacaattg	gcagcagcag	ctcggatagt	2160
atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	ccgccacact	tgctgccttg	2220
acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	tgtgtttgat	cttgtgtgta	2280
cgcgcttttg	cgagttgcta	gctgcttggt	ctatttgca	ataccacccc	cagcatcccc	2340
ttccctcggt	tcatatcgct	tgcatcccaa	ccgcaactta	tctacgctgt	cctgctatcc	2400
ctcagcgcgt	ctcctgctcc	tgtcactgc	ccctcgaca	gccttggttt	gggctccgcc	2460
tgtattctcc	tggtagctga	acctgtaaac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaatg	gaggatcccc	cgtctcgaac	agagcgcgca	gaggaacgct	2580
gaaggctctg	cctctgtcgc	acctcagcgc	ggcatacacc	acaataacca	cctgacgaat	2640
gcgcttggtt	cttcgtccat	tagcgaagcg	tccggttcac	acacgtgcca	cgttggcgag	2700
gtggcaggtg	acaatgatcg	gtggagctga	tggtcgaaac	gttcacagcc	tagggatatc	2760
gaattccttt	cttgcgctat	gacacttcca	gcaaaaggta	gggcgggctg	cgagacggct	2820
tcccggcgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgctgtttta	atagccaggc	ccccgattgc	2940

32378198_1.TXT

```

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg ggcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 3120
cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gagggcccctc 3180
cccgtgcgcg ggcgcgcccc cgactgggtcc atgctgttcg ccgtgatcac caccatcttc 3240
agcgccgccc agaagcagtg gaccaacctg gagtggaaag ccaagcccaa gctgcccag 3300
ctgctggacg accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc 3360
tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag 3420
gaggccaccc tgaaccacgc caagagcgtg ggcacccctg gcgacggctt cggcaccacc 3480
ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag 3540
cgctacccca cctggggcga caccgtggag gtggagtgtt ggatcggcgc cagcggcaac 3600
aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgcg 3660
tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac 3720
gagggtgcgc gcgagatcgg ccccgccctt atcgacaacg tggccgtgaa ggacgacgag 3780
atcaagaagc tgcagaagct gaacgactcc accgcccact acatccaggg cggcctgacc 3840
ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg 3900
gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcacctg 3960
gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc 4020
ggcagctccg aggcggcctt ggtgtgcgac cacctgctgc agctggaggg cggcagcgag 4080
gtgctgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 4140
gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260
atagtatcga cacactctgg acgctggctg tgtgatggac tgttgccgcc acacttgctg 4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 4380
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 4440
tccccctccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc 4500
tatccctcag cgctgtcctt gtcctgtctc actgcccctc gcacagcctt ggtttgggct 4560
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620
aagtagtggg atgggaacac aaatggaaaag ctt 4653

```

<210> 88

<211> 4647

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 88

```

ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgt 120
gccatgtgta ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

```


32378198_1.TXT

gccgggcaat gcgcatgggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
 tccagggtggt tcaatcgcg gaggcagagg gatttcagat gatcgcgcg acagggttgag 300
 cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcggggcgtga 540
 tgcaagcatg cctggcgaa acacatgggtg tgcggatgct gccggctgct gcctgctgcg 600
 cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtgggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
 tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
 gccaaaccgg agcgcaccga gtgtccacac tgcaccagg cccgcaagct ttgcagaacc 840
 atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt 1020
 cacattcatt tgcatgcctg gagaagcgag gctggggcctt ttgggctggt gcagcccga 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
 ccatagtga ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
 cgcgccagct cagcaacctg tcccgtgggt ccccgctgcc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgcccggct cgcgaagtag gcgcccctct ttctgctcgc cctctctccg 1320
 tcccgctctc agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc 1380
 ctgggtggag cgctgttctg gctacgactg ggcccagcag accatcggtt gctccgacgc 1440
 cgccgtgttc cgctgtctcg cccagggccg ccccgctgctg ttctgtaaga ccgacctgtc 1500
 cggcgccctg aacgagctgc aggacgaggc cgcccgctg tcctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 gggcgagggtg cccggccagg acctgctgtc ctccacctg gcccccgccg agaagggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
 gaaggcccg atgcccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgtttctc cggtttctac gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggcgg 2040
 cgagtgggccc gaccgcttcc tgggtgctga cggcatcgcc gccccgact cccagcgcat 2100
 cgctttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagt 2160
 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
 cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc 2340
 ttccctcggt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
 ctgagcgtg ctctgctcc tgctcactgc ccctgcaca gccttggtt gggctccgcc 2460

32378198_1.TXT

tgtattctcc	tggtactgca	acctgtaa	ac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaatg	gaggatcccc	cg	tctcgaac	agagcgcgca	gaggaacgct	2580
gaaggtctcg	cctctgtcgc	acctcagcgc	gg	catacacc	acaataacca	cctgacgaat	2640
gcgcttggtt	cttcgtccat	tagcgaagcg	tcc	ggttcac	acacgtgcca	cgttggcgag	2700
gtggcaggtg	acaatgatcg	gtggagctga	tggt	cgaaac	gttcacagcc	tagggatatt	2760
gaattccttt	cttgcgctat	gacacitcca	gcaaa	aggta	gggcgggctg	cgagacggct	2820
tcccggcgct	gcatgcaaca	ccgatgatgc	ttc	gaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgct	gtttta	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaagccat	att	caaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgcactc	cgct	aagggg	gcgccctctt	ctcttcgttt	3060
cagtcacaac	ccgcaaacac	tagtatggct	tcc	gcggcat	tcaccatgtc	ggcgtgcccc	3120
gcgatgactg	gcagggcccc	tggggcacgt	cgct	ccggac	ggccagtcgc	cacccgcctg	3180
agggggcgcg	cccccgactg	gtccatgctg	ttc	gccgtga	tcaccacat	cttcagcgcc	3240
gccgagaagc	agtggaccaa	cctggagtgg	aag	cccaagc	ccaagctgcc	ccagctgctg	3300
gacgaccact	tcggcctgca	cgccctgggtg	ttcc	gccgca	ccttcgccat	ccgctcctac	3360
gaggtggggc	ccgaccgcag	cacctccatc	ctgg	ccgtga	tgaaccacat	gcaggaggcc	3420
accctgaacc	acgccaagag	cgtgggcac	ctgg	gcgacg	gcttcggcac	caccctggag	3480
atgtccaagc	gcgacctgat	gtgggtgggtg	cgcc	gcaccc	acgtggccgt	ggagcgctac	3540
cccacctggg	gcgacaccgt	ggagggtggag	tgct	ggatcg	gcgccagcgg	caacaacggc	3600
atgcgccgcg	acttcctggt	gcgcgactgc	aag	accggcg	agatcctgac	ccgctgcacc	3660
tccctgagcg	tgctgatgaa	cacccgcacc	cgcc	gcctga	gcaccatccc	cgacgaggtg	3720
cgcggcgaga	tcggccccgc	cttcacgcac	aac	gtggccg	tgaaggacga	cgagatcaag	3780
aagctgcaga	agctgaacga	ctccaccgcc	gact	acatcc	agggcggcct	gacccccgc	3840
tggaacgacc	tggacgtgaa	ccagcacgtg	aaca	acctga	agtacgtggc	ctgggtgttc	3900
gagaccgtgc	ccgacagcat	cttcgagtc	cacc	acatca	gctccttcac	cctggagtac	3960
cgccgcgagt	gcacccgcga	ctccgtgctg	cg	cagcctga	ccaccgtgag	cggcggcagc	4020
tccgaggccg	gcctgggtgtg	cgaccacctg	ctgc	agctgg	agggcggcag	cgaggtgctg	4080
cgcgcccgca	ccgagtggcg	ccccaaagctg	acc	gactcct	tccgcggcat	cagcgtgatc	4140
cccgcgcgagc	cccgcgtgat	ggactacaag	gacc	acgacg	gcgactacaa	ggaccacgac	4200
atcgactaca	aggacgacga	cgacaagtga	tgact	cgagg	cagcagcagc	tcggatagta	4260
tcgacacact	ctggacgctg	gtcgtgtgat	ggact	gtttgc	cgccacactt	gctgccttga	4320
cctgtgaata	tccctgccgc	ttttatcaaa	cag	cctcagt	gtgtttgatc	ttgtgtgtac	4380
gcgcttttgc	gagttgctag	ctgcttgtgc	tatt	tgcgaa	taccaccccc	agcatcccct	4440
tccctcgttt	catatcgctt	gcateccaac	cg	caacttat	ctacgctgtc	ctgctatccc	4500
tcagcgctgc	tcctgctcct	gctcactgcc	cct	cgacag	ccttggtttg	ggctccgcct	4560
gtattctcct	ggtactgcaa	cctgtaaacc	ag	cactgcaa	tgctgatgca	cggaagtag	4620
tgggatggga	acacaaatgg	aaagctt					4647

32378198_1.TXT

<210> 89

<211> 4721

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 89

```

gggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac      60
gtgcttgggc gcctgccgcc tgccctgccgc atgcttgtgc tggtagaggct gggcagtgtc      120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac      180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc      240
tccagggtggt tcaatcgcgg cagccagagg gatctcagat gatcgcgcg acagggttag      300
cagcagtgtc agcaaaggta gcagtttgcc agaattgatc gttcagctgt taatcaatgc      360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg      420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg      480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga      540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg      600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg      660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt      720
tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat      780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc      840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gtcctctcgc ggtgtgggaa      900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg      960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt     1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggg gcagcccgc      1080
atggaatgca ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt     1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa     1200
cgcgccagct cagcaacctg tcccgtgggt ccccgtgcc gatgaaatcg tgtgcacgcc     1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg     1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgcgcgc     1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc     1440
cgccgtgttc cgcctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc     1500
cggcgccctg aacgagctgc aggacgaggc cgcccgctg tcctggctgg ccaccaccgg     1560
cgtgccctgc gccgccgtgc tggacgtggg gaccgaggcc ggccgcgact ggctgctgct     1620
gggcgagggtg cccggccagg acctgctgtc ctccacactg gccccgccg agaagggtgtc     1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga      1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga      1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct     1860
gaaggcccg c atgcccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc     1920

```

32378198_1.TXT

caacatcatg	gtggagaacg	gccgcttctc	cggtttcatc	gactgcggcc	gcctgggctg	1980
ggccgaccgc	taccaggaca	tcgccctggc	cacccgcgac	atcgccgagg	agctgggctg	2040
cgagtgggccc	gaccgcttcc	tggtgctgta	cggtatcgcc	gccccgact	cccagcgcat	2100
cgctttctac	cgctgctgg	acgagttctt	ctgacaattg	gcagcagcag	ctcggatagt	2160
atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	ccgccacact	tgctgccttg	2220
acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	tgtgtttgat	cttgtgtgta	2280
cgcgcttttg	cgagtgtcta	gctgcttggt	ctatttgcga	ataccacccc	cagcatcccc	2340
ttccctcggt	tcatatcgct	tgcatcccaa	ccgcaactta	tctacgctgt	cctgctatcc	2400
ctcagcgctg	ctcctgctcc	tgctcactgc	ccctcgacac	gccttggttt	gggctccgcc	2460
tgtattctcc	tggtactgca	acctgtaaac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaatg	gaggatcccc	cgtctcgaac	agagcgcgca	gaggaacgct	2580
gaaggtctcg	cctctgtcgc	acctcagcgc	ggcatacacc	acaataacca	cctgacgaat	2640
gcgcttggtt	cttcgtccat	tagcgaagcg	tccggttcac	acacgtgcc	cggtggcgag	2700
gtggcaggtg	acaatgatcg	gtggagctga	tggtcgaaac	gttcacagcc	tagggatata	2760
gaattccttt	cttgcgctat	gacacttcca	gcaaaaggta	gggcgggctg	cgagacggct	2820
tcccgccgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgctgtttta	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaagccat	attcaaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgactc	cgctaagggg	gcgcctcttc	ctcttcgttt	3060
cagtcacaac	ccgcaaacgg	cgcgccatgg	ccaccacctc	cctggcctcc	gccttctgca	3120
gcatgaaggc	cgtgatgctg	gcccgcgacg	gccgcggcat	gaagccccgc	tccagcgacc	3180
tgcagctgcg	cgccggcaac	gcccagacct	ccctgaagat	gatcaacggc	accaagttct	3240
cctacaccga	gagcctgaag	aagctgcccc	actgggtccat	gctgttcgcc	gtgatcacca	3300
ccatcttctc	cgccgccgag	aagcagtgga	ccaacctgga	gtggaagccc	aagcccaacc	3360
ccccccagct	gctggacgac	cacttcggcc	cccacggcct	ggtgttcgcg	cgcaccttcg	3420
ccatccgcag	ctacgagggt	ggccccgacc	gctccaccag	catcgtggcc	gtgatgaacc	3480
acctgcagga	ggccgcccctg	aaccacgcca	agtccgtggg	catcctgggc	gacggcttcg	3540
gcaccaccct	ggagatgtcc	aagcgcgacc	tgatctgggt	ggtgaagcgc	accacgtgg	3600
ccgtggagcg	ctaccccgcc	tggggcgaca	ccgtggaggt	ggagtgctgg	gtgggcgcct	3660
ccggcaacaa	cgcccgccgc	cacgacttcc	tggtgcgcga	ctgcaagacc	ggcgagatcc	3720
tgacccgctg	cacctccctg	agcgtgatga	tgaacacccg	cacccgccgc	ctgagcaaga	3780
tccccgagga	ggtgcgcggc	gagatcgggc	ccgccttcat	cgacaacgtg	gccgtgaagg	3840
acgaggagat	caagaagccc	cagaagctga	acgactccac	cgccgactac	atccagggcg	3900
gcctgacccc	ccgctggaac	gacctggaca	tcaaccagca	cgtgaacaac	atcaagtacg	3960
tggactggat	cctggagacc	gtgcccgcga	gcatcttcga	gagccaccac	atctcctcct	4020
tcaccatcga	gtaccgccgc	gagtgcacca	tggacagcgt	gctgcagtcc	ctgaccaccg	4080
tgagcggcgg	ctcctccgag	gccggcctgg	tgtgcgagca	cctgctgcag	ctggagggcg	4140
gcagcgaggt	gctgcgcgcc	aagaccgagt	ggcgccccaa	gctgaccgac	tccttccgcg	4200

32378198_1.TXT

gcatcagcgt gatccccgcc gagtccagcg tgatggacta caaggaccac gacggcgact	4260
acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgatgactc gaggcagcag	4320
cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac	4380
acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt	4440
gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac	4500
ccccagcatc cccttccctc gtttcatatc gcttgcaccc caaccgcaac ttatctacgc	4560
tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg	4620
tttgggctcc gcctgtattc tcctggact gcaacctgta aaccagcact gcaatgctga	4680
tgcacgggaa gtagtgggat gggaacacaa atggaaagct t	4721

<210> 90

<211> 4650

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 90

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgccctgccgc atgcttgtgc tggtagaggct gggcagtgtc	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
gccggggcaat gcgcatgggt gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgat acaggttag	300
cagcagtgtc agcaaaggta gcagtttgcc agaattgatc gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagt	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttgc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaag acacatgggt tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggcctcagc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcattggat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgc	1080
atggaatgca ggaccgccag gctagcagca aaggcgctc ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccgctgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcccctct ttctgctcgc cctctctccg	1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cagccggct ccccgccgcg	1380

32378198_1.TXT

ctgggtggag	cgctgttcg	gctacgactg	ggcccagcag	accatcggct	gctccgacgc	1440
cgccgtgttc	cgctgtccg	cccagggccg	ccccgtgctg	ttcgtgaaga	ccgacctgtc	1500
cggcgccctg	aacgagctgc	aggacgaggc	cgcccgctg	tcctggctgg	ccaccaccgg	1560
cgtgccctgc	gccgccgtgc	tggacgtggt	gaccgaggcc	ggccgcgact	ggctgctgct	1620
gggcgagggtg	cccggccagg	acctgctgtc	ctcccacctg	gccccgccg	agaagggtgc	1680
catcatggcc	gacgccatgc	gccgcctgca	cacctggac	cccgccacct	gccccttcga	1740
ccaccaggcc	aagcaccgca	tcgagcgcg	ccgcacccgc	atggaggccg	gcctggtgga	1800
ccaggacgac	ctggacgagg	agcaccaggg	cctggcccc	gccgagctgt	tcgcccgcct	1860
gaaggcccg	atgcccgcg	gcgaggacct	ggtggtgacc	cacggcgacg	cctgcctgcc	1920
caacatcatg	gtggagaacg	gccgcttctc	cggttctatc	gactgcggcc	gcctgggcgt	1980
ggccgaccgc	taccaggaca	tcgccctggc	caccgcgac	atcgccgagg	agctgggcgg	2040
cgagtggg	gaccgcttcc	tgggtgctgta	cgccatcgcc	gcccccgact	cccagcgcat	2100
cgcttctac	cgctgtggtg	acgagttctt	ctgacaattg	gcagcagcag	ctcgatagt	2160
atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	ccgccacact	tgctgccttg	2220
acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	tgtgtttgat	cttgtgtgta	2280
cgcgcttttg	cgagttgcta	gctgcttggtg	ctatttgca	ataccacccc	cagcatcccc	2340
ttccctcggt	tcatatcgct	tgcattccaa	ccgcaactta	tctacgctgt	cctgctatcc	2400
ctcagcgctg	ctcctgctcc	tgctcactgc	ccctcgaca	gccttggttt	gggctccgcc	2460
tgtattctcc	tgggtactgca	acctgtaaac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaatg	gaggatcccc	cgtctcgaac	agagcgcgca	gaggaacgct	2580
gaaggctctg	cctctgtcgc	acctcagcgc	ggcatacacc	acaataacca	cctgacgaat	2640
gcgcttggtt	cttcgtccat	tagcgaagcg	tccgggtcac	acacgtgcca	cggtggcgag	2700
gtggcagggtg	acaatgatcg	gtggagctga	tggtcgaaac	gttcacagcc	tagggatattc	2760
gaattccttt	cttgcgctat	gacacttcca	gcaaaaggta	gggcgggctg	cgagacggct	2820
tcccggcgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgctgtttaa	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaagccat	attcaaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgcactc	cgctaagggg	gcgcctcttc	ctcttcgttt	3060
cagtcacaac	ccgcaaacac	tagtatgacg	ttcggggtcg	ccctcccggc	catgggccgc	3120
ggtgtctccc	ttccccggcc	cagggtcgcg	gtgcgcgccc	agtcggcgag	tcaggttttg	3180
gagagcgggc	gcgccccga	ctgggtccatg	ctgttcgccg	tgatcaccac	catcttctcc	3240
gccgccgaga	agcagtggac	caacctggag	tggaaagcca	agcccaaccc	ccccagctg	3300
ctggacgacc	acttcggccc	ccacggcctg	gtgttcgcc	gcaccttcgc	catccgcagc	3360
tacgagggtg	gccccgaccg	ctccaccagc	atcgtggccg	tgatgaacca	cctgcaggag	3420
gccgccctga	accacgcca	gtccgtgggc	atcctgggcg	acggcttcgg	caccacctg	3480
gagatgtcca	agcgcgacct	gatctgggtg	gtgaagcgca	cccacgtggc	cgtggagcgc	3540
taccccgctt	ggggcgacac	cgtggagggtg	gagtgtgtgg	tgggcgcctc	cggcaacaac	3600

32378198_1.TXT

```

ggccgcccgc acgacttcct ggtgcgcgac tgcaagaccg gcgagatcct gacccgctgc 3660
acctccctga gcgatgatgat gaacacccgc acccgccgcc tgagcaagat ccccgaggag 3720
gtgcgcggcg agatcgcccc cgccttcacg gacaacgtgg ccgtgaagga cgaggagatc 3780
aagaagcccc agaagctgaa cgactccacc gccgactaca tccagggcgg cctgaccccc 3840
cgctggaacg acctggacat caaccagcac gtgaacaaca tcaagtacgt ggactggatc 3900
ctggagaccg tgcccagacag catcttcgag agccaccaca tctctcctt caccatcgag 3960
taccgcccgc agtgcaccat ggacagcgtg ctgcagtccc tgaccaccgt gagcggcggc 4020
tcctccgagg ccggcctggt gtgcgagcac ctgctgcagc tggagggcgg cagcgagggtg 4080
ctgcgcgcca agaccgagtg gcgccccaaag ctgaccgact ccttccgcgg catcagcgtg 4140
atccccgccg agtccagcgt gatggactac aaggaccacg acggcgacta caaggaccac 4200
gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata 4260
gtatcgacac actctggacg ctggctcgtg gatggactgt tgccgccaca cttgctgcct 4320
tgacctgtga atatccctgc cgtttttatc aaacagcctc agtgtgtttg atcttgtgtg 4380
tacgcgcttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc 4440
ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgct gtctgtctat 4500
ccctcagcgc tgctcctgct cctgctcact gccctcgcga cagccttggg ttgggctccg 4560
cctgtattct cctggtactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620
tagtgggatg ggaacacaaa tggaaagctt 4650

```

<210> 91

<211> 4653

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 91

```

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgctg tggtagggt gggcagtgt 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccagggtgg tcaatcgagg cagccagagg gatttcagat gatcgcgctg acagggttag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcttgcctgc 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

```

32378198_1.TXT

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcgggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt 1020
 cacattcatt tgcattgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgc 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
 ccatagtgc tggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
 cgcgccagct cagcaacctg tcccgtgggt ccccggtgcc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgcccggct cgcaagtag gcgcccctct ttctgctcg cctctctccg 1320
 tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
 cgccgtgttc cgctgtccg cccagggccg ccccggtgctg ttctgaaga ccgacctgtc 1500
 cggcgccttg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggg gaccgaggcc ggccgcgact ggctgctgct 1620
 gggcgagggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740
 ccaccaggcc aagcaccgca tcgagcgcg ccgcaccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgcccgcct 1860
 gaaggccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg 2040
 cgagtgggccc gaccgcttcc tgggtgctga cggcatgcc gccccgact cccagcgcat 2100
 cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagt 2160
 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgcctg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtga 2280
 cgcgcttttg cgagttgcta gctgcttgtg ctatttgca ataccacccc cagcatcccc 2340
 ttccctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc 2400
 ctcagcgctg ctctgctcc tgctcactgc cctcgcaca gccttggttt gggctccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgt 2580
 gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
 gtggcagggtg acaatgatcg gtggagctga tggctgaaac gttcacagcc tagggatatc 2760
 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
 atgggcgctc cgatgccgct ccaggcgag cgctgtttaa atagccaggc ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
 cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 3060
 cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120

32378198_1.TXT

```

cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg      3180
gcgcttcgtg ggcgcgcccc cgactgggtcc atgctgttcg ccgtgatcac caccatcttc      3240
tccgccgccc agaagcagtg gaccaacctg gagtggaagc ccaagcccaa cccccccag      3300
ctgctggacg accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc      3360
agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag      3420
gaggccgccc tgaaccacgc caagtccgtg ggcatcctgg gcgacggctt cggcaccacc      3480
ctggagatgt ccaagcgcg a cctgatctgg gtggtgaagc gcacccacgt ggccgtggag      3540
cgctaccccc cctggggcg a caccgtggag gtggagtgtt ggggtgggcgc ctccggcaac      3600
aacggccgccc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgcg      3660
tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag      3720
gaggtgcgcg gcgagatcgg ccccgccctt atcgacaacg tggccgtgaa ggacgaggag      3780
atcaagaagc ccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc      3840
ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg      3900
atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcttc cttcaccatc      3960
gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc      4020
ggctcctccg aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag      4080
gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc      4140
gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac      4200
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg      4260
atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg      4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt      4380
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcaataacc acccccagca      4440
tccccctccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc      4500
tatccctcag cgctgtcctt gctcctgtc actgcccctc gcacagcctt ggtttgggct      4560
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg      4620
aagtagtggg atgggaacac aaatggaaag ctt                                     4653

```

<210> 92

<211> 1893

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 92

```

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct      60
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc      120
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc      180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta      240
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt      300
cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc      360

```

32378198_1.TXT

```

cgctgcggcg acctgcgtcg ctgcggcggc tccgggcccc ggcgcccagc gagggcccctc 420
cccgtgcgcg ggcgcgcccc cgactgggtcc atgctgttcg ccgtgatcac caccatcttc 480
tccgccgccc agaagcagtg gaccaacctg gagtggaagc ccaagcccaa cccccccag 540
ctgctggacg accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc 600
agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag 660
gaggccgccc tgaaccacgc caagtccgtg ggcacccctg gcgacggctt cggcaccacc 720
ctggagatgt ccaagcgcg cctgatctgg gtggtgaagc gcacccacgt ggccgtggag 780
cgctaccccc cctggggcga caccgtggag gtggagtgtt ggggtggcgc ctccggcaac 840
aacggccgccc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgcg 900
tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag 960
gaggtgcgcg gcgagatcgg ccccgccctt atcgacaacg tggccgtgaa ggacgaggag 1020
atcaagaagc ccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 1080
ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg 1140
atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcttc cttcaccatc 1200
gagtaccgcc gcgagtgac catggacagc gtgctgcagt ccctgaccac cgtgagcggc 1260
ggctcctccg aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag 1320
gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 1380
gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac 1440
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 1500
atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 1560
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctgagtgtgt ttgatcttgt 1620
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcaataacc acccccagca 1680
tccccctccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 1740
tatccctcag cgctgctcct gtcctgctc actgcccctc gcacagcctt ggtttgggct 1800
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 1860
aagtagtggg atgggaacac aaatggaaaag ctt 1893

```

<210> 93

<211> 1887

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 93

```

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
atgggcgctc cgatgccgct ccaggcgag cgctgtttaa atagccaggc ccccgattgc 180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 300
cagtcacaac ccgcaaacac tagtatggct tccgcggcat tcaccatgtc ggcgtgcccc 360

```

32378198_1.TXT

gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc caccgcctg	420
agggggcgcg ccccgactg gtccatgctg ttcgccgtga tcaccaccat cttctccgcc	480
gccgagaagc agtggacca cctggagtgg aagcccaagc ccaaccccc ccagctgctg	540
gacgaccact tcggccccca cggcctggtg ttccgccgca ccttcgccat ccgcagctac	600
gagggtgggccc ccgaccgctc caccagcatc gtggccgtga tgaaccacct gcaggaggcc	660
gccctgaacc acgccaagtc cgtgggcatc ctggggcgag gcttcggcac caccctggag	720
atgtccaagc gcgacctgat ctgggtggtg aagcgcaccc acgtggccgt ggagcgctac	780
cccgctggg gcgacaccgt ggagggtggag tgctgggtgg gcgcctccgg caacaacggc	840
cgccgccacg acttcctggt gcgcgactgc aagaccggcg agatcctgac ccgctgcacc	900
tccctgagcg tgatgatgaa caccgcacc cgccgcctga gcaagatccc cgaggagggtg	960
cgcgggcgaga tcggccccgc cttcatcgac aacgtggccg tgaaggacga ggagatcaag	1020
aagccccaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc	1080
tggaacgacc tggacatcaa ccagcacgtg aacaacatca agtacgtgga ctggatcctg	1140
gagaccgtgc ccgacagcat cttcgagagc caccacatct cctccttcac catcgagtac	1200
cgccgcgagt gcaccatgga cagcgtgctg cagtccctga ccaccgtgag cggcggctcc	1260
tccgaggccg gcctgggtg cgagcacctg ctgcagctgg agggcggcag cgagggtgctg	1320
cgcgccaaga ccgagtggcg cccaagctg accgactcct tccgcggcat cagcgtgatc	1380
cccgcggagt ccagcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac	1440
atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta	1500
tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga	1560
cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac	1620
gcgcttttgc gaggttgctag ctgcttgtgc tatttgcaa taccaccccc agcatccct	1680
tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc	1740
tcagcgctgc tcctgtcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct	1800
gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag	1860
tgggatggga acacaaatgg aaagctt	1887

<210> 94
 <211> 32
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 94	
gatcagaatt ccgcctgcaa cgcaagggca gc	32

<210> 95
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

32378198_1.TXT

<400> 95
gcatactagt ggcgggacgg agagagggcg

30

<210> 96
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 96
Lys Asp Glu Leu
1

HOSA

or

BLANCO

165A

CA

BLANCO

HOSA EN BLANC

Listado de secuencias

<110> SOLAZYME, INC.

<120> FRACCIONAMIENTO DE LA BIOMASA MICROBIANA QUE CONTIENE ACEITE

<130> 026172-002630PC

<140> PCT/US2010/036238

<141> 2010-05-26

<150> 61/181,252

<151> 2009-05-26

<160> 96

<170> PatentIn versión 3.5

<210> 1

<211> 1187

<212> ADN

<213> Chlorella sp.

<400> 1

gatcagacgg gcctgacctg cgagataatc aagtgtctgt aggcaaccaa ctcagcagct	60
gcttggtgtt gggctctgcag gatagtgttg cagggcccca aggacagcag gggaacttac	120
acctgtccc cgacccagtt ttatggagtg cattgccica agagcctagc cggagcgcta	180
ggctacatac ttgccgcacc ggtatgaggg gatatagtac tcgcactgeg ctgtctagt	240

agatgggcag tgctgccat aaacaactgg ctgctcagcc atttgttggc ggaccattct	300
ggggggggcca gcaatgcctg acttctgggt agggtgaaaa ctgaacaaag actaccaaaa	360
cagaatttct tcttccttgg aggtaagcgc aggccggccc gcctgcgccc acatggcgct	420
ccgaacacct ccatagtgt aaggcgcaa acatggccgg actgtgtca gcactcttc	480
atggccatac aaggatcatg cgagattagt gctgagtaag acactatcac cccatgttg	540
attgaagccg tgactcatg ccaacctgcc cctgggcgta gcagacgtat gccatcatga	600
ccactagccg acatgcgtg tcttttgcca ccaaaacaac tggtaaccg ctgaagtcg	660
tgcgcacac ctccgggagt gagtccggcg actctcccc gccggggcgc gccctacct	720
gggtagggtc gccataccgc cagacccaaa cgacgcagga ggggatggg gtagggaatc	780
ccaaccagcc taaccaagac ggcacctata ataatagggtg gggggactaa cagccctata	840
tcgaagctt tgggtgccia tcttgagaag cagagttgg agtggctgtg tacggtcgac	900
cctaagggtg gtgtgccga gcctgaaaca aagcgtctag cagctgcttc tataatgtg	960
cagccgttgi gtttcagtta tattgtatgc tattgtttgt tctgtctagg gtggcgagg	1020
cccacctact gtggcgggcc attggttgg gcttgaattg cctcaccatc taaggctcga	1080
acgtcactc aaacgccttt gtacaactgc agaactttcc ttggcgctgc aactacagt	1140
tgcaaaccag cacatagcac tcccttcat caccagcag tacaaca	1187

<210> 2
<211> 1414
<212> ADN
<213> Chlorella ellipsoidea

<400> 2	
cgctgegcac cagggccgcc agctcgtga tgcetcca aatcggtcc cccgatttt	60
tgtttcat ctctccacc ttgtggcct tctggccag ggcctcagc tgcagcgca	120
cagaccgtg agctcctgat cagcctcctc aggaggccct ttgacaagca agccctgtg	180
caagccatt cacgggtac cagtgggtg gaggtagatg gggtgaaa ggattgctg	240
gtcgattgt gctcatgaa ttggcatgt catgcatgt cacaatatg caccaggct	300
tggagcaaga gagcatgaat gcctcaggc aggtgaaag ttcctggggg tgaaggca	360
gggccgagga ttgaggagg aaagcatcaa gtcgtcgtc atgctcatgt ttcagtcag	420
agtttgcaa gctcacagga gcagagaca gactggctgc tcagggttg catcgtgtg	480
gtggtggggg ggggggggtt aatacgttac gaaatgact tgaattccc acctcatgcc	540
agcggacca catgctgaa ttcaggcct gtgggtgag aaatgctac tctgccctg	600
ttgtgaggt acttcaggcc gctgagctc aagtcgatg cctgctcgtc ttcagggcc	660
tgcacctctg ggctgaccg ctcagcctc ttcggggca tggagtaggc gccggcagcg	720
ttcatgtccg ggcccagggc agcgggtgtg ccataaatgt cgggtatgt ggggagggg	780

gccgtcgcca caccattgcc gttgctggct gacgcatgca catgtggcct ggctggcacc 840
 ggcagcactg gtcctcagcc agccagcaag tggctgttca ggaaageggc catgttgttg 900
 gtccctgcgc atgtaattcc ccagatcaaa ggagggaaca gcttggattt gatgtagtgc 960
 ccaaccggac tgaatgtgcg atggcaggtc ctttgagtc tcccgaatta ctagcagggc 1020
 actgtgacct aacgcagcat gcccaaccgca aaaaaatgat tgacagaaaa tgaagcgggtg 1080
 tgtcaatatt tgctgtattt attcgtttta atcagcaacc aagttcgaaa cgcaactatc 1140
 gtggatgatca agtgaacctc atcagactta cctcgttcgg caaggaaacg gaggcaccaa 1200
 attccaattt gatattatcg ctgccaagc tagagctgat ctttgggaaa ccaactgcca 1260
 gacagtggac tgtgatggag tgecccgagt ggtggagcct cttegatcgc gttagtcatt 1320
 actaacgtga accctcagtg aagggacat cagaccagaa agaccagatc tctctctoga 1380
 caccgagaga gtgttgcggc agtaggacga caag 1414

<210> 3

<211> 1113

<212> ADN

<213> Botryococcus braunii

<400> 3

aattggaaac cccgcgcaag accgggttgt ttggccgcct gaccggaaag ggggggcctg 60

tcccgaaggg ggtctatctc ttgggggatg tcgggcgcgg aaagtcgatg ttgatggacc 120
 tcttcttcga ccatgtcggg gtcgaggcca agagccgcgt ccatttcgcc gagttcatga 180
 tggaggtgaa tgaccgcac gccaccgaac gcgccaagaa gcgggcgacc gatcgcccc 240
 gtcgctgcag cccttgccga ggaagtccgg ctgctggcgt tcgacgagat gatggtgacg 300
 aacagcccgg acgcgatgat cctgtcgcgg ctgttcaccg cgtgatcga ggccgggggtg 360
 acgategtca ccacctcaa ccggccgccc agggatctct ataagaacgg gctcaaccgc 420
 gagcatttcc tgeccfcat cgcgctgac gaggcgcggc tggacgtgct ggcgctgaac 480
 ggcccgaccg actatcggcg cgaccggctg gggcggctgg acacgtgggt ggtgcccaat 540
 ggccccaagg cgacgattac cttgtcggcg gcgttcttc gcctgaccga ctatccgtc 600
 gaggatgccg cgcatgtgcc ctctgaggac ctgaagggtg gcgggcgcgt gctgaatgtc 660
 cccaaggcgc tgaagggcgt cgcggtcttc tegtcaagc ggttgctgcg cgaagcgcgg 720
 ggggcgcggc actatctggc ggtcgcgcgg ggcttcaca ccgtcactt ggtcggaatc 780
 cccaagctgg gggcggagaa ccgcaacgag gcggggcgct tcgtccagct gatcgacgcg 840
 ctctacgaac ataaggtaaa gctgctgcc gcagccgatg ccagccgcc gaactctatg 900
 aaaccggcga cggccgggtc gagttgagc gcagatcagc cggttggaag agatgcgctc 960
 cgaggattat ctggcccaag gccalggctc ggaggggcct tgatcaggcc ttaatgcact 1020

tegcaacat tategtttaa aatcttaaac tctgtggaat aacgggtccc cgacgcgcga 1080

atacacgtac gtccactacg gagtaggatt gga 1113

<210> 4

<211> 253

<212> ADN

<213> Chlamydomonas reinhardtii

<400> 4

cgcttagaag atttcgataa ggcgccagaa ggagcgcagc caaaccagga tgatgttga 60

tggggtattt gagcactgc aacccttacc cggaagcccc ctggcccaca aaggetaggc 120

gccaatgcaa gcagttcgca tgcagcccct ggagcgggtgc cctcctgata aaccggccag 180

ggggcctatg ttctttactt ttitacaaga gaagtcactc aacatcttaa acggtcttaa 240

gaagtctatc cgg 253

<210> 5

<211> 22

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 5

tgtgaagaa tgagccggcg ac 22

<210> 6
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: cebador sintético

<400> 6
cagtgagcta ttacgcactc 20

<210> 7
<211> 565
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 7
tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60

cttagcgaaa gcgagtctga atagggcgat caaatatttt aatatttaca atttagteat 120

ttttctaga cccgaaccgg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180

aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg 240

```

aaataccagt cgaacccgga gctagctggg tctccccgaa atgcgttgag gcgcagcagt  300

acatctagtc tatctagggg taaagcactg ttctgggtgcg ggctgtgaaa acggtaccaa  360

atcgtggcaa actctgaata ctagaaatga cgggttagta gtgagactgt gggggataag  420

ctccattgtc aagagggaaa cageccagac caccagctaa ggccccaaaa tggtaatgta  480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggt ggcttagaag cagccatcct  540

ttaaagagtg cgtaatagct cactg                                     565

```

```

<210> 8
<211> 546
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

```

```

<400> 8
tgttgaagaa tgagccggcg acttagaaaa cgtggcaagg ttaaggaaac gtatccggag  60

ccgaagcgaa agcaagtctg aacagggcga ttaagtcatt tttctagac ccgaaccgg  120

gtgatctaac catgaccagg atgaagcttg ggtgacacca agtgaaggte cgaaccgacc  180

gatgttgaaa aatcggcgga tgagttgtgg ttagcgggtga aataccagtc gaactcggag  240

```


ctagctgggt cccccgaaa tgcgttgagg cgcagcgggt cataaggctg tctaggggta	300
aagcaactgtt tcggtgcggg ctgcgaaagc ggtaccaaatt cgtggcaaac tctgaatact	360
agatatgcta ttatggggc agtgagacgg tgggggataa gcttcacgt cgagaggga	420
acagcccaga tcactagcta agggcccaaa atgatcgta agtgacaaag gaggtgagaa	480
tgcagaaaca accaggaggt ttgcttagaa gcagccaccc tttaaagagt gcgtaatagc	540
tcactg	546

<210> 9

<211> 565

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 9

tgtgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc	60
cttagcgaag gcgagctga ataggcgcat caaatatttt aatatttaca atttagtc	120
ttttctaga ccgaaccgg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc	180
aagtgaaggt ccgaaccgac cgaatgtgaa aaatcggcgg atgagttgtg gttagcgggt	240
aaataccagt cgaaccggga gctagctggt tctccccgaa atgcgttgag gcgcagcagt	300

acatctagtc tatctagggg taaagcactg ttctgggtgcg ggctgtgaaa acggtaccaa 360

atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420

ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggaggtt ggcttagaag cagccatcct 540

ttaaagagtg cgtaatagct cactg 565

- <210> 10
- <211> 565
- <212> ADN
- <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 10

tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60

cttagcgaag gcgagtctga atagggcgat caaatatatt aatattaca atttagtc 120

ttttctaga cccgaaccgg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180

aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg 240

aaataccagt cgaaccggga gctagctggt tctccccgaa atgcgttgag ggcgagcagt 300

acatctagtc tatctagggg taaagcactg ttctgggtgcg ggctgtgaaa acggtaccaa 360

atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420

ctccattgtc aagagggaag cagcccagac caccagctaa ggccccaaaa tggtaatgta 480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggt ggcttagaag cagccatcct 540

ttaaagagtg cgtaatagct cactg 565

<210> 11

<211> 548

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 11

tgttgaagaa tgagccggcg acttagaaga agtggttgg ttaaggataa ctatccggag 60

ccagagcgaa agcaagtctg aatagggcgc ttaaaggta cttttctag acccgaaacc 120

gggtgatcta accatgacca ggaatgaagct tgggtaacac cacgtgaagg tcgaaccga 180

ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240

agctagctgg ttctccccga aatgcgttga ggcgcagcgg ttatgaggc tgtctagggg 300

taaagcactg ttctgggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360

ctagatatgc tattcatgag ccagtgagac ggtgggggat aagcttcac gtcaagaggg 420

aaacagccca gatcaccagc taaggcccca aaatggctgt taagtggcaa aggaggtgag 480

aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaga gtgcgtaata 540

gctcactg 548

<210> 12

<211> 548

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 12

tgttgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag 60

ccagagcgaa agcaagtctg aatagggcgc ttaaaggtea cttttctag acccgaacce 120

gggtgatcta accatgacca ggatgaagct tggglaacac cacgtgaagg tccgaaccga 180

ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240

agctagctgg ttctccccga aatgcgttga ggcgcagcgg tttatgaggc tgtctagggg 300

taaagcactg ttctgggtcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360

ctagatatgc tatcatgag ccagtgagac ggtgggggat aagcttcac gtcaagaggg 420

aaacagccca gatcaccagc taaggcccca aaatggctgt taagtggcaa aggaggtgag 480

aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540

gctcactg 548

<210> 13

<211> 548

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 13

tgttgaagaa tgagccggcg acttagaaga agtggttgg ttaaggataa ctatccggag 60

ccagagcgaa agcaagtctg aatagggcgc ttaaaggta cttttctag acccgaaccc 120

gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga 180

ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240

agctagctgg ttctccccga aatgcgttga ggcgcagcgg ttatgaggc tgtctagggg 300

taaagcactg ttctgggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360

ctagatatgc tattcatgag ccagtgagac ggtgggggat aagcttcac gtcaagaggg 420

aaacagccca gatcaccagc taaggcccca aaatggtcgt taagtggcaa aggaggtgag 480

aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540

gctcactg 548

<210> 14

<211> 548

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 14

tgttgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag 60

ccagagcgaa agcaagtctg aatagggcgc ttaaaggta cttttctag acccgaaacc 120

gggtgatcta accatgacca ggatgaagct tgggtaacac cacgtgaagg tccgaaccga 180

ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240

agctagctgg ttctccccga aatgcgttga ggcgcagcgg ttatgaggc tgtctagggg 300

taaagcactg ttctgggtcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360

ctagatatgc tattcatgag ccagtgaagc ggtgggggat aagcttcac gtcaagaggg 420

aaacagccca gatcaccagc taaggcccca aaatggctgt taagtggcaa aggaggtgag 480

aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540

gctcactg 548

<210> 15
<211> 565
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 15
tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60

cttagcgaag gcgagtctga atagggcgat caaatatttt aatattaca atttagcat 120

ttttctaga cccgaaccgg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180

aagtgaaggt ccgaaccgac cgalgttgaa aaatcggcgg atgagttgtg gttagcggtg 240

aaataccagt cgaaccggga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300

acatctagtc tatctagggg taaagcactg ttfcggtgcg ggctgtgaaa acggtaccaa 360

atcgtggcaa actctgaata ctagaaatga cgggttagta gtgagactgt gggggataag 420

ctccattgtc aagagggaaa cagccagac caccagctaa ggccccaaaa tggtaatgta 480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggt ggcttagaag cagccatcct 540

ttaagagtg cgtaatagct cactg 565

<210> 16
<211> 541
<212> ADN
<213> Prototheca stagnora

<400> 16	
tgttgaagaa tgagccggcg agttaaaaaa aatggcatgg ttaaagatat ttctctgaag	60
ccatagcgaa agcaagtttt acaagctata gtcatttttt ttagaccgga aaccgagtga	120
tctacccatg atcaggggtga agtgttggtc aaataacatg gagggcccgaa ccgactaatg	180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag	240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaage	300
actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta	360
tttagatata tactagttag accttggggg ataagctcct tggcaaaag ggaaacagcc	420
cagatcacca gttaaggccc caaatgaaa atgatatgta ctaaggacgt gagtatgtca	480
aaacctccag caggtttagct tagaagcagc aatcctttca agagtgcgta atagctcact	540
g	541

<210> 17
<211> 573
<212> ADN
<213> Prototheca moriformis

<400> 17	
tggtgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga	60
aacctaagcg aaagcaagtc ttaatagggc gclaatltaa caaaacatta aataaaatct	120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg	180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt	240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc	300
gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat	360
ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg	420
gggataagct ccatagtcga gagggaaaca gccagacca ccagttaagg ccccaaatg	480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca	540
gccatccttt aaagagtgcg taatagctca ctg	573

- <210> 18
<211> 573
<212> ADN
<213> Secuencia artificial
- <220>
<223> Descripción de la secuencia artificial: polinucleótido sintético
- <400> 18

```

tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactga      60

aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaacatta aataaaatct    120

aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg    180

gtgacaccaa gtggaagtec gaaccgaccg atgttgaaaa atcggcggat gaactgtggt    240

tagtggtgaa ataccagtcg aactcagagc tagctgggtc tccccgaaat gcgttgaggc    300

gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat    360

ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg    420

gggataagct ccatagtcca gagggaaaca gcccagacca ccagttaagg ccccaaatg    480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca    540

gccatccttt aaagagtgcg taatagctca ctg                                  573

```

<210> 19

<211> 565

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 19

```

tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc      60

```

```

cttagcgaaa gcgagctctga atagggcgat caaatatttt aatatttaca atttagtcat    120

ttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc    180

aagtgaaggt ccgaaccgac cgaatgtgaa aaatcggcgg atgagttgtg gttagcggtg    240

aaataccagt cgaacccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt    300

acatctagtc tatctagggg taaagcactg tttegggtgcg ggctgtgaaa acggtaccaa    360

atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag    420

ctccattgtc aagaggggaaa cagcccagac caccagctaa ggccccaaaa ttgtaatgta    480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggtt ggcttagaag cagccatcct    540

ttaaagagtg cgtaatagct cactg                                           565

```

```

<210> 20
<211> 565
<212> ADN
<213> Secuencia artificial

```

```

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

```

```

<400> 20
tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc    60

cttagcgaaa gcgagctctga atagggcgat caaatatttt aatatttaca atttagtcat    120

```

ttttctaga ccgaaccgg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180

aagtgaaggt ccgaaccgac cgaatgttgaa aaatcgggcg atgagttgtg gttagcggtg 240

aaataccagt cgaaccggga gctagctggt tctcccgaa atgcgttgag gcgcagcagt 300

acatctagtc tatctagggg taaagcactg ttccggtgcg ggctgtgaaa acggtaccaa 360

atcgtggcaa actctgaata ctagaaatga cgggttagta gtgagactgt gggggataag 420

ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggt ggcttagaag cagccatcct 540

ttaaagagtg cgtaatagct cactg 565

<210> 21

<211> 546

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 21

tgtgaagaa tgagccggcg acttagaaaa cgtggcaagg ttaaggacat gtatccggag 60

ccgaagcgaa agcaagtctg aatagggcgc ctaagtcatt tttctagac ccgaaccggg 120

gtgatctaac catgaccagg atgaagcttg ggtgacacca agtgaaggtc cgaaccgacc 180

gatgtgaaa aatcggcgga tgagttgtgg ttagcgggtga aataccagtc gaactcggag	240
ctagctgggt cccccgaaa tgcgttgagg cgcagcgggt cataaggctg tctaggggta	300
aagcactgtt tcggtgcggg ctgcgaaagc ggtaccaaat cgtggcaaac tctgaatact	360
agatatgcta ttatgagcc agtgagacgg tgggggataa gcttcacgt cgagaggga	420
acagcccaga tcactagcta aggcccctaa atgacgtta agtgacaaag gaggtgagaa	480
tcagaaaaca accaggaggt ttgcttagaa gcagccaccc tttaaagagt gcgtaatagc	540
tcactg	546

<210> 22

<211> 550

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 22

tgttgaagaa tgagccggcg acttatagga agtggcaggg ttaaggaaga atctccggag	60
cccaagcgaa agcgagctcg aaaagggcga ttgggtcact tcttatggac ccgaacctgg	120
atgatctaata catggccaag ttgaagcatg ggtaacacta tgcgaggac tgaaccacc	180
gatgtgaaa aatcggggga tgagctgtga ttagcgggtga aattccaatc gaattcagag	240

ctagctggat ctccccgaaa tgcgttgagg cgcagcggcg acgatgtcct gtctaagggt 300

agagcgactg tttegggtgcg ggctgcgaaa gcggtacca gtcgtggcaa actccgaata 360

ttaggcaaag gattccgtga gccagtgaga ctgtggggga taagcttcat agtcaagagg 420

gaaacagccc agaccatcag ctaaggcccc taaatggctg ctaagtggaa aaggatgtga 480

gaatgctgaa acaaccagga ggttcgctta gaagcagcta ttccttgaaa gagtgcgtaa 540

tagctcactg 550

<210> 23

<211> 548

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 23

tgttgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag 60

ccagagcgaa agcaagtctg aatagggcgc ttaaaggta cttttctag acccgaacc 120

gggtgatcta accatgacca ggatgaaget tgggtaacac cacgtgaagg tccgaaccga 180

ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240

agctagctgg ttctccccga aatgcgttga ggcgcagcgg ttatgaggc tgtctagggg 300

taaagcactg ttccggcgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360

ctagatatgc tattcatgag ccagtgagac ggtgggggat aagcttcac gtcaagaggg 420

aaacagccca gacaccagc taaggcccca aaatggctgt taagtgcaa aggaggtgag 480

aatgctgaaa caaccaggag gtttgcttag aagcagccac ccttaaaga gtgcgtaata 540

gctcactg 548

- <210> 24
- <211> 556
- <212> ADN
- <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 24

tgttgaagaa tgagccggcg acttataggg ggtggcgtgg ttaaggaagt aatccgaagc 60

caaagcgaaa gcaagtttc aatagagcga ttgtgcacc ccttatggac cgaacccgg 120

gtgatctaac ctgaccagg atgaagctg ggtaacacca agtgaaggtc cgaactcac 180

gacttgaaa aatcgtggga tgagttgggg ttagttggtt aaatgctaata cgaactcgga 240

gctagctggt tctccccgaa atgtgtgag gcgcagcgat taacgaaata tttgtacgg 300

tttaggggta aagcactgtt tcggtgcggg ctgcgaaagc ggtaccaaata cgtggcaaac 360

tctgaatact aagcctgtat accgttagtc agtgagagta taggggataa gctctatact 420

caagagggaa acagcccaga tcaccagcia aggccccaaa atgacagcta agtggcaaag 480

gaggtgaaag tgcagaaaca accaggaggt tcgcttagaa gcagcaaccc tttaaagagt 540

gcgtaatagc tcaactg 556

<210> 25

<211> 548

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 25

tgttgaagaa tgagccggcg acttagaaga agtggcttgg ttaaggataa ctatccggag 60

ccagagcgaa agcaagtctg aatagggcgc ttaaaggta cttttctag acccgaaccc 120

gggtgatcta accatgacca ggatgaaget tgggtaacac cacgtgaagg tccgaaccga 180

ccgatgttga aaaatcggcg gatgagttgt ggtagcggg gaaataccaa tcgaactcgg 240

agctagctgg ttctccccga aatgcgttga ggcgcagcgg ttatgaggc tgtctagggg 300

taaagcactg ttccggtgcg ggctgcgaaa gcggtaccaa atcgtggcaa actctgaata 360

ctagalatgc tattcatgag ccagtgagac ggtgggggat aagcttcac gtcaagaggg 420

aaacagccca gatcaccagc taaggcccca aaatggctgt taagtggcaa aggaggtgag 480

aatgctgaaa caaccaggag gtttgcttag aagcagccac cctttaaaga gtgcgtaata 540

gctcactg 548

<210> 26

<211> 565

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 26

tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60

cttagcgaaa gcgagtctga atagggcgat caaatatttt aatattaca atttagcat 120

ttttctaga cccgaaccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180

aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcggtg 240

aaataccagt cgaaccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300

acatctagtc tatctagggg taaagcactg ttctggtgcg ggctgtgaaa acggtaccaa 360

atcgtggcaa actctgaata ctagaaatga cgggtgtagta gtgagactgt gggggataag 420

ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480

gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggtt ggcttagaag cagccatcct	540
ttaaagagtg cgtaatagct cactg	565
<210> 27	
<211> 573	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> Descripción de la secuencia artificial: polinucleótido sintético	
<400> 27	
tgttgaagaa tgagccggcg acttataggg ggtggcttgg ttaaggacta caatccgaag	60
cccaagcgaa agcaagtttg aagtgtacac acattgtgtg tctagagcga tttgtcact	120
ccttatggac ccgaaccggg gtgactctatt catggccagg atgaagcttg ggtaacacca	180
agtgaaggtc cgaactcctc gatgttgaaa aatcgtggga tgagttgtga ataggggtga	240
aatgccaatc gaactcggag ctagctgggt cccccgaaa tgtgttgagg cgcagcgatt	300
cacgatctaa agtacgggtt aggggtaaag cactgtttcg gtgcgggctg ttaacgcggt	360
accaaactgt ggcaaactaa gaatactaaa cttgtatgcc gtgaatcagt gagactaaga	420
gggataagct tcttagtcaa gagggaaaca gcccatgca ccagctaagg ccccaaatg	480
acagctaagt ggcaaaggag gtgagagtgc agaaacaacc aggaggtttg cttagaagca	540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 28

<211> 573

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 28

tgttgaagaa tgagccggcg acttataggg ggtggcttgg ttaaggacta caatccgaag 60

cccaagcgaa agcaagtttg aagtgtacac acgttgtgtg tctagagcga ttltgtcact 120

ccttatggac ccgaaccggg gtgatctatt catggccagg atgaagcttg ggtaacacca 180

agtgaaggtc cgaactcadc gatgttgaaa aatcgtggga tgagttgtga ataggggtga 240

aatgccaatc gaactcggag ctagctgggt ctecccgaaa tgtgttgagg cgcagcgatt 300

cacgatctaa agtacggttt aggggtaaag cactgtttcg gtgcgggctg ttaacgcggt 360

accaaactgt ggcaaaactaa gaatactaaa ctgttatgcc gtgaatcagt gagactaaga 420

gggataagct tcttagtcaa gagggaaaca gccagatca ccagctaagg ccccaaatg 480

acagctaagt ggcaaaggag gtgagagtgc agaaacaacc aggaggtttg cttagaagca 540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 29

<211> 573

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 29

tgttgaagaa tgagccggcg acttataggg ggtggcttgg ttaaggacta caatccgaag 60

cccaagcgaa agcaagtttg aagtgtacac acattgtgtg tctagagcga tttgtcact 120

ccttatggac ccgaaccggt gtgatctatt catggccagg atgaagcttg ggtaacacca 180

agtgaaggtc cgaactcatc gatgttgaaa aatcgtggga tgagttgtga ataggggtga 240

aatgccaatc gaactcggag ctagctgggt cccccgaaa tgtgttgagg cgcagcgatt 300

cacgatctaa agtacgggtt aggggtaaag cactgtttcg gtgcgggctg ttaacgcggt 360

accaaactgt ggcaaactaa gaatactaaa ctgtatgcc gtgaatcagt gagactaaga 420

gggataagct tcttagtcaa gagggaaaca gccagatca ccagctaagg ccccaaatg 480

acagctaagt ggcaaaggag gtgagagtgc agaaacaacc aggaggttg cttagaagca 540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 30

<211> 546
<212> ADN
<213> Chlorella protothecoides

<400> 30
tggtgaagaa tgagccggcg acttagaaaa cgtggcaagg ttaaggaaac gtatccggag 60

ccgaagcgaa agcaagtctg aacagggcga ttaagtcatt tttctagac ccgaaccggg 120

gtgatctaac catgaccagg atgaagcttg ggtgacacca agtgaaggtc cgaaccgacc 180

gatgttgaaa aatcggcgga tgagtgtgg ttagcgggtga aataccagtc gaactcggag 240

ctagctggtt ctccccgaaa tgcgttgagg cgcagegggt cataaggctg tctaggggta 300

aagcactggt tcggtgcggg ctgcgaaagc ggtaccaaact cgtggcaaac tctgaatact 360

agatatgcta ttatgggcc agtgagacgg tgggggataa gcttcacgtg cgagagggaa 420

acagcccaga tcactagcta aggccccaaa atgatcgfta agtgacaaag gaggtgagaa 480

tgcagaaaca accaggaggt ttgcttagaa gcagccaccc tttaaagagt gcgtaatage 540

tcactg 546

<210> 31
<211> 565
<212> ADN
<213> Chlorella protothecoides

<400> 31

tgttgaagaa tgagccggcg acttagaaaa agtggcgtgg ttaaggaaaa attccgaagc 60
 cttagcgaag gcgagtctga atagggcgat caaatatit aatattaca atttagtcat 120
 ttttctaga cccgaaccg ggtgatctaa ccatgaccag gatgaaactt gggtgatacc 180
 aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcggtg 240
 aaataccagt cgaaccggga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300
 acatctagtc tatctagggg taaagcactg ttccggcg ggctgtgaaa acggtaccaa 360
 atcgtggcaa actctgaata ctagaaatga cgggtgtgta gtgagactgt gggggataag 420
 ctccattgtc aagagggaaa cageccagac caccagctaa ggccccaaaa tggtaatgta 480
 gtgacaaagg aggtgaaaat gcaaacacaa ccaggagggt ggcttagaag cagecatcct 540
 ttaaagagtg cgtaatatgct cactg 565

<210> 32
 <211> 565
 <212> ADN
 <213> Prototheca moriformis

<400> 32
 tgttgaagaa tgagccggcg acttagaaaa ggtggcatgg ttaaggaaat attccgaagc 60
 cgtagcaaaa gcgagtctga atagggcgat aaaatatatt aatattaga atctagtcatt 120
 ttttctaga cccgaaccg ggtgatctaa ccatgaccag gatgaagctt gggtgatacc 180

aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg 240
 aaataccagt cgaacccgga gctagctggg tctcccgaa atgcgttgag gcgcagcagt 300
 acatctagtc tatctagggg taaagcactg ttccggtgcg ggctgtgaga acggtaccaa 360
 atcgtggcaa actctgaata ctagaaatga cgatgtagta gtgagactgt gggggataag 420
 ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa ttgtaatgta 480
 gtgacaaagg aggtgaaaat gcaaatacaa ccaggagggt ggcttagaag cagccatcct 540
 ttaaagagtg cgtaatagct cactg 565

<210> 33

<211> 573

<212> ADN

<213> Prototheca moriformis

<400> 33

tggtgaagaa tgagccggcg acttaaaala aatggcaggc taagagaatt aataactga 60

aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaacatta aataaaatct 120

aaagtcattt atttagacc cgaacctgag tgatcfaacc atggtcagga tgaaacttg 180

gtgacaccaa gtggaagtc gaaccgaccg atgtgaaaa atcggcggat gaactgtggt 240

tagtggtgaa ataccagtcg aactcagagc tagctgggtc tccccgaaat gcgttgaggc 300

gcagcaatat atctcgtcta tctaggggta aagcactgtt icgggtcggg ctatgaaaat 360

ggtaccaaatt cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420

gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca 540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 34

<211> 541

<212> ADN

<213> Prototheca kruegani

<400> 34

tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa tactctggag 60

ccatagcgaa agcaagttta gtaagcttag gtcattcttt ttagaccga aaccgagtga 120

tctacccatg atcagggtga agtggttagta aaataacatg gaggcccgaa ccgactaatg 180

ttgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcag tacaataga ggggtaaagc 300

actgtttctt ttgtgggctt cgaaagtgt acctcaaagt ggcaaactct gaatactcta 360

tttagatatc tactagttag acctgggggg ataagctcct tggcaaaag ggaaacagcc 420

cagatcacca gttaaagccc caaatgaaa atgatatga ctaaggatgt gggtatgtca 480

aaacctccag caggtagct tagaagcagc aatccttca agagtgcgta atagctcact 540

g 541

<210> 35

<211> 573

<212> ADN

<213> Prototheca wickerhamii

<400> 35

tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagattt aataactcga 60

aacctaaagcg aaagcaagtc ttaatagggc gtcaatttaa caaaacttta aataaattat 120

aaagtcattt atttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180

gtgacaccaa gtggaagtc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240

tagtggtgaa ataccagtcg aactcagagc tagctgggtc tccccgaaat gcgttgaggc 300

gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtcggg ctatgaaaat 360

ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420

gggataagct ccatagtcga gagggaaaca gccagacca ccagttaagg ccccaaatg 480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca 540

gccatcctt aaagagtgcg taatagctca ctg 573

<210> 36

<211> 541

<212> ADN

<213> Prototheca moriformis

<400> 36

tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagataa ttctctggag 60

ccatagcgaa agcaagtta acaagctaaa gtcacccttt ttagaccga aaccgagtga 120

tctacccatg atcagggtga agtgttggtg aaataacatg gagggccgaa ccgactaatg 180

gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaa tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300

actgtttctt ttgtgggctt cgaaagtgt acctcaaagt ggcaaactct gaatactcta 360

tttagatc tactagttag acctggggg ataagctcct tggcaaaag ggaaacagcc 420

cagateacca gtaaggccc caaatgaaa atgatatga ctaaggatgt gggatatgta 480

aaacctccag caggttagct tagaagcagc aatccttca agagtgcgta atagctcact 540

g 541

<210> 37

<211> 573

<212> ADN

<213> Prototheca wickerhamii

<400> 37

tggtgaagaa tgagccgtcg acttaaaata aatggcaggc taagagaatt aataactcga 60

aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct 120

aaagtcattt attttagacc cgaacctgag tgaictaacc atggtcagga tgaaacttgg 180

gtgacaccaa gtggaagtc gaaccgaccg atgttgaata atcgccggat gaactgtggt 240

tagtggtgaa ataccagtcg aactcagagc tagctgggtc tccccgaaat gcgttgaggc 300

gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtcggg ctatgaaat 360

ggtaccaa at cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420

gggataagct ccatagtcga gagggaaaca gccagacca ccagttagg ccccaaatg 480

ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 38

<211> 541

<212> ADN

<213> Prototheca moriformis

<400> 38

tggtgaagaa tgagccggcg agttaaaaag agtggcgtgg ttaaagaaaa ttctctggaa 60

ccatagcgaa agcaagtta acaagcttaa gtcactttt ttagaccga aaccgagtga 120

tctacccatg atcaggggtga agtggttgga aaataacatg gaggcccgaa ccgactaatg	180
gtgaaaaatt agcgggatgaa ttgtgggtag gggcgaaaaa ccaatcgaa tcggagttag	240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc	300
actgtttctt ttgtgggctc cgaaagtgt acctcaaagt ggcaaactct gaatactcta	360
tttagatc tactagtgag acctggggg ataagtcct tggcgaag ggaacagcc	420
cagatcacca gtaaggccc caaatgaaa atgatatga ctaaggatgt gattatgca	480
aaacctccag caggttagct tagaagcagc aatccttca agagtgcga atagtcact	540
g	541

- <210> 39
- <211> 541
- <212> ADN
- <213> Prototheca zopfii

<400> 39	
tggtgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa ttctctggag	60
ccatagegaa agcaagttta acaagcttaa gtcactttt ttagaccga aaccgagtga	120
tctacccatg atcaggggtga agtggttgga aaataacatg gaggcccgaa ccgactaatg	180
gtgaaaaatt agcgggatgaa ttgtgggtag gggcgaaaaa ccaatcgaa tcggagttag	240

ctgggtctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaage 300

actgtttctt tcgtgggcct cgaaagtgt acctcaaagt ggcaaactct gaatactcta 360

tttagatac tactagtgag accttggggg ataagctcct tggfcaaaag ggaaacagcc 420

cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca 480

aaacctccag caggttagct tagaagcagc aatccttca agagtgcgta atagctcact 540

g 541

<210> 40

<211> 39

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 40

Met Thr Phe Gly Val Ala Leu Pro Ala Met Gly Arg Gly Val Ser Leu

1 5 10 15

Pro Arg Pro Arg Val Ala Val Arg Ala Gln Ser Ala Ser Gln Val Leu

20 25 30

Glu Ser Gly Arg Ala Gln Leu

35

<210> 41
<211> 40
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 41
Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr
1 5 10 15

Ile Gly Thr Leu Arg Lys Ala Ile Pro Ala His Cys Phe Glu Arg Ser
 20 25 30

Ala Leu Arg Gly Arg Ala Gln Leu
 35 40

<210> 42
<211> 36
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 42
Met Ala Ser Ala Ala Phe Thr Met Ser Ala Cys Pro Ala Met Thr Gly

1 5 10 15

Arg Ala Pro Gly Ala Arg Arg Ser Gly Arg Pro Val Ala Thr Arg Leu
 20 25 30

Arg Gly Arg Ala
 35

<210> 43

<211> 40

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 43

Met Ala Thr Ala Ser Thr Phe Ser Ala Phe Asn Ala Arg Cys Gly Asp
1 5 10 15

Leu Arg Arg Ser Ala Gly Ser Gly Pro Arg Arg Pro Ala Arg Pro Leu
 20 25 30

Pro Val Arg Gly Arg Ala Gln Leu
 35 40

<210> 44

<211> 87

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 44

Met Val Ala Ala Ala Ala Ser Ser Ala Phe Phe Pro Val Pro Ala Pro
1 5 10 15

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
 20 25 30

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val
 35 40 45

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser
 50 55 60

Leu Lys Ser Gly Ser Leu Asn Thr Gln Glu Asp Thr Ser Ser Ser Pro
65 70 75 80

Pro Pro Arg Thr Phe Leu His
 85

<210> 45
<211> 60
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético
<400> 45

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
 20 25 30

Gln Leu Arg Ala Gly Asn Ala Pro Thr Ser Leu Lys Met Ile Asn Gly
 35 40 45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Arg Leu
 50 55 60

<210> 46
<211> 60
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 46

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
 20 25 30

Gln Leu Arg Ala Gly Asn Ala Gln Thr Ser Leu Lys Met Ile Asn Gly
 35 40 45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Lys Leu
 50 55 60

<210> 47

<211> 50

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 47

tgacctaggt gattaattaa ctcgaggcag cagcagctcg gatagtatcg 50

<210> 48

<211> 45

<212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: cebador sintético

<400> 48
 ctacgagctc aagcttcca ttgtgttc catccacta ctcc 45

<210> 49
 <211> 1568
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 49
 gaattccttt ctgcgctat gacactcca gcaaaggta gggcgggctg cgagacgget 60
 tcccggeget gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tcggggetgc 120
 atgggcgctc cgatccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcaact accacttcta 240
 cacaggccac tegagcttgt gatcgactc cgctaagggg ggcctcttc ctctcgttt 300
 cagtcacaac ccgcaaacgg cgcgccatat caatgattga acaagatgga ttgcacgcag 360
 gttctccggc cgcttgggtg gagaggctat tcggetatga ctgggcacaa cagacaatcg 420

gctgctctga tgccgcctg ttccgctgt cagcgcaggg gcgcccggtt cttttgtca 480
 agaccgacct gtccggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc 540
 tggccacgac gggcggttct tgcgcagctg tgctcgacgt tgcactgaa gcgggaaggg 600
 actggctgct atggggcgaa gtgccggggc aggatctct gtcactcac ctgtctctg 660
 ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgtt gatccggcta 720
 cctgccatt cgaccaccaa gcgaaacatc gcacgagcg agcacgtact cggatggaag 780
 ccggtctgt cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac 840
 tgttgccag gctcaaggcg cgcctgccg acggcgagga tcctgtctg acctatggcg 900
 atgctgctt gccgaatac atggtggaaa atggccgctt ttctggattc atcactgtg 960
 gccggctggg tgtggcggac cgctatcagg acatagcgtt ggctaccgt gatattgtg 1020
 aagagcttg cggcgaatgg gctgaccgt tctcgtgtt ttacggatc gccgtcccg 1080
 attgcagcg catcgcttc tctcgcttc ttacgagtt cttetaagat ctgtcgatg 1140
 acaagtgact cgaggcagca gcagctcgga tagtatcgac acactctgga cgctggctgt 1200
 gtgatggact gttgcccca cactgtgtg ctgacctgt gaatatccct gccgcttta 1260
 tcaaacagcc tcagtgtgtt tgatctgtg tgcacgcgt ttgcgagtt gctagctgt 1320
 tgtctattt gcgaatacca ccccagcat ccccttccct cgtttcatt cgcttgcatc 1380

ccaaccgcaa cttatctacg ctgtcctgct atccctcage gctgctcctg ctctctgetca 1440

ctgcccctcg cacagccttg gtttgggctc cgcctgtatt ctctgggtac tgcaacctgt 1500

aaaccagcac tgcaatgctg atgcacggga agtagtgga tggaacaca aatggaaagc 1560

ttgagctc 1568

- <210> 50
- <211> 2571
- <212> ADN
- <213> Secuencia artificial

- <220>
- <223> Descripción de la secuencia artificial: polinucleótido sintético

- <220>
- <221> modificado_base
- <222> (997)..(999)
- <223> a, c, t, g, desconocido u otro

<400> 50

gaattccgcc tgcaacgcaa gggcagccac agecgctccc acccgccgct gaaccgacac 60

gtgcttgggc gctgcccgc tgctgcccgc atgcttgtgc tggtaggct gggcagtgt 120

gccaatgctga ttgaggcttg gttcatcggg tggaagctta tgttgtgtgt gggcttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcgggtgcc 240

tccaggtggt tcaatcgcg gaggcagagg gatttcagat gatcgcggt acaggttgag 300
 cagcagtgtc agcaaaggta gcagtttgc agaagatcg gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacac ggcatgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgc ttgctgttc agctacaacg 480
 gcaggagtca tccaactaac catagctgat caacactgca atcatcggcg gctgatgcaa 540
 gcatcctgca agacacatgc tgtcgatgc tgcgtgctg cctgctgcgc acgccgttga 600
 gttggcagca gtcagccat gcactggatc aggctgggt gccactgcaa tgtgttgat 660
 aggatgcaag tggagcgaat accaaacct ctggtgctt gctgggttc atggcatgc 720
 accatcagca ggagcgcatg cgaagggact ggcccatgc acgcatgcc aaaccggagc 780
 gcaccgagtg tccacactgt caccaggccc gcaagcttg cagaacctg ctcatggacg 840
 catgtagcgc tgacgtccct tgacggcgt cctctcgggt gtgggaaacg caatgcagca 900
 caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcgggggccc accctcttg 960
 cggggtcgcg cccagccag cggatgatgc ctgatcnnc caaacgagtt cacattcatt 1020
 tgcagcctgg agaagcgagg ctggggcctt tgggtgttg cagcccgcaa tggatgcgg 1080
 gaccgccagg ctacgagcaa aggcgcctcc cctactccgc atcgatgtc catagtgeat 1140
 tggactgcat ttgggtgggg cgcccggtg ttcttctgt gttgcaaac gcgccacgtc 1200

agcaacctgt cccgtgggtc ccccgtagcg atgaaatcgt gtgcacgccg atcagctgat 1260
 tgcccgggtc gcgaagtagg cgcctcttt ctgctgccc tctctcgtc ccgccactag 1320
 tggcgcgcca tatcaatgat tgaacaagat ggattgcacg caggttctcc ggccgcttgg 1380
 gtggagaggc tattcggcta tgactgggca caacagacaa tcggctgctc tgatccgcc 1440
 gtgtccggc tgcagcgca ggggcgccc gttcttttg tcaagaccga cctgtccgt 1500
 gccctgaatg aactgcagga cgaggcagcg cggtatcgt ggctggccac gacgggcgtt 1560
 cctgcgcag ctgtgctga cgtgtcact gaagcgggaa gggactggct gctattggc 1620
 gaagtgccgg ggcaggatct cctgtcatct cacttctc ctgccgagaa agtatccatc 1680
 atggctgatg caatgcggcg gctgcatacg ctgatccgg ctacctgcc attegaccac 1740
 caagcgaac atgcacga gcgagcaegt actcggatgg aagccgtct tgcgatcag 1800
 gatgatctgg acgaagagca tcaggggtc gcgccagccg aactgttcgc caggctcaag 1860
 gcgcgcatgc ccgacggcga ggatctctc gtgacccatg gcgatgcctg cttgccgaat 1920
 atcatggtgg aaaatggccg ctttctgga ttcacgact gtggccggct ggtgtggcg 1980
 gaccgtatc aggacatagc gttggtacc cgtgatattg ctgaagagct tggcggcgaa 2040
 tgggctgacc gcttctcgt gctttacgt atcgccgtc ccgattcgca gcgcategcc 2100
 ttctatgcc ttctgacga gttcttctaa gatctgtga tcgacaagt actcgaggca 2160

gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 2220

ccacacttgc tgccttgacc tgtgaataatc cctgccgctt ttatcaaaca gcctcagtgt 2280

gtttgatctt gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata 2340

ccacccccag catccccctc cctcgtttca tatcgcttgc atcccaaccg caacttatct 2400

acgtgtgctc gctatccctc agcgtgctc ctgctcctgc tcaactgccc tcgcacagcc 2460

ttggtttggg ctccgcctgt attctcctgg tactgcaacc tglaaaccag cactgcaatg 2520

ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agcttgagct c 2571

<210> 51

<211> 2550

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<220>

<221> modificado_base

<222> (997)..(999)

<223> a, c, t, g, desconocido u otro

<400> 51

gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgc 120
 gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggttgc 180
 gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc 240
 tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
 cagcagtgc agcaaaggta gcagttgcc agaatgateg gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttc agctacaacg 480
 gcaggagtea tccaactaac catagctgat caaactgca atcatcggcg gctgatgcaa 540
 gcatcctgca agacacatgc tgtcgatgc tgcgtgctg cctgctgcgc acgccgttga 600
 gttggcagca gctcagccat gcaactggatc aggtgggct gccactgcaa tgtggtggat 660
 aggatgcaag tggagcgaat accaaaccct ctggtgctt gctgggttgc atggcatgc 720
 accatcagca ggagcgcgtg cgaagggaact ggccccatgc acgcatgcc aaaccggagc 780
 gcaccgagtg tccactgt caccaggccc gcaagctttg cagaacctg ctcattggacg 840
 catgtagcgc tgacgtccct tgacggcgt cctctcgggt gtgggaaacg caatgcagca 900
 caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcgggggcc accctcttg 960
 cggggtcgcg cccagccag cgggtgatgc ctgatennnc caaacgagtt cacattcatt 1020

tgcagcctgg agaagcgagg ctggggcctt tgggctggg cagcccgcaa tggaatgcgg 1080

gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcac 1140

tggactgcat ttgggtgggg cgcccggtg ttctttcgt gttgcaaac gcgccacgtc 1200

agcaacctgt cccgtgggtc ccccggtccg atgaaatcgt gtgcacgcg atcagctgat 1260

tgcccggtc gcgaagtagg cgcctcttt ctgctgccc ttctccgtc ccgccactag 1320

tggcgcgcca tatcaatgat cgagcaggac ggcctccacg ccggtcccc cgcgcctgg 1380

gtggagcgcc tgttcggcta cgactgggccc cagcagacca tcggtgctc cgacgccgcc 1440

gtgtccgcc tgtccgcca gggccgcccc gtgctgttcg tgaagaccga cctgtccggc 1500

gccctgaacg agctgcagga cgaggccgcc cgctgtctt ggctggccac caccggcgtg 1560

ccctgcgcg ccgtgtgga cgtggtgacc gaggcggccc gcgactggct gctgtgggc 1620

gaggtgccc gccaggacct gctgtctcc cacctggccc ccgccagaa ggtgtccatc 1680

atggccgacg ccatgcgcg cctgcacacc ctggaccccg ccacctgcc cttegaccac 1740

caggccaagc accgcatga gcgcgcccgc acccgcatgg aggcggcct ggtggaccag 1800

gacgacctgg acgaggagca ccagggcctg gccccgcgc agctgttcgc ccgcctgaag 1860

gcccgcacg ccgacggcga ggacctggtg gtgacccacg gcgacgcctg cctgcccaac 1920

atcatggtgg agaacggccg ctctccggc tcatcgact gggccgcct gggcgtggcc 1980

gaccgctacc aggacategc cctggccacc cgcgacatcg ccgaggagct gggcggcgag 2040

tgggcggacc gcttctggt gctgtacggc atcgccgcc ccgactccca gcgcacgcc 2100

ttctaccgcc tgctggacga gttcttctga ctgaggcag cagcagctcg gatagtatcg 2160

acacactctg gacgctggc gtgtgatgga ctgttgcgc cacacttget gccttgacct 2220

gtgaatatec ctgcgcgttt tatcaaacag cctcagtgtg ttgatcttg tgtgtacgcg 2280

cttttgcgag ttgctagctg ctgtgtctat ttgcgaatac cacccccage atccccttc 2340

ctcgttcat atcgcttgea tcccaaccgc aacttatcta cgtgttcctg ctatccctea 2400

gcgctgctcc tgcctctgct cactgcccct cgcacagcct tggtttgggc tccgcctgta 2460

ttctctggt actgcaacct glaaaccage actgcaatgc tgatgcacgg gaagtagtgg 2520

gatgggaaca caaatggaaa gcttgagctc 2550

<210> 52

<211> 1547

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 52

gaattccttt ctgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 cacagggcac tegagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 300
 cagtcacaac ccgcaaacgg cgcgccatat caatgatcga gcaggacggc ctccacgccg 360
 gctccccgc cgctgggtg gaggcgctgt tcggctacga ctgggccag cagaccateg 420
 gctgctccga cgccgccgtg ttccgcctgt ccgccaggg ccgcccggtg ctgltcgtga 480
 agaccgacct gtccggcgcc ctgaacgagc tgcaggacga ggccgccccg ctgtcctggc 540
 tgccaccac cggcgtgccc tgcgcgcgcg tgctggacgt ggtgaccgag gccggccgcg 600
 actggctgct gctgggcgag gtgcccggcc aggacctgct gtctccac ctggcccccg 660
 ccgagaaggt gtccatcatg gccgacgcca tgcgcgcct gcacacctg gaccccgcca 720
 cctgcccctt cgaccaccag gccaaagcacc gcacgagcgc cgcccgacc cgcatggagg 780
 ccggcctggt ggaccaggac gacctggacg aggagcacca gggcctggcc ccgcccagac 840
 tgttcgcccg cctgaaggcc cgcatgcccg acggcgagga cctggtggtg acccacggcg 900
 acgcctgect gcccaacatc atggtggaga acggccgctt ctccggttc atcgactgcg 960
 gccgcctggg cgtggccgac cgctaccagg acatgcacct ggccacccgc gacatgcgcg 1020

aggagctggg cggcgagtgg gccgaccgct tectggtgct gtacggcacc gccgcccccg 1080

actcccagcg catcgcttc taccgcctgc tggacgagtt cttctgactc gaggcagcag 1140

cagctcggat agtatcgaca cactctggac gctggctgtg tgatggactg ttgccgccac 1200

acttctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt 1260

gatcttgtgt gtacgcgctt ttgcgagltg ctagctgctt gtgctatttg cgaataccac 1320

ccccagcacc ccttccctc gtttcatac gcttgcaccc caaccgcaac ttatctacgc 1380

tgctctgcta tcccacagcg ctgctctgc tectgctcac tgcctctcgc acagccttgg 1440

tttgggctcc gctgtattc tectggtact gcaacctgta aaccagcact gcaatgctga 1500

tgacggggaa gtagtgggat gggaacacaa atggaaagct tgagctc 1547

<210> 53

<211> 21

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 53

gccgcgactg gctgctgctg g

<210> 54
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: cebador sintético

<400> 54
aggtcctcgc cgtcgggcat g 21

<210> 55
<211> 2357
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 55
ctttcttgcg ctatgacct tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60

cgctgcatgc aacaccgatg atgcttcgac cccccaage tcttcgggg ctgcatgggc 120

gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc agggcccccga ttgcaaagac 180

attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240

ccactcgage ttgtgatgc actccgctaa gggggcgcct ctctcttc gtctcagtc 300

caaccgcaa acggcgcgcc atatcaatgc ttcttcagge cttcttttt cttcttgctg 360

gttttgctgc caáatcagc gcctctatga cgaacgaaac ctcgataga ccacttgtgc 420
 actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaag 480
 atgccaagtg gcatctgtac ttcaataca acccgaacga tactgtctgg gggacgccat 540
 tgtttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta 600
 tcgctccgaa gaggaacgac tccggagcat tctcgggtc catggtggtt gactacaaca 660
 atactccgg cttttcaac gataccattg acccgagaca acgctgcgtg gccatatgga 720
 cttacaacac accggagtcc gaggagcagt acatctcgtg tagcctggac ggtggataca 780
 ctttacaga gtatcagaag aacctgtgc ttctgcaaa ttgactcag ttccgagatc 840
 cgaaggtctt ttggtacgag ccctcgcaga agtggatcat gacagcggca aagtcacagg 900
 actacaagat cgaaattlac tcgtctgacg acctaaate ctggaagctc gaatccgcgt 960
 tcgcaaacga gggctttctc ggctaccaat acgaatgccc aggcctgata gaggtcccaa 1020
 cagagcaaga tccagcaag tctactggg tgatgtttat ttccattaat ccaggagcac 1080
 cggcaggagg ttctttaat cagtactcg tcggaagctt taacggaact cattcgagg 1140
 catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt 1200
 tcttaatac tgaccgacc tatgggagcg ctcttggeat tgcgtgggct tctaactggg 1260
 agtatccgc attcgttctt acaaaccctt ggaggctctc catgtcgtc gtgaggaaat 1320

tctctctcaa cactgagtac caggccaacc cggaaaccga actcataaac ctgaaagccg 1380
 aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgaacc aacaccacgt 1440
 tgacgaaagc caacagctac aacgtcgatc tttegaatag caccggtaca ctgaatttg 1500
 aactgggtga tgccgtcaat accacccaaa cgatctcgaa gtcgggtgtc gcggacctct 1560
 ccctctgggt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt 1620
 ctgcgtcttc cttcttcctt gatcgcgagg acagcaaagt aaaatttgtt aaggagaacc 1680
 catattttac caacaggatg agcgtaaca accaaccatt caagagcgaa aacgacctgt 1740
 cgtactacaa agtgtatggt ttgcttgatc aaaatatect ggaactctac ttcaacgatg 1800
 gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga 1860
 acatgacgac ggggtgtggat aacctgttct acatcgacaa attccaggtg aggggaagtc 1920
 agtgagatct gtcgatcgac aagctcgagg cagcagcagc tcggatagta tcgacacact 1980
 ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata 2040
 tcctgcegc tttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgttttgc 2100
 gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcateccct tcctcggtt 2160
 catatgcctt gcacccaac cgcaacttat ctacgctgtc ctgtatccc tcagcgtgc 2220
 tcctgtcctt gctcactgcc cctcgacag ccttggtttg ggtccgctt gtattctct 2280

ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga 2340

acacaaatgg aaagctt 2357

<210> 56

<211> 2335

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 56

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60

cgctgcatgc aacaccgatg atgcttcgac cccccaaga tecttcgggg ctgcatgggc 120

gctccgatgc cgctccaggg cgagcgctgt taaatagcc agggccccga ttgcaaagac 180

attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240

ccactcgagc ttgtgatgc actccgctaa gggggcgccct ctctctcttc gtttcagtea 300

caaccgcga acggcgcgcc atgctgctgc aggccttctt gtctctgctg gccggttcg 360

ccgccaagat cagcgctcc atgacgaacg agacgtccga ccgccccctg gtgcacttca 420

cccccaacaa gggctggatg aacgaccca acggcctgtg gtacgacgag aaggacgcca 480

agtggcacct gtacttccag tacaaccga acgacaccgt ctgggggacg cccttgttct 540

gggggcacgc cacgtccgac gacctgacca actgggagga ccagcccatc gccatgccc 600
 cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac aacaacacct 660
 ccggcttctt caacgacacc atcgaccgc gccagcgctg cgtggccatc tggacctaca 720
 acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc tacacctca 780
 ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc gaccgaagg 840
 tcttctgga cgagccctcc cagaagtga tcatgaccgc ggccaagtcc caggactaca 900
 agatcgagat ctactctcc gacgacctga agtctggaa gctggagtc gcgttcgcca 960
 acgagggctt cctcggtac cagtacgagt gccccggcct gatcgaggtc cccaccgagc 1020
 aggaccccag caagtctac tgggtgatgt tcatctccat caaccccggc gccccggccg 1080
 gcggctcctt caaccagtac ttctgggca gctcaacgg caccacttc gaggccttcg 1140
 acaaccagtc ccgctgggtg gacttcggca aggactacta cgccctgcag accttctca 1200
 acaccgacc gacctacggg agcgccctgg gcatcgctg ggccccaac tgggagtact 1260
 ccgcttcgt gccaccaac cctggcgct cctcatgtc cctcgtcgc aagtctccc 1320
 tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag gccgagccga 1380
 tctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc acgttgacga 1440
 aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag ttcgagctgg 1500

tgtacgccgt caacaccacc cagacgatct ccaagtcctg gtgcgcggac ctctccctct 1560

ggttcaaggg cctggaggac cccgaggagt acctccgat gggcttcgag gtgtcccgct 1620

cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag aaccctact 1680

tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac ctgtctact 1740

acaagggtga cggcttgctg gaccagaaca tcttgagct gtacttcaac gacggcgacg 1800

tctgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc gtgaacatga 1860

cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag gtcaagtgt 1920

taattaactc gaggcagcag cagctcggat agtatcgaca cactctggac gctggctgtg 1980

tgatggactg ttgcgccac acttgcgcc ttgacctgtg aatatccctg ccgctttat 2040

caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt 2100

gtgctatttg cgaataccac cccagcacc ccttccctc gtttcataac gcttgcacc 2160

caaccgcaac ttatclacgc tgtcttgcta tccctcagcg ctgctcctgc tectgtcac 2220

tgccctcgc acagccttgg ttgggctcc gcctgtattc tcttggtact gcaacctgta 2280

aaccagcact gcaatgctga tgcacgggaa gtagtgggat gggaacacaa atgga 2335

<210> 57

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 57

ccgccgtgct ggacgtggtg 20

<210> 58

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 58

ggtggcgggg tccaggggtg 20

<210> 59

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 59

cggccggcgg ctcttcaac 20

<210> 60
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: cebador sintético

<400> 60
ggcgctcccg taggtcgggt 20

<210> 61
<211> 1335
<212> ADN
<213> Chlorella sorokiniana

<400> 61
cgctgcaac gcaagggcag ccacagccgc tcccacccgc cgctgaaccg acacgtgctt 60

gggcgcctgc cgctgcctg ccgcatgctt gtgctgggtga ggctgggcag tgctgcatg 120

ctgattgagg cttggttcat cgggtggaag cttatgtgtg tgctgggctt gcatgccggg 180

caatgcgcat ggtggcaaga gggcggcagc acttgctgga gctgccgcgg tgcctccagg 240

tggttcaatc gcggcagcca gagggatttc agatgatcgc gcgtacaggt tgagcagcag 300

tgtcagcaaa ggtagcagtt tgccagaatg atcgggtcag ctgttaatca atgccagcaa 360

gagaaggggt caagtgcaaa cacgggcatg ccacagcacg ggcaccgggg agtggaatgg 420
 caccaccaag tgtgtgcgag ccagcatgcg cgcctggctg ttccagctac aacggcagga 480
 gtcaccaac gtaacatga gctgatcaac actgcaatca tcgggcgggc gtgatgcaag 540
 catgcctggc gaagacacat ggtgtgcgga tgctgccggc tgctgcctgc tgcgcacgcc 600
 gttgagttgg cagcaggctc agccatgcac tggatggcag ctgggctgcc actgcaatgt 660
 ggtggatagg atgcaagtgg agcgaatacc aaacctctg gctgcttgcg gggttgcatg 720
 gcctgcacc atcagcagga gcgcatgcga agggactggc cccatgcacg ccatgcaaaa 780
 ccggagcgca ccgagtgtcc aactgtcac caggcccgca agctttgcag aaccatgctc 840
 atggacgcac gtacgcctga cgtcccttga cggcgctcct ctgggtgtg ggaaacgcaa 900
 tgcagcacag gcagcagagg cggcggcagc agagcggcgg cagcagcggc gggggccacc 960
 ctcttgcgg ggtcgcgcc cagccagcgg tgatgcctg atcccaaag agttcacatt 1020
 catttgcag cctggagaag cgaggtctgg gcccttgggc tggcgcagcc cgcaatggaa 1080
 tgcgggaccg ccaggctagc agcaaaggcg cctcccctac tccgcatga tgtccatag 1140
 tgcatggac tgcatctgg tgggcgggcc ggctgtttct ttctgttgc aaaacgcgcc 1200
 agctcagcaa cctgtccctg gggcccccg tgccgatgaa atcgtgtgca cgccgatcag 1260
 ctgattgcc ggtcgcgaa gtaggcgcc tctttctgc tgcctctc tccgtccgc 1320

cactagtggc gcgcc

1335

<210> 62

<211> 1146

<212> ADN

<213> Umbellularia californica

<400> 62

atggccacca ccagcctggc ctccgccttc tgctecatga aggccgtgat gctggcccg 60

gacggccg cgatgaagcc ccgcagctcc gacctgcagc tgcgcgccgg caacgcccc 120

acctccctga agatgatcaa cggcaccaag ttcagctaca ccgagagcct gaagcgctg 180

cccgactggt ccatgctgtt cgccgtgatc accacatct tcagcgccgc cgagaagcag 240

tggaccaacc tggagtggaa gccaagccc aagctgcccc agctgctgga cgaccactc 300

ggcctgcacg gcctggtgtt ccgcgcacc ttgccatcc gctcctacga ggtgggcccc 360

gaccgcagca cctccatcct ggccgtgatg aaccacatgc aggaggccac cctgaaccac 420

gccaagagcg tgggcatcct gggcgacggc ttggcacca cctggagat gtccaagcgc 480

gacctgatgt ggggtggtgc ccgcacccac gtggccgtgg agcgctaccc cacctggggc 540

gacaccgtgg aggtggagtg ctggatcggc gccagcggca acaacggcat gcgccgcgac 600

ttcctggtgc gcgactgcaa gaccggcgag atcctgaccc gctgcacctc cctgagcgtg 660

ctgatgaaca cccgcacccg ccgctgagc accatccccg acgaggtgcg cggcgagatc 720

ggccccgct tcatgacaa cgtggccgtg aaggacgacg agatcaagaa gctgcagaag	780
ctgaacgact ccaccgccga ctacatccag ggcggcctga cccccgctg gaacgacctg	840
gacgtgaacc agcacgtgaa caacctgaag tacgtggcct ggggtgtcga gaccgtgccc	900
gacagcatct tcgagtccea ccacatcagc tccttcaccc tggagtaccg ccgcgagtgc	960
acccgcgact ccgtgctgcg cagcctgacc accgtgagcg gcggcagctc cgaggccggc	1020
ctgggtgctg accacctgct gcagctggag ggcggcagcg aggtgctgcg cgcgccacc	1080
gagtggcgcc ccaagctgac cgactccttc cgcggcatca gcgtgatecc cgccgagccc	1140
cgcgctg	1146

- <210> 63
- <211> 1146
- <212> ADN
- <213> Cinnamomum camphora

<400> 63	
atggccacca cctccctggc ctccgccttc tgcagcatga aggccgtgat gctggccccg	60
gacggccgcg gcatgaagcc ccgctccagc gacctgcagc tgcgcgccgg caacgccag	120
acctccctga agatgatcaa cggcaccaag ttctctaca ccgagagcct gaagaagctg	180
cccgactggt ccatgctgtt cgccgtgalc accaccatct tctccgccgc cgagaagcag	240

tggaccaacc tggagtggaa gccaagccc aaccccccc agctgctgga cgaccactc 300
 ggccccacg gcctgggtgtt ccgccgcacc ttgccatcc gcagctacga ggtgggcccc 360
 gaccgtcca ccagcatcgt ggccgtgatg aaccacctgc aggaggccgc cctgaaccac 420
 gccaaagtcg tgggcatcct gggcgacggc ttggcacca ccctggagat glccaagcgc 480
 gacctgatct ggggtggtgaa gcgcacccac gtggccgtgg agcgctaccc cgctgggggc 540
 gacaccgtgg aggtggagtg ctgggtgggc gcctccggca acaacggccg ccgccacgac 600
 ttctgggtgc gcgactgcaa gaccggcgag atcctgaccc gctgcacctc cctgagcgtg 660
 atgatgaaca ccgcacccg ccgcctgagc aagatccccg aggaggtgcg cggcgagatc 720
 ggccccgcct tcctcgacaa cgtggccgtg aaggacgagg agatcaagaa gccccagaag 780
 ctgaacgact ccaccgccga ctacatccag ggcggcctga cccccgctg gaacgacctg 840
 gacatcaacc agcagtgaa caacatcaag tacgtggact ggatcctgga gaccgtgccc 900
 gacagcatct tegagagcca ccacatctcc tcttcacca tegagtaccg ccgcgagtgc 960
 accatggaca gcgtgctgca gtccctgacc accgtgagcg gcggctctc cgaggccggc 1020
 ctggtgtgcg agcacctgct gcagctggag ggcggcagcg aggtgctgcg cgccaagacc 1080
 gagtggcgcc ccaagctgac cgactcttc cgcgcatca gcgtgatecc cgccgagtc 1140
 agcgtg 1146

<210> 64
 <211> 72
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de la secuencia artificial: oligoneuclótido sintético

<400> 64
 atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 60
 gacgacaagt ga 72

<210> 65
 <211> 408
 <212> ADN
 <213> Chlorella vulgaris

<400> 65
 ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga 60
 ctgttgccgc cacacttgct gccctgacct gtgaatatcc ctgccgcttt tatcaaacag 120
 cctcagtgtg ttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat 180
 ttggaatac caccceccagc atccccttcc ctgcttccat atcgcttgca tcccaaccgc 240
 aacttaleta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgccct 300
 cgcacagcct tggtttgggc tccgcctgta ttctctgggt actgcaacct gtaaaccagc 360

actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatgga 408

<210> 66
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: cebador sintético

<400> 66
ctgggcgacg gcttcggcac 20

<210> 67
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: cebador sintético

<400> 67
aagtcgcggc gcatgccgtt 20

<210> 68
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 68

taccccgctt gggcgacac 20

<210> 69

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 69

cttgctcagg cggcgggtgc 20

<210> 70

<211> 1317

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 70

atgggtggcgc ccgcgcctc cagcgcttc ttcctcgtgc cgcgcgcgcgc gcgcgcgcgc 60

aagcccggea agtcggcaa ctggccctcc agcctgagcc cctcctcaa gcccaagtc 120

atcccacacg gcggttcca ggtgaaggcc aacgacagcg cccaccccaa ggccaacggc 180
 tccgccgtga gcctgaagag cggcagcctg aacacccagg aggacacctc ctccagcccc 240
 cccccccgca ccttctgca ccagctgccc gactggagcc gcctgctgac cgccatcacc 300
 accgtgttcg tgaagtcaa gcgccccgac atgcacgacc gcaagtcaa gcgccccgac 360
 atgttggtgg acagcttcgg cctggagtc accgtgcagg acggcctggt gtccgccag 420
 tccttctcca tccgtccta cgagatcggc accgaccgca ccgccagcat cgagaccctg 480
 atgaaccacc tgcaggagac ctccctgaac cactgcaaga gcaccggcat cctgctggac 540
 ggcttcggcc gcaccctgga gatgtgcaag cgcgacctga tctgggtggt gatcaagatg 600
 cagatcaagg tgaaccgcta ccccgcttg ggcgacaccg tggagatcaa caccgcttc 660
 agccgcctgg gcaagatcgg catgggccgc gactggctga tctccgactg caacaccggc 720
 gagatcctgg tgcgcgccac cagcgctac gccatgatga accagaagac ccgccgctg 780
 tccaagctgc cctacgaggt gcaccaggag atcgtgcccc tgttcgtgga cagccccgtg 840
 atcgaggact ccgacctgaa ggtgcacaag ttcaaggatga agaccggcga cagcatccag 900
 aagggcctga cccccggctg gaacgacctg gacgtgaacc agcacgtgtc caacgtgaag 960
 tacatcggtt ggatcctgga gagcatgccc accgaggtgc tggagacca ggagctgtgc 1020
 tccttgcccc tggagtaccg ccgcgagtgc ggccgcgact ccgtgctgga gacgtgacc 1080

gccatggacc ccagcaaggt gggcgtgccc tccagtacc agcacctgct gcgcctggag 1140

gacggcaccg ccatcgtgaa cggcgccacc gaggggccc ccaagaacgc cggcgccaac 1200

ggcgccatct ccaccggcaa gaccagcaac ggcaactccg tgcctatgga ctacaaggac 1260

cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtga 1317

<210> 71

<211> 1170

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 71

atggccaccg catccacttt ctggcggttc aatgcccgtc gggcgacct gcgtcgctcg 60

gggggtccg ggccccggcg ccagcgagg cccctccccg tgcggggcg cggccagctg 120

ccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcgcc 180

gacatgcacg accgcaagtc caagcgcgcc gacatgctgg tggacagctt cggcctggag 240

tccaccgtgc aggacggcct ggtgttcgc cagtccttct ccatcgcctc ctacgagatc 300

ggcaccgacc gcaccgccag catcgagacc ctgatgaacc accfgcagga gacctccctg 360

aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcacct ggagatgtgc 420

aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc	480
tggggcgaca cctgggagat caacacccgc ttcagccgcc tgggcaagat cggcattggc	540
cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tgggtgcgcg caccagcgcc	600
tacgcatga tgaaccagaa gaccgcgcgc ctgtccaagc tgcctacga ggtgcaccag	660
gagatcgtgc cctgttcgt ggacagcccc gtgategagg actccgacct gaaggtgcac	720
aagtcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac	780
ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg	840
cccaccgagg tgctggagac ccaggagctg tgctccctgg cctggagta ccgccgcgag	900
tgcggccgcg actccgtgct ggagagcgtg accgccatgg acccagcaa ggtgggcgtg	960
cgtcccagt accagacct gctgcgcctg gaggacggca ccgccatct gaacggcgcc	1020
accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc	1080
aacggcaact cctgtccat ggactacaag gaccagacg gcgactaaa ggaccacgac	1140
atcgactaca aggacgacga cgacaagtga	1170

<210> 72

<211> 1170

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 72

atggctatca agacgaacag gcagcctgtg gagaagcctc cggtcacgat cgggacgctg	60
cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg cgcccagctg	120
cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc	180
gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag	240
tccaccgtgc aggacggcct ggtgttcgc cagtccttct ccatccgctc ctacgagate	300
ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctcctg	360
aaccactgca agagcaccgg catcctgctg gacggttcg gccgcaccct ggagatgtgc	420
aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc	480
tggggcgaca ccgtggagat caacacccgc ttcagccgcc tgggcaagat cggeatgggc	540
cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tggcgcgcc caccagcgcc	600
tacgccatga tgaaccagaa gaccgcgcgc ctgtccaagc tgcctacga ggtgcaccag	660
gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac	720
aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac	780
ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg	840

cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag 900

tgcggccgcg actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg 960

cgctcccagt accagcacct gctgcgcctg gaggacggca ccgccatcgt gaacggcgcc 1020

accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc 1080

aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac 1140

atcgactaca aggacgacga cgacaagtga 1170

<210> 73

<211> 1167

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 73

atgacgttcg gggtcgccct ccgggccatg ggccgcgggtg tctcccttc ccgcccagg 60

gtcgcgggtgc gcgccagtc ggcgagtcag gtttggaga gcgggcgcgc ccagctgcc 120

gactggagcc gcctgtgac cgccatcacc accgtgttcg tgaagtcaa gcgccccgac 180

atgcacgacc gcaagtcaa gcgccccgac atgtgggtgg acagcttcgg cctggagtc 240

accgtgcagg acggcctggt gttccgccag tcttctcca tccgtccta cgagatcggc 300

accgaccgca ccgccagcat cgagaccctg atgaaccacc tgcaggagac ctccctgaac 360
 cactgcaaga gcaccggcat cctgctggac ggcttcggcc gcaccctgga gatgtgcaag 420
 cgcgacctga tctgggtggt gatcaagatg cagatcaagg tgaaccgcta ccccgctgg 480
 ggcgacaccg tggagatcaa caccgcttc agccgcctgg gcaagatcgg catgggccgc 540
 gactggctga tctccgactg caacaccggc gagatcctgg tgcgcgccac cagcgcctac 600
 gccatgatga accagaagac ccgccgcctg tccaagctgc cctacgaggt gcaccaggag 660
 atcgtgcccc tgttcgtgga cagccccgtg atcgaggact ccgacctgaa ggtgcacaag 720
 ttcaaggatga agaccggcga cagcatccag aagggcctga ccccggtg gaacgacctg 780
 gacgtgaacc agcacgtgtc caacgtgaag tacatcggct ggatectgga gagcatgccc 840
 accgaggtgc tggagacca ggagctgtgc tccctggccc tggaglacg ccgcgagtgc 900
 ggccgcgact ccgtgtgga gagcgtgacc gccatggacc ccagcaaggt gggcgtgcgc 960
 tccagtagc agcacctgtc gcgcctggag gacggcaccg ccatcgtgaa cggcgccacc 1020
 gactggcgcc ccaagaacgc cggcgccaac ggcgccatct ccaccggcaa gaccagcaac 1080
 ggcaactccg tgtccatgga ctacaaggac cagcagggcg actacaagga ccacgacatc 1140
 gactacaagg acgacgacga caagtga 1167

<210> 74

<211> 1149

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 74

atgacgttcg gggtoeccct cccggccatg ggccgcgggtg tctcccttc cggccccagg 60

gtcgcgggtgc ggcgccagtc ggcgagtcag gttttggaga gcgggcgcgc ccccgactgg 120

tccatgctgt tcgccgtgat caccaccatc ttcagcgccg ccgagaagca gtggaccaac 180

ctggagtgga agcccaagcc caagctgcc cagctgctgg acgaccactt cgccctgcac 240

ggcctgggtgt tcgcgcgcac ctteggcatc cgtctctacg aggtggggccc cgaccgcagc 300

acctecatcc tggccgtgat gaaccacatg caggaggcca ccctgaacca cgccaagagc 360

gtgggcatcc tgggcgacgg ctteggcacc acctggaga tgcgaagcg cgacctgatg 420

tgggtgggtgc gccgcaccca cgtggccgtg gagegctacc ccacctgggg cgacaccgtg 480

gaggtggagt gctggatcgg cgccagcggc aacaacggca tgcgccgcga cttcctgggtg 540

cgcgactgca agaccggcga gatctgacc cgtgcacct ccctgagcgt gctgatgaac 600

accgcacccc gccgcctgag caccatcccc gacgaggtgc gcggcgagat cgcccccgcc 660

ttcatcgaca acgtggccgt gaaggacgac gagatcaaga agctgcagaa gctgaacgac 720

tccaccgccg actacatcca gggcggcctg accccccgct ggaacgacct ggacgtgaac 780

cagcacgtga acaacctgaa gtacgtggcc tgggtgttcg agaccgtgcc cgacagcatc 840

ttcagatccc accacatcag ctcttcacc ctggagtacc gccgcgagtg caccgcgac 900

tccgtgctgc gcagcctgac caccgtgagc ggcggcagct ccgaggccgg cctggtgtgc 960

gaccacctgc tgcagctgga gggcggcagc gaggtgctgc gcgcccgcac cgagtggcgc 1020

cccaagctga ccgactcctt ccgcggcatc agcgtgatcc ccgcgagcc ccgcgtgatg 1080

gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa ggacgacgac 1140

gacaagtga 1149

<210> 75

<211> 1146

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 75

atggcttcg cggcattcac catgtcggcg tgccccgcga tgactggcag ggcccctggg 60

gcacgteget ccggacggcc agtcgccacc cgctgaggg ggcgcgcccc cgactggtcc 120

atgctgttcg ccgtgatcac caccatcttc agcgcgcgagc agaagcagtg gaccaacctg 180

gagtggaagc ccaagcccaa gctgccccag ctgctggacg accacttcgg cctgcacggc 240

ctggtgttcc gccgcacctt cgccatccgc tectacgagg tgggccccga ccgcagcacc 300

tccatcctgg ccgtgatgaa ccacatgcag gagggcacc cgaaccacgc caagagcgtg 360

ggcatcctgg gcgacggctt cggcaccacc ctggagatgt ccaagcgga cctgatgtgg 420

gtggtgcgcc gcacccacgt ggccgtggag cgctacccca cctggggcga caccgtggag 480

gtggagtgtt ggatcggcgc cagcggcaac aacggcatgc gccgcgactt cctggtgcgc 540

gactgcaaga ccggcgagat cctgaccgc tgcacctccc tgagcgtgtt gatgaacacc 600

cgcacccgcc gcctgagcac catcccgac gaggtgcgcg gcgagatcgg ccccgcttc 660

atcgacaacg tggccgtgaa ggacgacgag atcaagaagc tgcagaagct gaacgactcc 720

accgccgact acatccaggg cggcctgacc ccccgctgga acgacctgga cgtgaaccag 780

cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga ccgtgcccga cagcatcttc 840

gagtcccacc acatcagctc ctccacctg gactaccgcc gcgagtgcac ccgcgactcc 900

gtgtgcgca gcctgaccac cgtgagcggc ggcagctccg aggccggcct ggtgtgcgac 960

cacctgtgtc agctggaggg cggcagcgag gtgtgcgcg cccgcaccga gtggcgcccc 1020

aagctgaccg actccttcg cggcaccgag gtgaccccc ccgagcccc cgtgatggac 1080

tacaaggacc acgacggcga ctacaaggac cagcactcg actacaagga cgacgacgac 1140

aagtga 1146

<210> 76

<211> 1155

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 76

atggccaccg catccacttl ctggcggttc aatgcccgtt gggcgacct gcgtcgctcg 60

gcgggctccg ggccccggcg ccacgcgagg cccctccccg tgcgcggggcg cgcccccgac 120

tggtccatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc 180

aacctggagt ggaagcccaa gcccagctg cccagctgc tggacgacca ctctggcctg 240

cacggcctgg tgttcgccg cacttcgcc atccgtcctt acgaggtggg ccccgaccgc 300

agcacctcca tcttgccgt gatgaaccac atgcaggagg ccacctgaa ccacgccaag 360

agcgtgggca tcttgggcga cggttcggc accaccctgg agatgtccaa gcgcgacctg 420

atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgct accccacctg gggcgacacc 480

gtggaggtgg agtctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttctg 540

gtgcgcgact gcaagaccgg cgagatcctg acccgtgca cctccctgag cgtgtgatg 600

aacacccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc 660

gccttcatcg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac 720

gactccaccg ccgactacat ccagggeggc ctgaccccc gctggaacga cctggacgtg 780

aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgacagc 840

atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcaccgcg 900

gactccgtgc tgcgcagcct gaccaccgtg agcggcggca gctccgaggc cggcctggtg 960

tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcccg caccgagtgg 1020

cgccccaagc tgaccgactc ctccgcggc atcagcgtga tccccccga gccccgcgtg 1080

atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140

gacgacaagt gatga 1155

<210> 77

<211> 1152

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 77

atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg 60

cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttctggtggcg cgcccccgac 120
 tgggccatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc 180
 aacctggagt ggaagcccaa gccaagctg cccagctgc tggacgacca ctfcggcctg 240
 cagggcctgg tgttcgccg cacttcgcc atccgtcct acgaggtggg ccccgaccgc 300
 agcacctcca tcttgccgt gatgaaccac atgcaggagg ccacctgaa ccacgccaag 360
 agcgtgggca tctggggcga cggttcggc accacctgg agatgtcaa gcgcgacctg 420
 atgtgggtgg tgcgcgcac ccacgtggc gtggagcgt accccacctg gggcgacacc 480
 gtggaggtgg agtgctggat cggcgccagc ggcaacaacg gcatgcgcg cgacttctg 540
 gtgcgcgact gcaagaccgg cgagatctg acccgtgca cctccctgag cgtgctgatg 600
 aacaccgca ccgcgcct gagcaccatc cccgacgagg tgcgcggcga gatcgcccc 660
 gccttcateg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac 720
 gactccaccg ccgactacat ccaggcggc ctgaccccc gctggaacga cctggacgtg 780
 aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgacagc 840
 atcttcgagt ccaccacat cagtccttc acctggagt accgccgca gtgcaccgc 900
 gactccgtgc tgcgcagcct gaccaccgtg agcggcgga gctccgagge cggcctggtg 960
 tgcgaccacc tctgcagct ggagggcggc agcgaggtgc tgcgcgccg caccgagtgg 1020

cgccecaage tgaccgactc ctcccgcggc atcagegtga tccccgccga gcccgcgtg 1080

atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140

gacgacaagt ga 1152

<210> 78

<211> 1155

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 78

atggccaccg catccacttt ctggcggttc aatgcccgt gggcgacct gcgtcgctcg 60

gcgggctccg ggccccggcg ccacgcgagg cccctccccg tgcgcgggcg ccccccgac 120

tggtccatgc tgttcgcgt gatcaccacc atcttctccg ccgcgcgagaa gcagtggacc 180

aacctggagt ggaagcccaa gcccaacccc cccagctgc tggacgacca ctctggcccc 240

cacggcctgg tgttcgcgcg caccttcgcc atccgcagct acgaggtggg ccccgaccgc 300

tccaccagca tcgtggccgt gatgaaccac ctgcaggagg ccgccctgaa ccacgccaaag 360

tccgtgggca tcctggggcg cggttcgggc accaccctgg agatgtcaa gcgcgacctg 420

atctgggtgg tgaagcgac ccacgtggcc gtggagcgt acccgcctg gggcgacacc 480

gtggagggtg agtgctgggt gggcgctcc ggcaacaacg gccgccgcca cgacttctg 540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgatgatg 600
aacacccgca cccgccgcct gagcaagatc cccgaggagg tgcgcggcga gatcggtccc 660
gccttcateg acaacgtggc cgtgaaggac gaggagatca agaagcccca gaagctgaac 720
gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacatc 780
aaccagcacg tgaacaacat caagtacgtg gactggatcc tggagaccgt gcccgacagc 840
atcttcgaga gccaccacat ctctctctc accatcgagt accgccgca gtgcacatg 900
gacagcgtgc tgcagtcctt gaccacegtg agcggcggtt cctccgagge cggcctggtg 960
tgcgagcacc tgcctcagct ggaggcgggc agcgaggtgc tgcgcgcca gaccgagtgg 1020
cgccccaagc tgaccgactc ctcccgggc atcagcgtga tccccgcca gtccagcgtg 1080
atggactaca aggaccacga cggcgactac aaggaccagc acatcgacta caaggacgac 1140
gacgacaagt gatga 1155

<210> 79

<211> 4817

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 79

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgccg cagccagagg gatttcagat gatcgcgct acaggttgag 300

cagcagtgc agcaaaggta gcagttgcc agaatacgc gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag categccgcc tggctgttc agctacaacg 480

gcaggagtea tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aatacctaac cctctggtg cttgctgggt 720

tgcattggcat cgcacatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgaagtc cctgacggc gctccctcgc ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcgggat gcgctgatcc caaacgagtt 1020

cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccga 1080

atggaatgcg ggaccgccag gctagcagca aaggcgctc ccctactccg catcgatgtt 1140

ccatagtga ttgactgca ttgggtggg gcggccggct gtctcttcg tgttcaaaa 1200

cgcgccagct cagcaacctg tccgtgggt ccccggtcc gatgaaatcg tgtcacgcc 1260

gatcagctga ttgccggct cgcgaagtag gcgcctcct ttctgcgcg cctctctcg 1320

tccgcctct agaatacaa tgcagagca ggaaggcctc cagccggct ccccgccgc 1380

ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggt gctccgacgc 1440

cgcctgttc cgctgtccg cccaggccg ccccgctg ttctgaaga ccgacctgc 1500

cggcgccctg aacgagctgc aggacgagc cgcccgctg tctggtgg ccaccaccg 1560

cgtgcctgc gcccgctgc tggacgtgtt gaccgagcc gcccgcgact ggctgctgt 1620

gggcgaggtg cccggccagg acctgtgtc ctccacctg gccccgccg agaaggtgc 1680

catcatggc gacgccatgc gccgctgca caccctggac cccgccacct gcccttga 1740

ccaccaggcc aagcaccga tcgagcgcg ccgcacccg atggaggccg gctggtgga 1800

ccaggacgac ctggacgagg agcaccagg cctggcccc gccgagctgt tcgccgct 1860

gaaggcccg atgcccgcg gcgaggacct ggtgggtgacc cacggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggtcttcac gactgcggcc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgcttc tgggtctgta cggcatcgcc gccccgact cccagcgcat 2100

cgccttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgtg cgcacacact tgctgccttg 2220

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280

cgcgcttttg cgagttgcta gctgctgtg ctatttgcga ataccacccc cagcateccc 2340

tccctcggt tcatatgct tgcateccaa ccgcaactta tctacgtgt cctgtatec 2400

ctcagegtg ctctgtctc tgetcactgc cctcgcaca gccttgggtt gggtccgcc 2460

tgtattctc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaa agagcgcgca gaggaacgt 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttgggt ctctgtccat tagcgaagcg tccggltcac acacgtcca cgttggcgag 2700

gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatac 2760

gaattccttt ctgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccagge ccccgattgc 2940

aaagacattt tagcgagcta ccaaagccat attcaaacac ctagatcact accatttcta 3000

cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacgg cgcgccatgg tggcgcgcgc cgcctccage gccttcttc 3120

ccgtgccccg ccccggcgc tcccccaagc ccggcaagtt cggcaactgg ccctccagcc 3180

tgagccccct cttaagccc aagtcctcc ccaacggcgg ctccagggtg aaggccaacg 3240

acagcgccca cccaaggcc aacggctccg ccgtgagcct gaagagcggc agcctgaaca 3300

cccaggagga cactctctcc agccccccc ccgcacctt cctgcaccag ctgcccgaact 3360

ggagccgcct gctgaccgcc atcaccaccg tgttcgtgaa gtccaagcgc cccgacatgc 3420

acgaccgcaa gtccaagcgc cccgacatgc tggtaggacag cttcggcctg gaggccaccg 3480

tgcaggacgg cctggtgttc cgccagtct tctccatccg ctctacgag atcggcaccg 3540

accgcaccgc cagcatcgag accctgatga accacctgca ggagacctcc ctgaaccact 3600

gcaagagcac cggcctcctg ctggacggct tcggccgcac cctggagatg tgcaagcgcg 3660

acctgatctg ggtgggtgat aagatgcaga tcaagggtgaa ccgtacccc gcctggggcg 3720

acaccgtgga gatcaacacc cgttcagcc gcctggggcaa gatcggcctg ggccgcgact 3780

ggctgatctc cgactgcaac accggcgaga tectgggtgcg cgccaccagc gectacgcca 3840
 tgatgaacca gaagacccgc cgctgtcca agctgcccta cgagggtcac caggagateg 3900
 tgccctgtt cgtggacagc cccgtgatcg aggactccga cctgaagggtg cacaagtta 3960
 aggtgaagac cggcgacagc atccagaagg gectgacccc cggctggaac gacctggacg 4020
 tgaaccagca cgtgtccaac gtgaagtaca tcggctggat cctggagagc atgccaccg 4080
 aggtgctgga gaccaggag ctgtgtccc tggccctgga gtaccgccgc gaggcgggcc 4140
 cgcactccgt gctggagagc gtgaccgcca tggaccccag caagggtggc gtgcgtccc 4200
 agtaccagca cctgtgcgc ctggaggacg gcaccgcat cgtgaacggc gccaccgagt 4260
 ggcgccccaa gaacgccggc gccaaogggc ccateccac cggcaagacc agcaacggca 4320
 actccgtgtc catggactac aaggaccagc acggcgacta caaggaccac gacatcgact 4380
 acaaggacga cgacgacaag tgactcgagg cagcagcagc tcggatagta tcgacacact 4440
 ctggacgtg gtcgtgtgat ggactgttc cgccacactt gctgcctga cctgtgaata 4500
 tccctgccgc tttatcaaa cagcctcagt gtgttgatc ttgtgtgtac gcgctttgc 4560
 gagggtctag ctgctgtgc tatltcgaa taccacccc agcaccct tccctgttt 4620
 catatcgctt gcacccaac cgcaacttat ctacgtgtc ctgtatccc tcagcgtgc 4680
 tectgtcct gctactgcc cctcgacag ccttggttg gctccgct gtattcct 4740

ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga 4800

acacaaatgg aaagctt 4817

<210> 80

<211> 4665

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 80

ggtaccegcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgctga ttgaggcttg gtcatcgagg tggaagctta tgtgtgtgct ggcttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgagg cagccagagg gatctcagat gatcgcgct acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgc agaatgatcg gtacagtgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttgc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatgggtg tgcggatgct gccggctgct gcctgctgag 600
 cacgccgttg agtgggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtgggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
 tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcaagccat 780
 gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct tgcagaacc 840
 atgctcatgg acgcatgtag cgtgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgtgatcc caaacgagtt 1020
 cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggtggt gcagcccga 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgctc ccctactccg catgatgtt 1140
 ccatagtga ttggactgca ttgggtggg gcggccggct gtttcttcg tgttgcaaaa 1200
 cgcgccagct cagcaacctg tccgtgggt ccccggtgc gatgaaatcg tgtgcacgc 1260
 gatcagctga ttgccggct cgcgaagtag gcgcctctt ttctgctgc cctctctcg 1320
 tccgcctct agaatacaa tgcagagca ggacggcctc cagccggct ccccgccgc 1380
 ctgggtggag cgcctgttc gctacgactg ggcccagcag accatgggt gctccgacgc 1440
 cgccgtgtc cgcctgtcc cccagggccg ccccggtctg ttctgaaga ccgacctgc 1500

cggcgccctg aacgagctgc aggacgaggc cggccgcctg tctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 gggcgagggtg cccggccagg acctgctgtc ctccacctg gccccgccg agaaggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac ccgccacct gcccttga 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgccgcct 1860
 gaaggccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgttctc cggttcacg gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atgcccgagg agctgggcgg 2040
 cgagtgggcc gaccgttcc tggctgtga cgccatgcc gccccgact ccagcgcac 2100
 cgccttctac cgctgctgg acgagttct ctgacaattg gcagcagcag ctggatagt 2160
 atcgacacac tctggacgt ggtcgtgtga tggactgtg ccgccacact tctgccttg 2220
 acctgtgaat atccctgccg cttttatca acagcctcag tgtgtttgat cttgtgtga 2280
 cgcgctttg cgagttgcta gctgctgtg ctatttgcga ataccacccc cagcaccccc 2340
 tcccctggt tcatatgct tgcacccaa ccgcaacta tctacgtgt cctgctatcc 2400
 ctacgcgtg ctctgtcc tgcactgc ccctgcaca gccttggtt gggtccgcc 2460

tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt ctctgctcat tagcgaagcg tccgggtcac acacgtgcca cgttggcgag 2700

gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata 2760

gaattccttt ctgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tccgggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcaact accacttcta 3000

cacagggcac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatgacg ttccgggctcg cctcccggc catgggcgcg 3120

ggtgtctccc ttccccggcc cagggtcgcg gtgcgcgccc agtcggcgag tcaggttttg 3180

gagagcgggc gcgcccagct gcccgactgg agccgcctgc tgaccgccat caccaccgtg 3240

ttcgtgaagt ccaagcgcgc cgacatgcac gaccgcaagt ccaagcgcgc cgacatgctg 3300

gtggacagct tcggcctgga gtccaccgtg caggacggcc tgggttccg ccagtccttc 3360

tccatccgct cctacgagat cggcaccgac cgcaccgcca gcatcgagac cctgatgaac 3420

cacctgcagg agacctccct gaaccactgc aagagcaccg gcatacctgct ggacggettc 3480
 ggccgcaccc tggagatgtg caagcgcgac ctgatctggg tggatgataa gatgcagatc 3540
 aaggtgaacc gctaccccg cctggggcgac accgtggaga tcaacacccg cttcagccgc 3600
 ctgggcaaga tcggcatggg ccgcgactgg ctgatctccg actgcaaac cggcgagatc 3660
 ctggtgcgcg ccaccagcgc ctacgccatg atgaaccaga agaccgcgcg cctgtccaag 3720
 ctgcctacg aggtgcacca ggagatcgtg cccctgttcg tggacagccc cgtgatcgag 3780
 gactccgacc tgaaggtgca caagttcaag gtgaagaccg ggcacagcat ccagaagggc 3840
 ctgacccccg gctggaacga cctggacgtg aaccagcagc tgccaacgt gaagtacatc 3900
 ggctggatcc tggagagcat gccaccgag gtgctggaga cccaggagct gtgctccctg 3960
 gccctggagt accgccgcga gtgcggccgc gactccgtgc tggagagcgt gaccgccatg 4020
 gacccagca aggtgggctg gcgtcccag taccagcacc tgctgcgcct ggaggacggc 4080
 accgccatcg tgaacggcgc caccgagtgg cgcaccaaga acgccggcgc caacggcgcc 4140
 atctccaccg gcaagaccag caacggcaac tccgtgtcca tggactaaa ggaccacgac 4200
 ggcgactaca aggaccagca catcgactac aaggacgacg acgacaagt actcgaggca 4260
 gcagcagctc ggatagtatc gacacaactt ggacgttgt cgtgtgatgg actgttgccg 4320
 ccacacttgc tgccttgacc tgtgaatatc cctgcgcgtt ttatcaaca gcctcagtgt 4380

gtttgatctt gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta ttgcgaata 4440

ccacccccag catcccttc cctcgtttca tategttgc atcccaaccg caacttatct 4500

acgtgtcct gctatccctc agcgtgtctc ctgctcctgc tcaactcccc tcgcacagcc 4560

ttggtttggg ctccgcctgt attctctgg tactgcaacc tgtaaaccag cactgcaatg 4620

ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agctt 4665

<210> 81

<211> 4668

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 81

ggtacccgcc tgcaacgcaa gggcagccac agccgtccc acccgccgt gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgtga ttgaggcttg gtcatcggg tggaagctta tgtgtgtgt gggcttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtagt tcaatcgcgg cagccagagg gattcagat gatcgcgcgt acaggtagt 300

cagcagtgtc agcaaaggta gcagtttgc agaatgatcg gtacagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag categccgcc tggctgttgc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
 tgcaagcatg cctggcgaag acacatgggtg tgcggatgct gccggctgct gcctgctgcg 600
 cacgccgttg agtgggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
 tgcattggcat cgcaccatca gcaggagcgc atggaaggg actggcccca tgcacgcat 780
 gccaaaccgg agcgcaccga gtgtccacac tgcaccagg cccgcaagct ttgcagaacc 840
 atgctcatgg acgcatgtg cgtgacgtc cctgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgcccagc cagcgggtgat gcgctgatcc caaacgagtt 1020
 cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggtggt gcagcccgca 1080
 atggaatgag ggaccgccag gctagcagca aaggcgctc ccctactccg catgatgtt 1140
 ccatagtga ttggactga ttgggtggg gcggccgggt gttctttcg tgtgcaaaa 1200
 cgcgccagct cagcaacctg tccgtgggt ccccggtgcc gatgaaatcg tgtcacgcc 1260
 gatcagctga ttgccgggt cgcgaagtag gcgccctct tctgctcgc cctctctcg 1320

tcccgcctct agaataatcaa tgategagca ggaeggccctc cacgccggct cccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggt gctccgacgc 1440
 cgccgtgttc cgctgtccg cccagggccg ccccgctgtg ttcgtgaaga ccgacctgc 1500
 cggcgccctg aacgagctgc aggacgaggc cgcccgctg tctggctgg ccaccaccg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggcccgact ggctgtgtct 1620
 gggcgagggtg cccggccagg acctgtgtc ctccacctg gccccgccg agaaggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740
 ccaccaggec aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccagg cctggcccc gccgagctgt tcgccgcct 1860
 gaaggcccc atgccgacg ggcaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgcttctc cggttcac gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atgccgagg agctgggcgg 2040
 cgagtgggcc gaccgttcc tgggtgtgta cggcatgcc gccccgact cccagcgc 2100
 cgcttctac cgctgtctgg acgatttct ctgacaattg gcagcagcag ctggatagt 2160
 atcgacacac tctggacgt ggtcgtgtga tggactgtg ccgccacact tctgccttg 2220
 acctgtgaat atccctgccg ctftatcaa acagcctcag tctgtttgat ctgtgtgta 2280

cgcgcctttg cgagttgcta gctgcttggt ctatttgcga ataccacccc cagcatcccc 2340
 ttccctcggt tcatacgct tgcataccaa ccgcaactta tctacgctgt cctgctatcc 2400
 ctcagecgtg ctccgtctcc tgcacactgc cctcgcaca gccttgggtt gggtcccgcc 2460
 tgtattctcc tggtaactga accgttaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctgaac agagcgcgca gaggaacgt 2580
 gaaggctcgc cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttgggt ctccgtccat tagcgaagcg tccgggtcac acacgtgcc cgttggcgag 2700
 gtggcaggtg acaatgatcg gtggagctga tggcgaac gttcacagcc tagggatata 2760
 gaattccttt ctgcgctat gacactcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgt gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tgggggctgc 2880
 atgggcgctc cgatgccgt ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat atcaaacac ctagatcact accacttcta 3000
 cacaggccac tcagactgt gatcgactc cgctaagggg gcgcctcttc ctctcggtt 3060
 cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120
 cctccgttca cgatcgggac gctgcgcaag gccatcccc cgcactgttt cgagcgctcg 3180
 gcgcttcgtg ggcgcgcca gctgccgac tggagccgcc tgetgaccgc catcaccacc 3240

gtgttcgtga agtccaagcg ccccgacatg caccgaccgca agtccaagcg ccccgacatg 3300
 ctggtggaca gcttcggcct ggagtcacc gtgcaggacg gcctgggtgtt ccgccagtcc 3360
 ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gacctgatg 3420
 aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480
 ttgggcgcga ccctggagat gtgcaagcgc gacctgatct ggggtggat caagatgcag 3540
 atcaaggtga accgtaccc cgcctggggc gacaccgtgg agatcaaac ccgcttcagc 3600
 cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgca caccggcgag 3660
 atcctgggtgc gcgccaccag cgcctacgcc atgatgaacc agaagaccgc ccgctgtcc 3720
 aagctgccct acgaggtgca ccaggagatc gtccccctgt tcgtggacag cccgtgatc 3780
 gaggactccg acctgaaggt gcacaagtc aaggtgaaga ccggcgacag catccagaag 3840
 ggctgaccc ccggtggaa cgacctggac gtgaaccagc acgtgtcaa cgtgaagtac 3900
 atcggctgga tcttgagag catgccacc gaggtgctgg agaccagga gctgtgctcc 3960
 ctggccctgg agtaccgcgc cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020
 atggacceca gcaaggtggg cgtgcgctcc cagtaccagc acctgtgcg cctggaggac 4080
 ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc 4140
 gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200

gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgactcgag 4260

gcagcagcag ctoggatagt atcgacacac tctggacgct ggctgtgtga tggactgttg 4320

ccgccacact tgetgccttg acctgtgaat atccctgccg cttttatcaa acagccicag 4380

tgtgtttgat ctgtgtgtga cgcgcttttg cgagttgcta gctgcttgig ctatttgcga 4440

ataccacccc cagcatcccc ttccctcggt tcatactgct tgcacccaa ccgcaactta 4500

tctacgtgt cctgctatcc ctacagcgtg ctctgtctcc tgcactgc ccctgcaca 4560

gccttgggtt gggctccgcc tgtattctcc tggctactgca acctgtaaac cagcactgca 4620

atgtgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 4668

<210> 82

<211> 4668

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 82

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgtga ttgaggcttg gtcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

gccgggcaat ggcgatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
 tccaggtggt tcaatcgcg gaggcagagg gatttcagat gatcgcegt acaggttgag 300
 cagcagtgtc agcaaaggta gcagtttccc agaatacgc gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga 540
 tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgtcgcg 600
 cagcccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg ctgtctgggt 720
 tgcattggcat cgcacatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat 780
 gccaaaccgg agcgaccga gtgtccacac tgcaccagg cccgcaagct ttgcagaacc 840
 atgctcatgg acgcatgtag cgtgacgtc cctgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgcccagc cagcgggtgat gcgtgatcc caaacgagtt 1020
 cacattcatt tgcattcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgctc ccctactccg catcgatgtt 1140

ccatagtga ttggactga ttgggtggg gggcgggct gttttttcg tgttgcaaaa 1200
 cggccagct cagcaacctg tccgtgggt ccccggtcc gatgaaatcg tgtgcacgc 1260
 gatcagctga ttgccggct cgggaagtag gggccctct ttctctcgc cctctctcg 1320
 tccgcctct agaatacaa tgcagagca ggacggctc cagccggct ccccgccgc 1380
 ctgggtggag cgcctgtcg gctacgactg ggccagcag accatggct gctccgacg 1440
 cgccgtgtc cgcctgtcg cccaggcgc cccgtgtg ttcgtgaaga ccgacctgc 1500
 cggcgccctg aacgagctg aggacgaggc cggcgccctg tctggttg ccaccaccg 1560
 cgtgccctg gccgcctg tggacgtgt gaccgaggc ggccgcgact ggctgtgt 1620
 gggcgaggtg cccggccagg acctgtgt ctcacctg gcccgcgc agaaggtgc 1680
 catcatggc gacgccatg gccgcctga cccctggac cccgccact gcccttga 1740
 ccaccaggc aagcaccga tcagcgcgc ccgacccgc atggaggcc gcctggtga 1800
 ccaggacgac ctggacgagg agcaccagg cctggcccc gccgagctgt tcgccgcct 1860
 gaaggccgc atgccgacg gcgaggacct ggtggtgacc cagggcag cctgcctgc 1920
 caacatcatg gtggagaacg gccgttctc cggttctc gactgcggc gcctgggcgt 1980
 ggccgaccg taccaggaca tcgcctggc ccccgcgac atcgccagg agctgggcg 2040
 cgagtgggc gaccgttcc tgggtgtga cggcatgcc gcccgcact ccagcgcat 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggaatgt 2160
 atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
 acctgtgaat atccctgccg cttttataca acagccctcag tgtgtttgat cttgtgtgta 2280
 cgcgcttttg cgagttgcta gctgcctgtg ctatttgcga ataccacccc cagcatcccc 2340
 ttcctcgtt tcatatgct tgcateccaa ccgcaactta tctacgctgt cctgctatcc 2400
 ctcagecgtg ctctgtctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
 gaaggctcgc cctctgtcgc acctcagegc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
 gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760
 gaattccttt ctgcgctat gacactcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgct gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tcggggctgc 2880
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccagge ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctatgactact accacttcta 3000
 cacaggccac tcgagcttgt gatgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggcc accgcatcca cttctcggc gticaatgcc 3120
 cgctggggcg acctgcgleg ctggcggggc tccgggcccc ggcgcccage gagggccctc 3180
 cccgtgcgcg ggcgcgcccga gtgcgccgac tggagccgcc tgctgaccgc catcaccacc 3240
 gtgttcgtga agtccaagcg ccccgacatg caccgaccga agtccaagcg ccccgacatg 3300
 ctggtggaaca gcttcggcct ggagtcacc gtgcaggacg gcctgggtgt cgcacagtc 3360
 ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gacctgatg 3420
 aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480
 ttggccgca cctggagat gtgcaagcg gacctgatc gggfgggtgat caagatgcag 3540
 atcaaggtga accgtaccc cgcctggggc gacaccgtgg agatcaaac ccgcttcage 3600
 cgcctgggca agatcggcat gggccgcgac tggtgatct ccgactgcaa caccggcgag 3660
 atcctgggtgc ggcaccaccg cgcctacgcc atgatgaacc agaagaccgc ccgcctgtcc 3720
 aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgatc 3780
 gaggactccg acctgaaggt gcacaagtc aaggtgaaga ccggcgacag catccagaag 3840
 ggcctgaccc ccggtggaa cgacctggac gtgaaccagc acgtgtcaa cgtgaagtac 3900
 atcggctgga tcctggagag catgccacc gaggtgctgg agaccagga gctgtgctcc 3960
 ctggccctgg agtaccgcgc cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020

atggacccca gcaaggtggg cgtgcgtcc cagtaccage acctgctgcg cctggaggac 4080

ggcaccgcca tcgtgaacgg cgccaccgag tggegcceca agaacgccgg cgccaacggc 4140

gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200

gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgaaa gtgactcgag 4260

gcagcagcag ctoggatagt atcgacacac tctggacgtt ggtcgtgtga tggactgttg 4320

ccgccacact tgetgccttg acctgtgaat atccctgccg ctltatcaa acagcctcag 4380

tgtgtttgat ctgtgtgta cgcgcttttg cgagttgcta gctgcttggt ctatttcga 4440

ataccacccc cagcaccccc ttccctcgtt tcatacgtt tgcaccceaa ccgcaactta 4500

tctacgtgtt cctgctatcc ctacgcgtg ctctgtctcc tgctcactgc ccctcgcaca 4560

gccttggttt gggtccgcc tgtattctcc tggactgca acctgtaaac cagcactgca 4620

atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 4668

<210> 83

<211> 4656

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 83

ggtacccgcc tgcaacgcaa gggcagccac agcegetccc acccgceget gaaccgacac 60
 gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtget 120
 gccatgetga ttgaggcttg gttcatcggg tggaagctta tgttgtgtct gggcttgcac 180
 gccgggcaat gcgcattgtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
 tccagggtgt tcaatcgagg cagccagagg gatttcagat gatcgcgct acaggttgag 300
 cagcagtgtc agcaaaggta gcagttgcc agaatgatcg gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga 540
 tgcaageatg cctggcgaag acacatgggt tgcggatget gccggctget gcctgtgcg 600
 cagccgttg agttggcagc aggcacagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtgtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgtgggt 720
 tgcattggcat cgcacatca gcaggagcg atcggaaggg actggcccca tgcacgcat 780
 gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg ccgcaagct ttgcagaacc 840
 atgtcatgg acgcatgtag cgctgacgtc cctgacggc gctcctctcg ggtgtggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcgggggc gcgccccagc cagcgggat gcgctgatcc caaacgagtt 1020
 cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
 atggaatgcg ggaccgccag gctagcagca aaggegcctc ccctacccg catcgatgtt 1140
 ccatagtga ttggactgca ttgggtggg gcggccggct gttctttcg tgtgcaaaa 1200
 cgcgccagct cagcaacctg tcccgtaggt ccccggtgcc gatgaaatcg tgtgcacgcc 1260
 gatcagetga ttgccggct cgcgaagtag gcgcctcct ttctgtcgc cctctctccg 1320
 tcccgctct agaatacaa tgatcgagca ggacggcctc cagccggct ccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggt gctccgacgc 1440
 cgccgtgttc cgctgtccg cccagggccg ccccggtctg ttctgaaga ccgacctgc 1500
 cggcgccctg aacgagctgc aggacgaggc cggccgctg tctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgtgtct 1620
 gggcgaggtg cccggccagg acctgtgtc ctcccacctg gccccgccg agaaggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca cacctggac cccgccacct gcccttga 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgccgcct 1860
 gaagggccgc atccccgacg gcgaggacct ggtggtgacc cagggcgacg cctgcctgcc 1920

caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccciggc caccgcgac atcgccgagg agctgggcgg 2040
 cgagtgggcc gaccgcttc tgggtctgta cggcatcgcc gccccgact cccagegcat 2100
 cgccttctac cgctgtctgg acgagttctt ctgacaattg gcagcagcag ctggatagt 2160
 atcgacacac tctggacgct ggctgtgtga tggactgtg ccgccacact tctgacctg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
 cgcgctttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
 ttccctcgtt tcatacgtt tgcateccaa ccgcaactta tctacgtgt cctgctatcc 2400
 ctcagegtg ctctgtctcc tgcctactgc cctcgcaca gccttggttt gggtccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgtgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgt 2580
 gaaggctcgc cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttgggt ctctgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
 gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata 2760
 gaatctctt ctgcgctat gacactcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgt gcattgcaaca ccgatgatgc ttgaccccc cgaagctcct tcggggctgc 2880

atggggcgtc cgatgccgt ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcaact accacttcta 3000
 cacaggccac tcgagcttgt gatcgcactc cgctaagggg ggcctcttc ctcttcgttt 3060
 cagtcacaac ccgcaaacac tagtalggct tccgcggcat tcacatgtc ggcgtgcccc 3120
 gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc caccgcctg 3180
 agggggcgcg ccccgactg gagecgctg ctgaccgcca tcaccaccgt gtctgtgaag 3240
 tccaagcgcc ccgacatgca cgaccgcaag tccaagcgcc ccgacatgct ggtggacagc 3300
 ttggcctgg agtcaccgt gcaggacggc ctggtgttc gccagtcctt ctccatccgc 3360
 tcctacgaga tcggcaccga ccgcaccgcc agcatcgaga cctgatgaa ccacctgcag 3420
 gagacctccc tgaacctg caagagcacc ggcatcctgc tggacggctt cggccgcacc 3480
 ctggagatgt gcaagcgca cctgatctgg ggggatga agatgcagat caaggtgaac 3540
 cgctaccccg cctggggcga caccgtggag atcaacaccc gcttcagccg cctgggcaag 3600
 atcggcatgg gccgcgactg gctgatctcc gactgcaaca ccggcgagat cctggtgcgc 3660
 gccaccagcg cctacgccat gatgaaccag aagaccgcc gcctgtcaa gtgccttac 3720
 gaggtgcacc aggagatcgt gcccctgttc gtggacagcc ccgtgatga ggactccgac 3780
 ctgaagggtgc acaagttcaa ggtgaagacc ggcgacagca tccagaaggg cctgaccccc 3840

ggctggaacg acctggacgt gaaccagcac gtgtccaacg tgaagtacat cggttggatc 3900

ctggagagca tgcccaccga ggtgctggag acccaggagc tgtgtccct ggccctggag 3960

taccgccgag agtgccggcg cgactccgtg ctggagagcg tgaccgcat ggacccagc 4020

aagggtggcg tgcgtccca glaccagcac ctgctgcgc tggaggacgg caccgccatc 4080

gtgaacggcg ccaccgagtg gcgcccgaag aacgccggcg ccaacggcgc catctccacc 4140

ggcaagacca gcaacggcaa ctccgtgtcc atggactaca aggaccacga cggcgactac 4200

aaggaccacg acatcgacta caaggacgac gacgacaagt gactcgagcg agcagcagct 4260

cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgc gccacactg 4320

ctgccttgac ctgtgaatat cctgccgct ttatcaaac agcctcagtg tgttgatct 4380

tgtgtgacg cgttttgcg agttgctagc tcttgtgct atttgcgaat accacccca 4440

gcacccctt cctcgttgc atatcgctg catccaacc gcaacttgc tacgtgtcc 4500

tgctatccct cagcgtgct cctgctcctg ctactgcc ctgcacagc ctgtgttg 4560

gctccgctg tattctctg gtactgcaac ctgtaaacca gcaactgcaat gctgatgac 4620

gggaagtagt gggatgggaa cacaatgga aagctt 4656

<210> 84

<211> 4721

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 84

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgttgtgtgt gggcttgcac 180

gccgggcaat gcgcatgggt gcaagagggc ggcagcactt gctggagctg ccgcgggtgcc 240

tccaggtggt tcaatcgagg cagccagagg gatctcagat gatcgcgct acaggttgag 300

cagcagtgtc agcaaaggta gcagtttccc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccgggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcggggcgtga 540

tgcaagcatg cctggcgaag acacatgggt tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcatggcat cgcacccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt 1020
 cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgcctc cctactccg catgatgtt 1140
 ccatagtga ttgactgca ttgggtggg gcggccggct gtttcttcg tttgcaaaa 1200
 cgcgccagct cagcaacctg tccgtgggt ccccggtgcc gatgaaatcg tgtcacgcc 1260
 gatcagctga ttgccggct cgcgaagtag gcgcctcct ttctgctgc cctctctcg 1320
 tccgcctct agaatatcaa tgatcgagca ggacggctc cagccggct ccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggt gtcggacgc 1440
 cgccgtgttc cgctgtccg cccagggccg ccccggtctg ttctgaaga ccgacctgc 1500
 cggegccctg aacgagctgc aggacgagc cgccgcctg tctggttg ccaccaccg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 gggcgaggtg cccggccagg acctgtgtc ctcccacctg gccccgcg agaaggtgc 1680
 catcatggc gacgccatgc gccgctgca caccctggac cccgccacct gcccttga 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
 gaaggccgc atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgttctc cggcttcac gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgcac atgcgcgagg agctgggcgg 2040
 cgagtgggcc gaccgttcc tgggtctgta cggcatcgcc gccccgact cccagcgcct 2100
 cgccttctac cgctgtctgg acgagttctt ctgacaattg gcagcagcag ctcgatagt 2160
 atcgacacac tctggacgt ggtcgtgtga tggactgtg ccgccacact tctgccttg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgla 2280
 cgcgttttg cgagtgcta gctgcttggt ctatttgcga ataccacccc cagcaccccc 2340
 ttccctcgtt tcatacgtc tgcacccaa ccgcaactta tctacgtgt cctgctatcc 2400
 ctccagctg ctctgtctc tgcactgc cctcgcaca gccttggtt gggtccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgt 2580
 gaaggctcgc cctctgtgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgttggtt ctctgcat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
 gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gtacacagcc tagggatatc 2760

gaattccttt ctgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccagge ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctatgactact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacgg cgcgccatgg ccaccaccag cctggcctcc gccttctget 3120

ccatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc agctccgacc 3180

tgcagctgcg cgccggcaac gccccacct cctgaagat gatcaacggc accaagtta 3240

gctacaccga gagcctgaag cgcctgcccg actggtccat gctgttcgcc gtgatcacca 3300

ccatcttcag cgccgccgag aagcagtgga ccaacctgga gtggaagccc aagcccaagc 3360

tgccccagct gctggacgac cacttcggcc tgcacggcct ggtgttcgcc cgcaccttcg 3420

ccatccgctc ctacgaggtg ggccecgacc gcagcacctc catcctggcc gtgatgaacc 3480

acatgcagga ggccaccctg aaccacgcca agagcgtggg catcctgggc gacggcttcg 3540

gcaccaccct ggagatgtcc aagcgcgacc tgatgtgggt ggtgcgcgc acccacgtgg 3600

ccgtggagcg ctaccccacc tggggcgaca ccgtggaggt ggagtgtctgg atcggcgcca 3660

gcggcaacaa cggcatgcgc cgcgacttcc tgggtgcgca ctgcaagacc ggcgagatcc 3720

tgacccgctg cacctccctg agcgtgctga tgaacacccg caccgcccgc ctgagcacca 3780

tccccgacga ggtgcgcggc gagatcgccc ccgcttcat cgacaacgtg gccgtgaagg 3840

acgacgagat caagaagctg cagaagctga acgactccac cggcgactac atccagggcg 3900

gcctgacccc ccgctggaac gacctggacg tgaaccagca cgtgaacaac ctgaagtaag 3960

tggcctgggt gttcgagacc gtgcccgaca gcatttcca gtcccaccac atcagctcct 4020

tcacctgga gtaccgccc gagtgcaccc gcgactccgt gctgcgcagc ctgaccaccg 4080

tgagcggcgg cagctccgag gccggcctgg tgtgcgacca cctgctgcag ctggagggcg 4140

gcagcgaggt gctgcgcgcc cgcaccgagt ggcgccccaa gctgaccgac tecttcgcg 4200

gcatcagcgt gatccccgcc gageccccgc tgatggacta caaggaccac gacggcgact 4260

acaaggacca cgacatcgac tacaaggacg acgacgaaa gtgatgactc gaggcagcag 4320

cagctcggat agtatcgaca cactctggac gctggctcgt tgatggactg ttccgccac 4380

acttctgcc ttgacctgtg aatatccctg ccgctttat caaacagcct cagtgtgtt 4440

gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtctatttg cgaataccac 4500

ccccageatc ccttccctc gttcattatc gcttgcatcc caaccgcaac ttatctacgc 4560

tgctctgcta tccctcagcg ctgctctgc tectgtcac tgcctctgc acagccttgg 4620

ttgggctcc gcctgtattc tectgttact gcaacctgta aaccagcact gcaatgctga 4680

tgcaaggga gtagtgggat gggaacacaa atggaaagct t 4721

- <210> 85
- <211> 4650
- <212> ADN
- <213> Secuencia artificial

<220>
<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 85
ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgt gggcttgcac 180

gccgggcaat gcgcatggcg gcaagagggc ggcagcactt gctggagctg ccgcgggtgcc 240

tccaggtggt tcaatcgagg cagccagagg gatttcagat gatcgcgctg acaggttgag 300

cagcagtgtc agcaaaggta gcagtttccc agaagatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggtgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcggggcgtga 540

tgcaagcatg cctggcgaag acacatggcg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agtgggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctclggctg cttgctgggt 720

tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat 780

gccaaaccgg agcgcaccga gtgtccacac tgtaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgtgacgtc ccttgacggc gclctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgcccage cagcggatg gcgtgatcc caaacgagtt 1020

cacattcatt tgcattcctg gagaagcag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc cctactccg catgatgtt 1140

ccatagtgca ttgactgca ttgggtggg gcggccggct gttctttcg tttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtaggt ccccggtcc gatgaaatcg tgtgcacgc 1260

gatcagctga ttgccggct cgcgaagtag gcgcctctt ttctgtcgc cctctctcg 1320

tccgcctct agaatacaa tgcagagca ggaaggcctc cagccggct ccccgccgc 1380

ctgggtggag cgcctgttc gctacgactg ggcccagcag accatcggt gctccgacg 1440

cgcctgttc cgcctgtccg cccaggccg ccccggtctg ttctgaaga ccgacctgc 1500

cggcgcctg aacgagctgc aggaagagge cgccgcctg tctggctgg ccaccaccg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 gggcgagggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaagggtgc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttga 1740
 ccaccaggcc aagcaccgca tcgagcgcg cgcacccgc atggaggccg gcctgggtga 1800
 ccaggacgac ctggacgagg agcaccagg cctggcccc gccgagctgt tcgccgcct 1860
 gaaggccgc atgccgacg gcgaggacct ggtgggtgacc cagggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgttctc cggcttcac gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atgccgagg agctgggcgg 2040
 cgagtgggcc gaccgttcc tggltctga cggcatgcc gccccgact cccagcgc 2100
 cgccttctac cgcctgctgg acgagttct ctgacaattg gcagcagcag ctggatagt 2160
 atcgacacac tctggacgt ggtcgtgtga tggactgtg ccgccacct tctgccttg 2220
 acctgtgaat atccctgcc cttttataa acagcctcag tgtgtttgat cttgtgtga 2280
 cgcgcctttg cgagttgcta gctgctgtg clatttgcga ataccaccc cagcateccc 2340
 ttcctcgtt tcatatgct tgcateccaa ccgcaactta tctacgtgt cctgctatcc 2400
 ctcagegtg ctctgtcc tgcactgc cctcgaca gccctgggtt gggtccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgaacgaat 2640

gcgcttggtt ctctgtccat tagcgaagcg tccgggtcac aacgtgccca cgttggcgag 2700

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggataac 2760

gaattccttt ctgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggtc 2820

tcccggcgct gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatcccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctatgactact accacttcta 3000

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctctctgttt 3060

cagtcacaac ccgcaaacac tagtatgacg ttccgggtcg cctcccggc catgggcgcg 3120

ggtgtctccc ttcccggcc cagggtcgcg gtgcgcgcgc agtcggcgag tcaggttttg 3180

gagagcgggc gcgcctccga ctggtccatg ctgttcgcgc tgatcaccac catcttcagc 3240

gccgccgaga agcagtggac caacctggag tggaagccca agcccaagct gccccagctg 3300

ctggacgacc acttcggcct gcacggcctg gtgttcgcgc gcaacttcgc catcgctcc 3360

tacgaggtgg gccccgaccg cagcacctcc atcctggcgc tgatgaacca catgcaggag 3420

gccacctga accacgcca gagegtgggc atcctgggcg acggttcgg caccacctg 3480

gagatgtcca agcgcgacct gatgtgggtg gtgcgccgca cccacgtggc cgtggagcgc 3540
 taccacacct ggggcgacac cgtggaggtg gagtgctgga tcggcgccag cggcaacaac 3600
 ggcatgcgcc gcgacttctt ggtgcgcgac tgcaagaccg gcgagatcct gaccgcctgc 3660
 acctccctga gcgtgctgat gaacacccgc accgcgcgcc tgagcaccat ccccgacgag 3720
 gtgcgcggcg agatcgcccc cgccttcac gacaacgtgg ccgtgaagga cgacgagatc 3780
 aagaagctgc agaagctgaa cgactccacc gccgactaca tcaggggcgg cctgaccccc 3840
 cgctggaacg acctggacgt gaaccagcac gtgaacaacc tgaagtacgt ggcttgggtg 3900
 ttcgagaccg tgccccgacag catcttcgag tcccaccaca tcagctcctt caccctggag 3960
 taccgcgcgc agtgcacccg cgactccgtg ctgcgcagcc tgaccacgtg ggcggcggc 4020
 agctccgagg ccggcctggt gtgcgaccac ctgtgcagc tggagggcgg cagcgaggtg 4080
 ctgcgcgccc gcaccgagtg gcgccccaa ctgaccgact ccttcgcgg catcagcgtg 4140
 atccccgcgc agccccgcgt gatggactac aaggaccagc acggcgacta caaggaccac 4200
 gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata 4260
 gtatcgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgtgcct 4320
 tgacctgtga atatccctgc cgttttacc aaacagcctc agtgtgttg atctgtgtg 4380
 tacgcgttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc 4440

ccttccctcg ttcatatcg ctgcatccc aaccgcaact tatctacgt gtcctgctat 4500

ccctcagegc tgctctgct cctgctcact gccctcgca cagccttggc ttgggctccg 4560

cctgtattct cctggctactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620

tagtgggatg ggaacacaaa tggaaagctt 4650

<210> 86

<211> 4653

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 86

ggtaccgcc tgcaacgcaa gggcagecac agcgcctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgctgcgc atgcttgtgc tggtaggct gggcagtgt 120

gcatgctga ttgaggcttg gttcatcggg tggaaagcta tgtgtgtgct gggttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgagg cagccagagg gatttcagat gatcgcgct acaggttgag 300

cagcagtgc agcaaaggta gcagtttgc agaatgatc gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttcc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggagtg 540
 tgcaagcatg cctggcgaag acacatgggtg tgcggatgct gccggctgct gcctgctgcg 600
 cagcccggtg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtgggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
 tgcattggcat cgcaccatca gcaggagcgc atggaaggg actggcccca tgcaagccat 780
 gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg ccgcaagct ttgcagaacc 840
 atgctcatgg acgcatgtag cgtgacgtc cctgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgtgatcc caaacgagtt 1020
 cacatfcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccga 1080
 atggaatgcg ggaccgccag gctagcagca aaggegcctc ccctactccg catgatgtt 1140
 ccatagtgea ttggactgca ttgggtggg gcggccggct gttctttcg tgttgcaaaa 1200
 cgcgccagct cagcaacctg tccgtgggt ccccggtgcc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgccggct cgcgaagtag gcgccctctt ttctgctgc cctctctccg 1320
 tccgcctct agaatatcaa tgategagca ggaaggcctc cagccgggt ccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggt gctccgacgc 1440
 cgccgtgttc cgcctgtccg cccagggccg ccccggtctg ttctgaaga ccgacctgc 1500
 cggcgcctg aacgagctgc aggacgagc cggccgctg tcttggtgg ccaccaccg 1560
 cgtgcctgc gccgacctg tggacgtgtt gaccgagcc gcccggaact ggtgtgtct 1620
 gggcgaggtg cccggccagg acctgtgtc ctccacctg gccccgccg agaaggtgtc 1680
 catcatgcc gacgcatgc gccgctgca caccctggac cccgccacct gcccttga 1740
 ccaccaggcc aagcaccgca tcgagcgcg ccgcaccgc atggaggccg gctggtgga 1800
 ccaggacgac ctggacgagg agcaccagg cctggcccc gccgagctgt tcgccgct 1860
 gaaggccgc atgccgacg gcgaggacct ggtggtgacc caggcgacg cctgcctgcc 1920
 caacatcatg gtggagaac gccgttctc cggttctc gactgggcc gccgtggcgt 1980
 ggccgaccg taccaggaca tcgccctgg caccgagc atgcgagg agctgggcg 2040
 cgagtgggcc gaccgttcc tgggtgtgta cggcatgcc gccccgact ccagcgcat 2100
 cgcttctac cgcctgttg acgagttct ctgacaattg gcagcagcag ctcgatagt 2160
 atcgacacac tctggacgt ggtcgtgtg tggactgtg ccgccacct tctgccttg 2220
 acctgtgaat atccctgcc cttttatcaa acagctcag tgtgttgat cttgtgtgta 2280
 cgcgttttg cgagttgcta gctgtgtg ctatttgcga ataccaccc cagcatccc 2340

ttccctcggt tcatacgct tgcateccaa ccgcaactta tctacgctgt cctgetatec 2400

ctcagcgctg ctectgctec tgctcactgc cctcgcaca gccttgggtt gggetccgcc 2460

tgtattctcc tggtagtgc accgttaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggctcgc cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttgggt ctctgctcat tagcgaagcg tccgggtcac acacgtgcca cgttggcgag 2700

gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata 2760

gaattccttt ctgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggecac tegagcttgt gatcgcactc cgctaagggg gcgcctcttc ctctcgttt 3060

cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120

cctccgttca cgategggac gctgcgcaag gccatcccg cgcactgttt cgagcgctcg 3180

gcgcttcgtg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 3240

agcgccgccc agaagcagtg gaccaacctg gagtgggaag ccaagcccaa gctgccccag 3300

ctgctggacg accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc 3360

tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag 3420

gaggccaccc tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc 3480

ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag 3540

cgctacccca cctggggcga caccgtggag gtggagtgtt ggatcggcgc cagcggcaac 3600

aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 3660

tgcacctccc tgagcgtgtt gatgaacacc cgcacccgcc gcctgagcac catcccgac 3720

gaggtgcgcg gcgagatcgg ccccgccctc atcgacaacg tggccgtgaa ggacgacgag 3780

atcaagaagc tgcagaagct gaacgactcc accgcccact acatccaggg cggcctgacc 3840

ccccgtgga acgacctgga cgtgaaccag cactgaaca acctgaagta cgtggcctgg 3900

gtgttcgaga ccgtgccga cagcatcttc gagleccacc acatcagctc ctccacctg 3960

gagtaccgcc gcgagtgcac ccgcgactcc gtgtgcgca gcctgaccac cgtgagcggc 4020

ggcagctccg aggccggcct ggtgtgcgac cacctgtctc agctggaggg cggcagcag 4080

gtgtgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttcg cggcacgac 4140

gtgatcccc cagagcccc cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200

cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260

atagtatcga cacactctgg acgctggteg tctgatggac tcttgcgcc acacttctg 4320

ccttgacctg tgaatatccc tgccgtttt atcaaacagc ctcaagtgtg ttgatctgt 4380

gtglacgcgc ttttgcagt tgctagctgc ttgtgclatt tgcaatacc acccccagea 4440

tccccctccc tegtctcata tctgttgcac cccaaccgca acttatctac gctgtctgc 4500

tatccctcag cgtgtctct gctctgtct actgcccctc gcacagcctt ggtttgggct 4560

ccgctgtat tctctggta ctgcaacctg taaaccagea ctgcaatgt gatgcacggg 4620

aagtagtggg atgggaacac aaatggaaag ctt 4653

<210> 87

<211> 4653

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 87

ggtaccgcc tgcaacgcaa gggcagccac agccgtccc accgcccgt gaaccgacac 60

gtgttgggc gctgccgcc tgctgccgc atgttgtgc tggtaggct gggcagtgt 120

gcatgtga ttgaggctg gtcatcggtg tgaagctta tgtgtgtgt gggcttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgagg cagccagagg gatitcagat gategcgcgt acaggttgag 300
 cagcagtgtc agcaaaggta gcagttgcc agaatgateg gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccgggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag categccgcc tggctgttgc agctacaacg 480
 gcaggagtea tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga 540
 tgcaagcatg cctggcgaag acacatgggt tgcggatgct gccggctgct gcctgtgcg 600
 cagcccggtg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgtgggt 720
 tgcatggcat cgcaccatca gcaggagcgc atggaaggg actggcccca tgcacgcat 780
 gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
 atgtcatgg acgcatgtg cgtgacgtc cttgacggc gctccctcgc ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt 1020
 cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgccgc ccctactccg categatgtt 1140
 ccatagtgca ttggactgca ttgggtggg gcggccggct gtttcttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtaggt ccccggtgcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgccggct cggaagtag gcgcccctct ttctgctcgc cctctctccg 1320

tcccgctct agaatacaa tgatcgagca ggacggcctc cagccggct ccccgccgc 1380

ctgggtggag cgctgttcg gctacgactg ggcccagcag accatggct gctccgacgc 1440

cgccgtgttc cgctgtccg cccagggccg ccccggtctg ttctgaaga ccgacctgc 1500

cgcgccctg aacgagctgc aggacgagc cgcccgctg tctggctgg ccaccaccg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggtgctgct 1620

ggcgaggtg cccggccagg acctgtctgc ctcccactg gccccgccg agaaggtgc 1680

catcatggcc gacgcatgc gccgctgca cacctggac cccgccacct gcccttga 1740

ccaccaggcc aagcaccgca tcgagcggc cgcacccgc atggaggccg gcctggtgga 1800

ccaggacgac ctggacgagg agcaccagg cctggcccc gccgagctgt tcgccgcct 1860

gaaggccgc atccccgacg gcgaggacct ggtggtgacc cagggcgac cctgcctgcc 1920

caacatcatg gtggagaacg gccgttctc cggttctc gactgcggcc gctgggcgt 1980

ggccgaccgc taccaggaca tcgcccggc caccgcgac atcgccgagg agctgggcgg 2040

cgagtgggcc gaccgttcc tgggtctgta cggcatgcc gccccgact cccagcgc 2100

cgccttctac cgctgtctgg acgagttct ctgacaattg gcagcagcag ctggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgetgccttg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat ctgtgtgta 2280
 cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcaccccc 2340
 ttccctcggt tcatacgct tgcaccctaa ccgcaactta tctacgctgt cctgetatcc 2400
 ctccagcgtg ctctgtctcc tgetcactgc cctcgcaca gccttgggtt gggtccgcc 2460
 tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
 gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttgggt ctctgtccat tagcgaagcg tccggttcac acacgtgcca cgltggcgag 2700
 gtggcagggtg acaatgatcg gtggagctga tggcgaaac gtacacagcc tagggatata 2760
 gaattccttt ctctgctat gacactcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgct gcatgcaaca ccgatgatgc ttgaccccc cgaagctctt tcggggctgc 2880
 atgggcgctc cgatccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctatgactact accacttcta 3000
 cacagccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctctctgtt 3060
 cagtcacaac ccgcaaacac tagtatggcc accgcaccca cttctcggc gtcaatgcc 3120

cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccage gagggccctc 3180

cccgtgcgcg ggcgcgcccc cgactggtec atgctgttcg ccgtgateac caccatcttc 3240

agcggccgccc agaagcagtg gaccaacctg gagtgggaagc ccaagcccaa gctgcccag 3300

ctgctggacg accacttcgg cctgcacggc ctggtgttc ggcgcacctt cgccatcgc 3360

tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag 3420

gaggccaccc tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc 3480

ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgc gcacccacgt ggccgtggag 3540

cgctacccca cctggggcga caccgtggag gtggagtgt ggatcggcgc cagcggcaac 3600

aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgc 3660

tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac 3720

gaggtgcgcg gcgagatcgg ccccgccctc atcgacaacg tggccgtgaa ggacgacgag 3780

atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840

ccccgtgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg 3900

gtgttcgaga ccgtgcccga cagcatcttc gatlccacc acatcagctc ctccacctg 3960

gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcga gcctgaccac cgtgagcggc 4020

ggcagctccg aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag 4080

gtgctgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttcg cggcacacgc 4140

gtgatecccc ccgagccccc cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200

cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260

atagtatcga cacactctgg acgctggcgc tgtgatggac tgttgcgcgc acacttgctg 4320

ccttgacctg tgaatatecc tgccgcttct atcaaacagc ctcagtgtgt ttgatcttgt 4380

gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcaataacc acccccagca 4440

tccccctccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 4500

tatccctcag cgtgtctcct gctcctgctc actgcccctc gcacagcctt gggttgggct 4560

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620

aagtagtggg atgggaacac aaatggaaag ctt 4653

<210> 88

<211> 4647

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 88

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt 120
 gccatgtga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgc 180
 gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
 tccagggtgt tcaatcgcg cagccagagg gatttcagat gatcgcgct acaggtgag 300
 cagcagtgc agcaaaggta gcagttgcc agaattgatc gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagt 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcg gcgggcgtga 540
 tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggtgct gcctgtgcg 600
 cagccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgtgggt 720
 tgcattgcat cgcaccatca gcaggagcg atgcaaggg actggccca tgcagccat 780
 gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg ccgcaagct ttgcagaacc 840
 atgtcatgg acgcatgtag cgctgacgtc cctgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagagggcg gcgcagag cggcggcagc agcggcggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgtgatcc caaacagtt 1020

cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgctc ccctactccg catcgatgtt 1140

ccatagtga ttggactgca ttgggtggg gcggccggt gttctttcg tttgcaaaa 1200

cgcgccagct cagcaacctg tccgtgggt ccccggtcc gatgaaatcg tgtgcacgcc 1260

gatcagctga ttgccgggt cgcgaagtag ggcctcct tctgctcg cctctctcg 1320

tccgcctct agaatacaa tgategagca ggacggctc cagccgggt ccccgccgc 1380

ctgggtggag cgctgttc gctacgactg ggcccagcag accatcggt gctccgacgc 1440

cgccgtgtc cgctgtccg ccagggccg cccgtgtg ttctgaaga ccgacctgc 1500

cggcgccctg aacgagctgc aggacgagc cggccgctg tctggtgg ccaccaccg 1560

cgtgccctgc gccgccgtg tggacgtgt gaccgagcc ggccgcgact ggtgtgtgt 1620

gggcgaggtg cccggccagg acctgtgtc ctccacctg gccccgcg agaaggtgc 1680

catcatggc gacgcatgc gccgctgca caccctggac ccgcccact gcccttga 1740

ccaccaggcc aagcacgca tcgagcgcg ccgcacccg atggagggc gcctggtgga 1800

ccaggacgac ctggacgagg agcaccagg cctggcccc gccgagctgt tcggcgct 1860

gaagggcgc atggcgacg gcgaggacct ggtggtgacc cagggcgac cctgcctgcc 1920

caacatcatg gtggagaac gccgttctc cggttcac gactggggc gcctgggcgt 1980

ggccgaccgc taccaggaca tcgccctggc caccgcgac atgcccgagg agctgggcgg 2040

cgagtgggccc gaccgcttcc tgggtgctgta cggcatcgcc gcccccgact cccagcgcac 2100

cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagt 2160

atcgacacac tctggacgct ggctggtga tggactgtg ccgccacact tctgccttg 2220

acctgtgaat atccctgccg cttttataca acagcctcag tgtgttgat cttgtgtga 2280

cgcgcttttg cgagttgcta gctgctgtg ctatttgcga ataccacccc cagcaccccc 2340

tccctcgtt tcatacgtt tgcacccaa ccgcaactta tctacgtgt cctgctatcc 2400

ctcagcgtg ctctgctcc tctcactgc cctcgcaca gccttggtt gggtccgccc 2460

tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctogaac agagcgcgca gaggaacgt 2580

gaaggtctcg cctctgtcg acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt ctctgctcat tagcgaagcg tccgggtcac acacgtgcca cgttggcgag 2700

gtggcagggtg acaatgatcg gtggagctga tggcgaac gttcacagcc tagggatata 2760

gaattccttt ctgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgtt gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000

cacaggccac tcgagcttgt gatgcactc cgctaagggg ggcctcttc ctcttcgttt 3060

cagtcacaac ccgcaaacac tagtatggt tccgggcat tcacatgtc ggcgtgcccc 3120

gcgatgactg gcaggggccc tggggcacgt cgtccggac ggccagtcgc cacccgcctg 3180

agggggcgcg cccccgactg gtccatgtg ttcccgctga tcaccacat cttcagcgcc 3240

gccgagaagc agtggaccaa cctggagtgg aagcccaagc ccaagctgcc ccagctgctg 3300

gacgaccact tcggcctgca cggcctgggt ttccgcgca ccttcgcat cgcctctac 3360

gaggtgggcc ccgaccgcag cacctccatc ctggccgtga tgaaccacat gcaggaggcc 3420

acctgaacc acgccaagag cgtgggcac cgtggcgacg gcttcggcac caccctggag 3480

atgtccaagc ggcacctgat gtgggtggtg cgcgcaccc acgtggccgt ggagcgctac 3540

cccacctggg gcgacaccgt ggaggtggag tgctggatcg gcgccagcgg caacaacggc 3600

atgcgccgag acttcctggt gcgcgactgc aagaccggcg agatcctgac ccgtgcacc 3660

tcctgagcg tgetgatga caccgcacc cgcgcctga gcaccatcc cgacgaggtg 3720

cgcggcgaga tcggccccgc ctcatcgac aacgtggccg tgaaggacga cgagatcaag 3780

aagctgcaga agctgaacga ctccaccgcc gactacatcc agggcggeet gacccccgc 3840

tggaacgacc tggacgtgaa ccagcacgtg aacaacctga agtacgtggt ctgggtgttc 3900

gagaccgtgc ccgacagcat cttcgagtc caccacatca gctccttcac cctggagtac 3960

cgccgcgagt gcaccccgga clccgtgctg cgcagcctga ccaccgtgag cggcggcagc 4020

tccgaggccg gcctgggtgtg cgaccacctg ctgcagctgg agggcggcag cgaggtgctg 4080

cgcgcccgca ccgagtggcg ccccaagctg accgactcct tccgcggcat cagcgtgatc 4140

cccgccgagc cccgcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac 4200

atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta 4260

tcgacacact ctggacgctg gtcgtgtgat ggactgttc cggcacactt gctgcctga 4320

cctgtgaata tccctgccgc tttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac 4380

gcgcttttc gagttgctag ctgctgtgc tatttgcgaa taccaccccc agcatcccct 4440

tccctegtlt catatcgtt gcateccaac cgcaacttat ctacgtgtc ctgtatccc 4500

tcagcgtgc tctgtcctt gctactgcc cctcgcacag ccttggtttg ggctccgcct 4560

gtattctct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag 4620

tgggatggga acacaaatgg aaagctt 4647

<210> 89

<211> 4721

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 89

ggtacccgcc tgcaacgcaa gggcagccac agccgcctcc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

gccgggcaat gcgcatggcg gcaagagggc ggcagcactt gctggagctg ccgcgggtgcc 240

tccaggtggt tcaatcgccg cagccagagg gatttcagat gatcgccgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgc agaagatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcacgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttgc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcggggcgtga 540

tgcaagcatg cctggcgaag acacatggcg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgtctgggt 720

tgcattggcat cgcacccatca gcaggagcgc atgcgaaggg actggcccca tgcaagccat 780

gccaaaccgg agcgcaccca gtgtccacac tgcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatec caaacgagtt 1020
 cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgctc cctactccg catcgatgtt 1140
 ccatagtgca ttggactgca ttgggtggg gcggccggct gttctttcg tgttcaaaa 1200
 cgcgccagct cagcaacctg tcccgtaggt ccccgtagc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgccggct cgcgaagtag gcgcctcct ttctgctcg cctctctccg 1320
 tccgcctct agaatacaa tgatcgagca ggaaggctc cagccggct ccccgccgc 1380
 ctgggtggag cgcctgttc gctacgactg ggcccagcag accatcggt gtcgcagc 1440
 cgccgtgtc cgcctgtcc cccagggccg ccccgtagt ttctgaaga ccgacctgc 1500
 cggegccctg aacgagctgc aggacgagc cgccgcctg tctggctgg ccaccaccg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgagggc ggccgcgact ggtgctgct 1620
 gggcgaggtg cccggccagg acctgtgtc ctccacctg gccccgccg agaaggtgc 1680
 catcatggc gacgcatgc gccgctgca caccctggac cccgccct gcccttga 1740
 ccaccaggcc aagcaccgca tcgagcgcg ccgcaccgc atggagggcg gctggtgga 1800

ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
 gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgtttctc cggcttcatc gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgcac atgcccgagg agctgggcgg 2040
 cgagtgggcc gaccgttcc tgggtctgta cggcatcgcc gccccgact cccagcgcct 2100
 cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagt 2160
 atcgacacac tctggacgct ggtcgtgtga tggactgtg ccgccacact tctgccttg 2220
 acctgtgaat atccctgccg cttttatcaa acagcctcag tctgtttgat cttgtgtgta 2280
 cgcgttttg cgagttgcta gctgcttggt ctatttgcga ataccacccc cagcaccccc 2340
 ttccctcgtt tcatacgtt tgcacccaa ccgcaactta tctacgtgt cctgcctacc 2400
 ctacgcgtg ctctgtctcc tctcactgc cctcgcaca gccttggtt gggtccgcc 2460
 tctattctcc tggactgca acctgtaaac cagcactgca atgtgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctogaac agagcgcgca gaggaacgt 2580
 gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttggtt ctctgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
 gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccagge ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacaac ctagatcaact accacttcta 3000
 cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
 cagtcacaac ccgcaaacgg cgcgccatgg ccaccacctc cctggcctec gccttctgca 3120
 gcatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc tccagcgacc 3180
 tgcagctgcg cgccggcaac gcccagacct ccctgaagat gatcaacggc accaagttct 3240
 cctacaccga gagcctgaag aagctgcccc actgggtccat gctgttcgcc gtgatcacca 3300
 ccatcttctc cgccgccgag aagcagtgga ccaacctgga gtggaagccc aagcccaacc 3360
 cccccagct gctggacgac cacttcggcc cccacggcct ggtgttcgc cgcaccttcg 3420
 ccatccgcag ctacgagggt ggccccgacc gctccaccag catcgtggcc gtgatgaacc 3480
 acctgcagga ggccgccctg aaccacgcca agtccgtggg catcctgggc gacggcttcg 3540
 gcaccaccct ggagatgtcc aagcgcgacc tgatctgggt ggtgaagcgc acccacgtgg 3600
 ccgtggagcg ctacccccgc tggggcgaca ccgtggaggt ggagtgtctg gtgggcgcct 3660
 ccggcaacaa cggccgccgc cagacttcc tgggtgcgca ctgcaagacc ggcgagatcc 3720

tgaccgctg cacciccctg agcgtgatga tgaacacccg caccgcccgc ctgagcaaga 3780
 tccccgagga ggtgcgcggc gagatcgcc ccgcctcat cgacaacgtg gccgtgaagg 3840
 acgaggagat caagaagccc cagaagctga acgactccac cgccgactac atccagggcg 3900
 gcctgacccc ccgtggaac gacctggaca tcaaccagca cgtgaacaac atcaagtacg 3960
 tggactggat cclggagacc gtgcccgaca gcatcttga gagccaccac atctctct 4020
 tcaccatga gtaccgccc gagtgcacca tggacagcgt gctgcagtcc ctgaccaccg 4080
 tgagcggcgg ctctccgag gccggcctgg tgtgcgagca cctgctgcag ctggagggcg 4140
 gcagcgaggt gctgcgcgcc aagaccgagt ggcgcccac gctgaccgac tcttcccg 4200
 gcatcagcgt gatccccgc gagtccagcg tgatggacta caaggaccac gacggcgact 4260
 acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgatgactc gaggcagcag 4320
 cagctcggat agtatcgaca cactctggac gctggctgtg tgatggactg ttgccgccac 4380
 acttgctgcc ttgacctgtg aatatccctg ccgctttat caaacagcct cagtgtgtt 4440
 gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtctatttg cgaataccac 4500
 cccagcacc ccttccctc gttcctatc gcttgcaccc caaccgcaac ttatctacgc 4560
 tgctctgta tccctcagcg ctgctctgc tctgtctac tgcctctgc acagccttg 4620
 ttgggctcc gcctgtatc tctgtgtact gcaacctgta aaccagcact gcaatgctga 4680

tgacagggaa gtagtgggat gggaacacaa atggaaagct t

4721

<210> 90

<211> 4650

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 90

ggtacccgcc tgcaacgeaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgc 120

gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgc 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcgggtgcc 240

tccagggtgt tcaatcgagg cagccagagg gatctcagat gatcgcgct acaggttgag 300

cagcagtgc agcaaaggta gcagtttccc agaattgatc gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggtgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600

cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg 660

caatgtgtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720

tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcaagccat 780

gccaaccggg agcgcaccga gtgtccacac tgcaccagg cccgcaagct ttgcagaacc 840

atgctcatgg acgcatgtag cgtgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900

acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960

gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt 1020

cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc cctactccg catgatgtt 1140

ccatagtgca ttggactgca ttgggtggg gcggccgggt gtttcttcg tgttgcaaaa 1200

cgcgccagct cagcaacctg tcccgtaggt ccccggtgc gatgaaatc tgtgcacgc 1260

gatcagctga ttgccggct cgcgaagtag gcgccctct tctgctcgc cctctctcg 1320

tccgcctct agaatatcaa tgatgagca ggacggcctc cagccgggt ccccgccgc 1380

ctgggtggag cgcctgttc gctacgactg ggcccagcag accatcggt gctccgacgc 1440

cgcctgttc cgcctgttc cccaggccg ccccgctg tctgtgaaga ccgacctgc 1500

cggcgccctg aacgagctgc aggaagagc cggcgccctg tcttggtgg ccaccaccg 1560

cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 gggcgagggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgc 1680
 catcatggcc gacgccatgc gccgcctgca cacctggac cccgccacct gcccttcga 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgccgcct 1860
 gaaggccgc atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgectgcc 1920
 caacatcatg gtggagaacg gccgttctc cggttcacg gactgcggcc gcctgggcgt 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atgccgagg agctgggcgg 2040
 cgagtgggcc gaccgttcc tggctgtga cggcatgcc gccccgact cccagcgc 2100
 cgcttctac cgctgctgg acgagttct ctgacaattg gcagcagcag ctggatagt 2160
 atcgacacac tctggacgt ggtcgtgtga tggactgtg ccgccacct tctgccttg 2220
 acctgtgaat atccctgcc ctttatcaa acagcctcag tgtgttgat cttgtgtga 2280
 cgcgctttg cgagttgcta gctgctgtg ctatttgcga ataccaccc cagcacc 2340
 ttccctggt tcatacgt tgcaccaca ccgcaacta tctacgtgt cctgctatcc 2400
 ctacgcgtg ctctgctcc tctcactgc ccctgcaca gccttggtt gggtccgcc 2460
 tgtattctc tggactgca acctgtaa acagcactgca atgctgatgc acgggaagta 2520

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580

gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640

gcgcttggtt cttcgtccat tagcgaagcg tccgggtcac acacgtgcc a gttggcgag 2700

gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gltcacagcc tagggatata 2760

gaattccttt cttgcgtat gacactcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tccgggctgc 2880

atgggcgctc cgatcccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcaact accacttcta 3000

cacagggcac tegagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttgctt 3060

cagtcacaac ccgcaaacac tagtatgacg ttccgggtcg cctcccggc catgggcgc 3120

ggtgtctccc tccccggcc cagggtcgcg gtgcgcgcc agtcggcgag tcaggttttg 3180

gagagcgggc gcgccccga ctggtccatg ctgttcgcg tgatcaccac catctctec 3240

gccgccgaga agcagtggac caacctggag lggaagccca agcccaacc cccccagtg 3300

ctggacgacc acttcggccc ccacggcctg gtgtccgcc gcaccttcg catccgcagc 3360

tacgaggtgg gccccgaccg ctccaccagc atcgtggcgg tgatgaacca cctgcaggag 3420

gccgccctga accacgcca gtcgtgggc atctgggcg acggttcgg caccacctg 3480

gagatgtcca agcgcgacct gatctgggtg gtgaagcgca cccacgtggc cgtggagcgc 3540
 taccgccct ggggcgacac cgtggagggtg gagtgcctggg tgggcgcctc cggcaacaac 3600
 ggccgccgcc acgacttctt ggtgcgcgac tgcaagaccg gcgagatcct gaccgctgc 3660
 acctccctga gcgtgatgat gaacaccgc acccgccgcc tgagcaagat ccccgaggag 3720
 gtgcgcggcg agatcgccc cgcttcac gacaacgtgg cctgaagga cgaggagatc 3780
 aagaagcccc agaagctgaa cgactccacc gccgactaca tccagggcgg cctgaccccc 3840
 cgctggaacg acctggacat caaccagcac gtgaacaaca tcaagtacgt ggactggatc 3900
 ctggagaccg tgcccgacag catcttcgag agccaccaca tctctctctt caccatcgag 3960
 taccgcccg agtgcacat ggacagcgtg ctgcagtcct tgaccaccgt gageggcggc 4020
 tcctccgagg ccggcctggt gtgcgagcac ctgctgcagc tggagggcgg cagcgagggtg 4080
 ctgcgcgcca agaccgagtg gcgccccaa ctgaccgact cctccgcgg catcagcgtg 4140
 atccccgccg agtccagcgt gatggactac aaggaccacg acggcgacta caaggaccac 4200
 gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata 4260
 gtatcgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgctgcct 4320
 tgacctgtga ataccctgc cgttttacc aaacagcctc agtgtgttg atctgtgtg 4380
 tacgccttt tgcgagtgc tagctgcttg tctatttgc gaataccacc cccagcatcc 4440

ccttccctcg ttccatateg cttgcatccc aaccgcaact tatctacgct gtccctgctat 4500

ccctcagegc tgcctctgct cctgctcact gcccctcgca cagccttggt ttgggctccg 4560

cctgtattct cctgggtactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620

tagtgggatg ggaacacaaa tggaaagctt 4650

<210> 91

<211> 4653

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 91

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60

gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgt 120

gccatgctga ttgaggcttg gtccatcggg tggaaagctta tgtgtgtgct gggttgcac 180

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240

tccaggtggt tcaatcgagg cagccagagg gatctcagat gatcgcgcgt acaggttgag 300

cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360

cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgttcc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcggggcgtga 540
 tgcaagcatg cctggcgaag acacatgggtg tgcggatgct gccggctgct gcctgctgcg 600
 cagcccggtg agttggcagc aggtccagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
 tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat 780
 gccaaaccgg agcgcaccga gtgtccacac tgcaccagg ccgcaagct ttgcagaacc 840
 atgctcatgg acgeatgtag cgctgacgtc cctgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccage cagcgggtgat gcgtgatcc caaacgagtt 1020
 cacattcatt tgcattcctg gagaagegag gctggggcct ttgggctggt gcagcccgca 1080
 atggaatgcg ggaccgccag gctagcagca aaggcgctc ccctactccg catgatgtt 1140
 ccatagtga ttggactgca ttgggtggg gcggccggct gttctttcg tgttgcaaaa 1200
 cgcgccagct cagcaacctg tccgtgggt ccccggtgcc gatgaaatcg tgtcacgcc 1260
 gatcagctga ttgccggct cgcgaagtag gcgcctcct ttctgctgc cctctctcg 1320
 tccgcctct agaatatcaa tgategagca ggacggctc cagccggct ccccgccgc 1380

ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
 cgccgtgttc cgcctgtccg cccagggccg ccccgctctg ttcgtgaaga ccgacctgtc 1500
 cggcgccctg aacgagctgc aggacgaggc cgcccgctg tcttggtgg ccaccaccg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggcccgact ggtgtgtct 1620
 gggcgaggtg cccggccagg acctgtgtc ctccacctg gccccgccg agaaggtgtc 1680
 cateatggcc gacgccatgc gccgectgca cacctggac ccgccacct gcccttga 1740
 ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccg atggaggccg gcttgttga 1800
 ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgccgct 1860
 gaagggccgc atgccgacg gcgaggacct ggtgttgacc cagggcgacg cctgcctgcc 1920
 caacatcatg gtggagaacg gccgtttc cggttcac gactggggcc gcctggggct 1980
 ggccgaccgc taccaggaca tcgccctggc caccgcgac atgccgagg agctggggcg 2040
 cgagtgggcc gaccgttcc tgggtgtgta cggcatgcc gccccgact ccagcgcat 2100
 cgccttctac cgcctgttg acgagttct ctgacaattg gcagcagcag ctcgatagt 2160
 atcgacacac tctggaacct ggtcgtgtga tggactgtg ccgccacct tgetgcctg 2220
 acctgtgaat atccctgccg cttttatcaa acagctcag tgtgtttgat cttgtgtga 2280
 cgcgttttg cgagttgta gtgtgtgtg ctatttgcga ataccacccc cagcatcccc 2340

ttcctcgtt tcatacgtt tgcacccaa ccgaactta tetaeactgt cctgcatac 2400
 ctacagctg ctctgctcc tgcactgc cctgcaca gccttggtt gggtccgcc 2460
 tgtattccc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
 gtgggatggg aacacaaatg gaggatcccg cgtctegaac agagcgcgca gaggaacgt 2580
 gaaggtctcg cctctgtgc acctacgcgc ggcatacacc acaataacca cctgacgaat 2640
 gcgcttggtt ctctgctcc tagcgaagcg tccggtcac acacgtgcca cgttggcgag 2700
 gtggcaggtg acaatgatcg gtggagctga tggcgaac gttcacagcc tagggatac 2760
 gaattcctt ctgcgtat gacactcca gcaaaaggta gggcgggctg cgagacggct 2820
 tcccggcgt gcatacaaca ccgatgatgc ttcgacccc cgaagctcct tcggggctgc 2880
 atgggcgctc cgatccgct ccagggcgag cgtgtttaa atagccaggc ccccgattgc 2940
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctatgaact accacttcta 3000
 cacaggccac tcgagctgt gatgcactc cgctaagggg gcgcctctc ctctcgtt 3060
 cagtcacaac ccgcaaacac tagtatggtt atcaagacga acaggcagcc tgtggagaag 3120
 cctccgttca cgatcgggac gctgcgcaag gccatcccc cgcactgtt cgagcgtcg 3180
 gcgcttctg ggcgcgcgcc cgactggtcc atgctgttc cgtgatcac caccatctc 3240
 tccgccgcc agaagcagtg gaccaacctg gagggaagc ccaagccaa cccccccag 3300

ctgctggacg accacttcgg cccccaacggc ctggtgttcc gccgcacctt egccatccgc 3360
 agctacgagg tgggccccga ccgtccacc agcatcgtgg ccgtgatgaa ccacctgcag 3420
 gaggccgccc tgaaccacgc caagtccgtg ggcatacctgg gcgacggctt cggcaccacc 3480
 ctggagatgt ccaagcgcga cctgatctgg gtggtgaagc gcacccacgt ggccgtggag 3540
 cgctaccccg cctggggcga caccgtggag gtggagtgtt ggggtgggcgc ctccggcaac 3600
 aacggccgcc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgcg 3660
 tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa galccccgag 3720
 gaggtgcgcg gcgagatcgg ccccgccctc atcgacaacg tggccgtgaa ggacgaggag 3780
 atcaagaagc ccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840
 ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg 3900
 atcctggaga ccgtgccccg cagcatcttc gagagccacc acatctcttc ctccaccatc 3960
 gagtaccgcc gcgagtgcac catggacagc gtgctgcagt cctgaccac cgtgagcggc 4020
 ggctcctccg aggccggcct ggtgtgcgag cacctgtctc agctggaggg cggcagcgag 4080
 gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttcgg cggcatcagc 4140
 gtgatcccg ccgagtcag cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200
 cagcatcgc actacaagga cgacgacgac aagtgatgac tegaggcagc agcagctcgg 4260

atagtatcga cacactctgg acgetggteg tgtgatggac tttgcccgc acacttctg 4320

ccttgacctg tgaatatecc tgccgetttt atcaaacage ctcagtgtgt ttgatcttgt 4380

gtgtacgcgc ttttgcgagt tgctagctgc ttgtctatt tgcgaatacc acccccagca 4440

tccccctccc teglttcata tegettgcat cccaaccgca acttatctac gctgtcctgc 4500

tateccctcag cgetgtctct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620

aagtagtggg atgggaacac aaatggaaag ctt 4653

<210> 92

<211> 1893

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 92

gaattccttt ctgcgctat gacactcca gcaaaaggta gggcgggctg cgagacggct 60

tcccggcgct gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tcggggctgc 120

atgggcgctc cgatccgct ccagggcgag cgetgtttaa atagccaggc ccccgattgc 180

aaagacatta tagcgagcta ccaaagccat attcaaacac ctgatcact accacttcta 240

cacaggccac tcgagcttgt gatgcactc cgctaagggg gcgcctcttc ctcttcgttt 300
 cagtcacaac ccgcaaacac tagtatggcc accgcatcca cttctcggc gttcaatgcc 360
 cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggccctc 420
 cccgtgcgcg ggcgcccc cgactggtec atgctgttcg ccgtgateac caccatcttc 480
 tccgccgcg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa cccccccag 540
 ctgtggacg accacttcgg cccccacggc ctggtgttc gccgcacct cgcctccgc 600
 agctacgagg tgggccccga ccgtccacc agcatcgtgg ccgtgatgaa ccacctgcag 660
 gaggccgcc tgaaccacgc caagtcctg ggcacctgg gcgacggctt cggcaccacc 720
 ctggagatgt ccaagcgga cctgatctgg gtggtgaagc geaccacgt ggccgtggag 780
 cgctaccccg cctggggcga caccgtggag gtggagtgt ggggtgggcgc ctccggcaac 840
 aacggccgcc gccacgactt cctggtgcgc gactgaaga ccggcgagat cctgaccgc 900
 tgcacctccc tgagcgtgat gatgaacacc cgcaccgcc gcctgagcaa gatcccgag 960
 gaggtgcgcg gcgagatcgg cccgccttc atcgacaacg tggccgtgaa ggacgaggag 1020
 atcaagaage ccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 1080
 cccgctgga acgacctgga catcaaccag cagtgaaaca acatcaagta cgtggactgg 1140
 atcctggaga ccgtgccga cagcatcttc gagagccacc acatctctc ctccaccatc 1200

gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc 1260
ggctcctccg aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcggag 1320
gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttcg cggcatcagc 1380
gtgatccccg ccgagtcag cgtgatggac tacaaggacc acgacggcga ctacaaggac 1440
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 1500
atagtatega cacactctgg acgtggteg tgtgatggac tgttgcgcc acactgctg 1560
cctgacctg tgaataccc tgcgccttt atcaaacagc ctacgtgtgt ttgatctgt 1620
gtgtacgcgc tttgcgagt tgctagctgc ttgtctatt tgcgaatacc acccccagca 1680
tcccttccc tcgttcata tcgcttgcac cccaaccga acttatctac gctgtcctgc 1740
tatccctcag cgctgctcct gctcctgctc actgccctc gcacagcctt ggtttgggct 1800
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 1860
aagtagtggg atgggaacac aaatggaaag ctt 1893

<210> 93

<211> 1887

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 93

```

gaattccttt cttgcgtat gacattcca gcaaaaggta gggcgggctg cgagacggct    60

tcccggegt gcatgcaaca ccgatgatgc ttgaccccc cgaagctcct tgggggtgc    120

atgggcgtc cgatgccgt ccagggcgag cgctgtttaa atagccaggc ccccgattgc    180

aaagacatta tagcgagcta ccaaagccat attcaaacac clagatcaact accattcta    240

cacaggccac tegagcttgt gatgcactc cgctaagggg gcgcctcttc ctctcgttt    300

cagtcacaac ccgcaaacac tagtatggt tccgcggcat tcacatgtc ggctgcccc    360

gcgatgactg gcagggtccc tggggcacgt cgctccggac ggccagtcgc caccgcctg    420

agggggcgcg ccccgactg gtccatgtg ttgcctga tcaccacat ctctccgcc    480

gccgagaagc agtggaccaa cctggagtgg aagcccaagc ccaaccccc ccagctgtg    540

gaagaccact tggccccca cggcctggtg ttccgcgca ccttcgcat ccgcagctac    600

gaggtgggcc ccgaccgtc caccagcatc gtggcctga tgaaccacct gcaggaggcc    660

gccctgaacc acgccaagtc cgtgggcac cttggcgac gcttcggcac caccctggag    720

atgtccaagc gcgacctgat ctgggtggtg aagcgcacc acgtggcct ggagcgctac    780

ccgcctggg gcgacaccgt ggaggtggag tgcgtgggtg gcgcctccg caacaacggc    840

cgccgccacg acttctggt gcgcgactgc aagaccggcg agatcctgac ccgtgcacc    900

```

tccctgagcg tgaatgatgaa cacccgcacc cgccgcctga gcaagatccc cgaggagggtg 960
 cgccggcgaga tcggccccgc ctccatcgac aacgtggccg tgaaggacga ggagatcaag 1020
 aagccccaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc 1080
 tggaaacgacc tggacatcaa ccagcacgtg aacaacatca agtacgtgga ctggatcctg 1140
 gagaccgtgc ccgacagcat ctccgagagc caccacatct cctccttcac catcgagtac 1200
 cgccgcgagt gcacatgga cagcgtgctg cagtcctga ccaccgtgag cggcggtcc 1260
 tccgaggccg gcctggtgtg cgagcacctg ctgcagctgg agggcggcag cgaggtgctg 1320
 cgcgccaaga ccgagtggcg ccccaagctg accgactcct tccgcggcat cagcgtgatc 1380
 cccgccgagt ccagcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac 1440
 atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta 1500
 tcgacacact ctggacgctg gtcgtgtgat ggactgttc cgccacactt gctgccttga 1560
 cctgtgaata tccctgccgc tttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac 1620
 gcgctttgc gatttgctag ctgcttgtgc tatttgcgaa taccacccc agcatcccct 1680
 tccctcgttt catatcgtt gcateccaac cgcaactat ctacgtgtc ctgtatccc 1740
 tcagcgtgc tcctgtcct gctactgcc cctgcacag ccttggttg ggtccgcct 1800
 gtatttctct ggtactgcaa cctgtaaacc agcactgcaa tgcgatgca cgggaagtag 1860

tgggatggga acacaaatgg aaagctt 1887

<210> 94

<211> 32

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 94

gacagaatt ccgcctgcaa cgcaagggca gc 32

<210> 95

<211> 30

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

<400> 95

gcatactagt ggcgggacgg agagagggcg 30

<210> 96

<211> 4

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<400> 96

Lys Asp Glu Leu

1

MODIFICACIÓN
EFECTUADA EN FASE
INTERNACIONAL

**MODIFICACIÓN EFECTUADA EN FASE
INTERNACIONAL BAJO EL ART. 34**

2. Expression in Mitochondria

[0171] In another preferred embodiment of the present invention, the expression of a polypeptide in a microorganism is targeted to mitochondria. Methods for targeting foreign gene products into mitochondria (Boutry *et al.*, *Nature* (London) (1987) 328:340-342) have been described, including in green microalgae (see for example *Mol Gen Genet.*, 1993 Jan;236(2-3):235-44).

[0172] For example, an expression vector encoding a suitable secretion signal can target a heterologous protein to the mitochondria. An exemplary expression vector for mitochondria targeting is the IMPACTVECTOR1.5 vector, from Wageningen UR-Plant Research International, which uses the yeast CoxIV secretion signal. This expression vector was shown to deliver proteins in the mitochondrial matrix. The protein is fused to the first 4 amino acids of the yeast CoxIV protein in order to allow proper processing of the signal peptide (Kohler *et al.*, *Plant J.* 11: 613-621 (1997)). Other mitochondrial targeting sequences are known, including those functional in green microalgae. For example, see *FEBS Lett.* 1990 Jan 29;260(2):165-8; and *J. Biol. Chem.* 2002 Feb 22;277(8):6051-8.

[0173] Proteins expressed in the nuclear genome of *Chlorella* can be targeted to the mitochondria using mitochondrial targeting signals. Details of mitochondrial protein targeting and transformation are provided herein.

3. Expression in Endoplasmic Reticulum

[0174] In another preferred embodiment of the present invention, the expression of a polypeptide in a microorganism is targeted to the endoplasmic reticulum (ER). The inclusion of an appropriate retention or sorting signal in an expression vector ensure that proteins are retained in the endoplasmic reticulum (ER) and do not go downstream into Golgi. For example, the IMPACTVECTOR1.3 vector, from Wageningen UR-Plant Research International, includes the well known KDEL (SEQ ID NO:96) retention or sorting signal. With this vector, ER retention has a practical advantage in that it has been reported to improve expression levels 5-fold or more. The main reason for this appears to be that the ER contains lower concentrations and/or different proteases responsible for post-translational degradation of expressed proteins than are present in the cytoplasm. ER retention signals functional in green microalgae are known. For example, see *Proc. Natl Acad. Sci. USA.* 2005 Apr 26;102(17):6225-30.

5' TGACCTAGGTGATTAATTAAGTCGAGGCAGCAGCAGCTCGGATAGTATCG 3'
(SEQ ID NO:47)

Reverse:

5' CTACGAGCTCAAGCTTTCCATTTGTGTTC CCATCCCACTACTTCC 3'
(SEQ ID NO:48)

[0443] The *Chlorella sorokiniana* glutamate dehydrogenase promoter/UTR was obtained via PCR of genomic DNA isolated from UTEX strain 1230 via PCR using the following primer pairs:

Forward: 5' GATCAGAATTCCGCCTGCAACGCAAGG GCAGC 3' (SEQ ID NO:94)

Reverse: 5' GCATACTAGTGGCGGGACGGAGAGA GGGCG 3' (SEQ ID NO:95)

[0444] Codon optimization was based on the codons for *Prototheca moriformis*. The sequence of the non-codon optimized neomycin phosphotransferase (nptII) cassette was synthesized as an *AscI*-*XhoI* fragment and was based on upon the sequence of Genbank Accession No. YP_788126. The codon optimized nptII cassette was also based on this Genbank Accession number.

[0445] The three *Prototheca* strains were transformed using biolistic methods described above. Briefly, the *Prototheca* strains were grown heterophically in liquid medium containing 2% glucose until they reached the desired cell density (1×10^7 cells/mL to 5×10^7 cells/mL). The cells were harvested, washed once with sterile distilled water and resuspended at 1×10^8 cells/mL. 0.5mL of cells were then spread out on a non-selective solid media plate and allowed to dry in a sterile hood. The cells were bombarded with the PDS-1000/He Biolistic Particle Delivery System (BioRad). The cells were allowed to recover at 25° C for 24 hours. Upon recovery, the cells were removed by washing plates with 1 mL of sterile media and transferring to fresh plates containing 100 µg/mL G418. Cells were allowed to dry in a sterile hood and colonies were allowed to form on the plate at room temperature for up to three weeks. Colonies of UTEX 1435, UTEX 1439 and UTEX 1437 were picked and spotted on selective agar plates for a second round of selection.

[0446] A subset of colonies that survived a second round of selection described above, were cultured in small volume and genomic DNA and RNA were extracted using standard molecular biology methods. Southern blots were done on genomic DNA extracted from untransformed (WT), the transformants and plasmid DNA. DNA from each sample was run on 0.8% agarose gels after the following treatments: undigested (U), digested with *AvrII* (A),

cttgggttgggctccgctgtattctcctgtgtactgcaacctgtaaacagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaat
ggaaagcttgagctc

SEQ ID NO: 52

Beta-tubulin::neo-opt::nit

gaattccttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttccggcgctgcatgcaacaccgatgat
gcttcgacccccgaagctccttcgggctgcatggcgctccgatgccgctccaggcgagcgctgtttaaatagccaggccccc
gattgcaaagacattatagcgagctaccaaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttggatc
gcactccgtiaagggggcgcttcttcttctgttcagtcacaacccgaaacggcgcccatatcaatgatcgagcaggacggcct
ccacggcgctccccggcgctgggtggagcgcttcttcggctacgactggcccgagcagaccatcggtctccgacggcgcc
gtgttcgctgttcggccaggcgcccgcttctgtgaagaccgacctgtccggcgccctgaacgagctgcaggacgaggc
cgcccgctgtcctggctggccaccaccggcgcttcccgcccgctgtggacgtgtgtgaccgaggcgccggcgactggctg
ctgctggcgaggtggcgccaggacgtgtcttccacctggccccggcgagaggtgtccatcatggccgacgccatggc
cgctgcacacctggacccggccacctgcccccttcgaccaccaggccaagcaccgcatcgagcgcgcccgacccgcatgga
ggcgcgctgtgtggaccaggacgacctggacgaggagcaccaggcgctggccccggcgagctgttcggcgctgaaggccc
gcatggcgacggcgaggacctgtgtgtgaccacggcgacgctgctgcccacatcatgttgagaacggcgcttctccgg
cttcatgactggcgcgctgtgtgtgacggcctaccaggacatcgccctggccacccggacatcgccgaggagctggcg
ggcgagtgggcgacggcttctgtgtgtgacggcctaccaggacatcgccctggccacccggacatcgcccttaccgctgtggcga
gttcttctgactgaggcagcagcgtcggatagtatcgacacacttggacgctggctgtgtgatggactgttggccacactgc
tgccttgacctgtaatatccctggcgctttatcaaacagcctcagtggtttgatcttgtgtacgcgtttgcgagttgtagctgctt
gtgctatttgcaataccacccccagcatcccttccctcgtttcatatcgcttgcacccaaccgcaacttatcagctgtcctgctatcc
ctcagcgctgtcctgtcctgtcactgccccctgcacagccttgggttgggctccgctgtattctcctgtgtactgcaacctgtaaac
agcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagcttgagctc

SEQ ID NO: 53

GCCGCGACTGGCTGCTGCTGG

SEQ ID NO: 54

AGGTCCTCGCCGTCGGGCATG

SEQ ID NO: 55

Crβ-tub::NCO-suc2::CvNitRed

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCTTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATATCAATG
CTTCTTCAGGCCCTTCTTTTCTTCTTGCTGGTTTTGCTGCCAAGATCAGCGCCTCT
ATGACGAACGAAACCTCGGATAGACCACTTGTGCACTTTACACCAAACAAGGGC
TGGATGAATGACCCCAATGGACTGTGGTACGACGAAAAAGATGCCAAGTGGCAT
CTGTACTTTCAATAACAACCCGAACGATACTGTCTGGGGGACGCCATTGTTTTGGG
GCCACGCCACGTCCGACGACCTGACCAATTGGGAGGACCAACCAATAGCTATCG
CTCCGAAGAGGAACGACTCCGGAGCATTCTCGGGTTCCATGGTGGTTGACTACAA

II.

2. Expresión en la mitocondria

En otra modalidad preferida de la presente invención, la expresión de un polipéptido en un microorganismo se dirige a la mitocondria. Se han descrito métodos para la selección de productos de genes extraños en las mitocondrias (Boutry y otros, *Nature* (Londres) (1987) 328:340-342),
 5 incluyendo en microalgas verdes (ver, por ejemplo, *Mol Gen Genet.*, 1993 ene;236(2-3):235-44).

Por ejemplo un vector de expresión que codifica una señal de secreción adecuada puede dirigir una proteína heteróloga a la mitocondria. Un vector de expresión ilustrativo para dirigir la mitocondria es el vector IMPACTVECTOR1.5, de Wageningen UR-Plant Research International, que usa la señal de secreción CoxIV de levadura. Este vector de expresión demostró entregar
 10 proteínas en la matriz mitocondrial. La proteína se fusiona con los 4 primeros aminoácidos de la proteína CoxIV de levadura para permitir el procesamiento adecuado del péptido señal (Kohler y otros, *Plant J.* 11: 613-621 (1997)). Otras secuencias de dirección mitocondrial se conocen, incluyendo aquellas funcionales en las microalgas verdes. Por ejemplo, ver *FEBS Lett.* 29 de enero de 1990 ;260(2):165-8; y *J. Biol. Chem.* 22 de febrero de 2002;277(8):6051-8.

15 Las proteínas expresadas en el genoma nuclear de *Chlorella* se pueden dirigir a la mitocondria usando señales mitocondriales dirigidas. Los detalles de la dirección y transformación de la proteína mitocondrial se proporcionan en la presente.

3. Expresión en el retículo endoplasmático

En otra modalidad preferida de la presente invención, la expresión de un polipéptido en un
 20 microorganismo se dirige al retículo endoplasmático (ER). La inclusión de una señal de retención o clasificación apropiada en un vector de expresión garantiza que las proteínas se retengan en el retículo endoplasmático (RE) y no vayan corriente abajo adentro de Golgi. Por ejemplo, el vector IMPACTVECTOR1.3, de Wageningen UR-Plant Research International, incluye la señal KDEL de retención o clasificación bien conocida. Con este vector, la retención en el RE tiene una ventaja
 25 práctica ya que se reportó que mejora los niveles de expresión 5 veces o más. La razón principal de esto parece ser que el RE contiene concentraciones más bajas y/o proteasas diferentes responsables de la degradación post-traducciona de las proteínas expresadas que están presentes en el citoplasma. Se conocen las señales de retención del RE funcionales en microalgas verdes. Por ejemplo, ver *Proc. Natl Acad. Sci. Estados Unidos.* 26 de abril de 2005;102(17):6225-30.

II.

5' TGACCTAGGTGATTAATTAACTCGAGGCAGCAGCAGCTCGGATAGTATCG 3' (sec. con núm. de ident.:47)

Inverso:

5' CTACGAGCTCAAGCTTTCCATTTGTGTTC CCATCCCACACTACTTCC 3'

5 (sec. con núm. de ident.:48)

[0001] El promotor/UTR de la glutamato deshidrogenasa de *Chlorella sorokiniana* se obtuvo a través de la PCR del ADN genómico aislado de la cepa UTEX 1230 a través de la PCR mediante el uso de los pares de cebadores siguientes:

Directo:5' GATCAGAATTCCGCCTGCAACGCAAGG GCAGC 3' (sec. con núm. de ident.:94)

10 Inverso:5' GCATACTAGTGGCGGGACGGAGAGA GGGCG 3' (sec. con núm. de ident.:95)

[0002] La optimización del codón se basó en los codones para *Prototheca moriformis*. La secuencia del casete de la neomicina fosfotransferasa (nptII) en la que no se optimizaron los codones se sintetizó como un fragmento AscI-XhoI y se basó en la secuencia con núm. de acceso YP_788126 del GenBank. El casete nptII con los codones optimizados se basó también en este número de acceso del Genbank.

[0003] Las tres cepas de *Prototheca* se transformaron usando los métodos biolísticos descritos anteriormente. En resumen, las cepas de *Prototheca* se cultivaron heterotróficamente en medio líquido que contenía 2% de glucosa hasta alcanzar la densidad celular deseada (1×10^7 células/ml a 5×10^7 células/ml). Las células se cosecharon, se lavaron una vez con agua destilada estéril y se resuspendieron a 1×10^8 células/ml. Después, 0.5 ml de células se extendieron en una placa con medio sólido no selectivo y se dejó secar en una campana estéril. Las células se bombardearon con el sistema biolístico de entrega de partículas PDS-1000/He (BioRad). Las células se dejaron recuperar a 25°C durante 24 horas. Tras la recuperación, las células se retiraron mediante el lavado de las placas con 1 ml de medio estéril y se transfirieron a nuevas placas que contenían 100 mg/ml de G418. Las células se dejaron secar en una campana estéril y se dejó que se formaran colonias en la placa a temperatura ambiente por hasta tres semanas. Las colonias de UTEX 1435, UTEX 1439 y UTEX 1437 se recogieron y se aplicaron en forma de puntos en placas selectivas de agar para una segunda ronda de selección.

II.

[0004] Un subconjunto de las colonias que sobrevivieron a una segunda ronda de selección descrita anteriormente, se cultivaron en pequeño volumen y se extrajo el ADN y ARN genómico por medio de técnicas de biología molecular. Se hizo la transferencia de tipo Southern en el ADN genómico extraído de las no transformadas (WT), los transformantes y el ADN del plásmido. El

5 ADN de cada muestra se corrió en 0.8% de geles de agarosa después de los siguientes tratamientos: sin digerir (U), digerido con AvrII (A),

cttggtttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaat
ggaaagcttgagetc

Sec. con núm. ident.: 52

Beta-tubulina::neo-opt::nit

gaattccttcttgcgtatgacactccagcaaaaggtagggcgggctgcgagacggcttcccggcgtgcatgcaacaccgatgat
gettgcacccccgaagctccttcggggctgcatgggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggccccc
gattgcaaagacattatagcgagctaccaaaagccatattcaaacacctagatcactaccactctacacaggccactcgagcttgatc
gcactccgctaaagggggcgcccttctctctcgttgcagtcacaacccgcaaacggcgcgccatataatgatcgagcaggacggcct
ccacggcggtccccgcgcctgggtggagcgcccttctcggctacgactggcccgagcagaccatcgctgctccgacggcgcc
gttttccgctgtccgcccaggcgcccgcccgctgtcttctgtaagaccgacctgtccggcgccctgaacgagctgcaggacgaggc
cgcccgctgtcttggctggccaccaccggcgctgcccggcgctgctggacgtgttgaccgagcgccggcgcgactggctg
ctgctggcgaggtgcccggccagacctgctgtcttcccacctggcccccggcgagaggtgtccatcatggccgacggccatgg
cgccctgcacacctggaccccgccacctgccccttgcaccaccaggccaagcaccgcacgagcgcccgcccgccatgga
ggccggcctgtgtggaccaggacgacctggacgaggagcaccaggcgctggcccccggcgagctgttgcggcctgaagccc
gcatgcccgacggcgaggacctgtgtgtgacctacggcgacctgctgcccacatcatgggtggagaaaggcgccgttctccgg
cttcatcactgcccggcgctgtgtgtgcccgacctaccaggacatcgccctggccaccggcgacatcgccgaggagctggcg
ggcgagtgggcgaccgcttctgtgtgtgtacggcatcgccgccccgactcccaaggcatcgcccttctaccggcctgctggacga
gttcttctgactcgaggcagcagcagctcggatagtatgcacacactggacgctggctgtgtgatggactgttgcggccacactgc
tgccttgacctgtgaatatccctgccgcttttatcaaacagccctcagtgtgttgatctgtgtgtacggcgttttgcgagtgctagctgctt
gtgctatttgcgaataccacccccagcacccttctcctcgtttcatalcgttgcacccaaccgcaacttatctacgctgtctgctatcc
ctcagcgctgtctctgtctctgctcactgcccctgcacagccttgggttgggctccgctgtattctcctgggtactgcaacctgtaaacc
10 agcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagcttgagetc

Sec. con núm. ident.: 53

GCCGCGACTGGCTGCTGCTGG

Sec. con núm. ident.: 54

AGGTCTCGCCGTCGGGCATG

15 Sec. con núm. ident.: 55

Crβ-tub::NCO-suc2::CvNitRed

II.

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATATCAATG
CTTCTTCAGGCCTTTCTTTTTCTTCTTGCTGGTTTTGCTGCCAAGATCAGCGCCTCT
ATGACGAACGAAACCTCGGATAGACCACTTGTGCACTTTACACCAAACAAGGGC
TGGATGAATGACCCCAATGGACTGTGGTACGACGAAAAAGATGCCAAGTGGCAT
CTGTACTTTCAATACAACCCGAACGATACTGTCTGGGGGACGCCATTGTTTTGGG
GCCACGCCACGTCCGACGACCTGACCAATTGGGAGGACCAACCAATAGCTATCG
CTCCGAAGAGGAACGACTCCGGAGCATTCTCGGGTTCCATGGTGGTTGACTACAA

Hayden 556

State of California

SECRETARY OF STATE

This Apostille only certifies the authenticity of the signature and the title of the public official who has signed the public document, and, where appropriate, the identity of the seal or stamp which the public document bears. This Apostille does not certify the authenticity of the underlying document for which it was issued. This Apostille is not valid for use anywhere within the United States of America, its territories or possessions. To verify the issuance of this Apostille, go to www.sos.ca.gov/business/notary/apostille-search/

APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. Country: United States of America
This public document
2. has been signed by J. Cohen
3. acting in the capacity of Notary Public, State of California
4. bears the seal/stamp of J. Cohen, Notary Public, State of California
CERTIFIED
5. At Sacramento, California
6. the 21st day of January 2011
7. by Deputy Secretary of State, State of California
8. No. 914898
9. Seal/Stamp:



10. Signature

John Bowen
Secretary of State

BY

AR



314

6
557

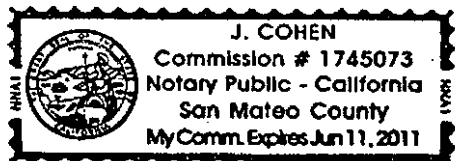
CALIFORNIA ALL-PURPOSE ACKNOWLEDGMENT

State of California

County of San Mateo

On Jan-07-2011 before me, J. Cohen, Notary Public

personally appeared Harrison F. Dillon



who proved to me on the basis of satisfactory evidence to be the person(s) whose name(s) is/are subscribed to the within instrument and acknowledged to me that he/she/they executed the same in his/her/their authorized capacity(ies), and that by his/her/their signature(s) on the instrument the person(s), or the entity upon behalf of which the person(s) acted, executed the instrument.

I certify under PENALTY OF PERJURY under the laws of the State of California that the foregoing paragraph is true and correct.

WITNESS my hand and official seal.

Signature: J. Cohen

Place Notary Seal and/or Stamp Above

Signature of Notary Public

OPTIONAL

Though the information below is not required by law, it may prove valuable to persons relying on the document and could prevent fraudulent removal and reattachment of this form to another document.

Description of Attached Document

Title or Type of Document: Power of Attorney

Document Date: Jan. 07, 2011 Number of Pages: 1+ Cert

Signer(s) Other Than Named Above: none

Capacity(ies) Claimed by Signer(s)

Signer's Name: Dillon Signer's Name: _____

☒ Corporate Officer — Title(s): President

☐ Individual # Chief Technology Off.

☐ Partner — ☐ Limited ☐ General

☐ Attorney in Fact

☐ Trustee

☐ Guardian or Conservator

☐ Other: _____

Signer Is Representing: _____

Solasyme, Inc.

RIGHT THUMBPRINT OF SIGNER

Top of thumb here

☐ Corporate Officer — Title(s): _____

☐ Individual

☐ Partner — ☐ Limited ☐ General

☐ Attorney in Fact

☐ Trustee

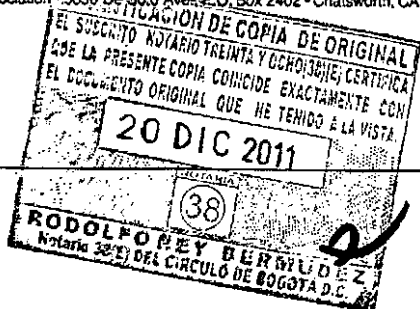
☐ Guardian or Conservator

☐ Other: _____

Signer Is Representing: _____

RIGHT THUMBPRINT OF SIGNER

Top of thumb here



315

POWER OF ATTORNEY

The undersigned,

SOLAZYME, INC.located at
225 Gateway Boulevard, South San Francisco, California
94080, Estados Unidos

declares that it hereby grants to

ÁLVARO CORREA-ORDÓÑEZ and/or
RICARDO METKE MENDEZ and/or
JUAN PABLO CONCHA-DELGADO and/or
OLGA GEORGETTE OTEROof the firm Baker & McKenzie Colombia S.A. members of Baker & McKenzie, domiciled at Av. 82 No. 10-62, 6th floor, of Bogotá D.C., Colombia, full and sufficient power of attorney to pursue before the competent authorities the following matters:

Trademark registrations and renewals; deposit and registration of commercial names and commercial ensigns; registration and renewal of domain names; obtainment of patents, utility models, industrial designs, integrated circuits layout-designs and vegetal varieties; records of mergers, changes of name, assignments, changes of domicile and address; security interests; recordal of license of use agreements, liens and other acts that affect or limit the mentioned industrial property rights; filing and prosecution of all matters relating to domain names; filing and prosecution of oppositions, cancellation actions, annulment actions, annulment and reinstatement of right action, precautionary injunctions, infringement actions, and all administrative and judicial actions available for the prosecution of proceedings related to industrial property matters, unfair competition and protection of consumers,

to which end it empowers them to file requests, petitions and lawsuits, appeals and objections, pay taxes, receive, including faculties to enter into settlement agreements, conciliations and transactions, abandon applications and renounce to intellectual property rights and, in general, to take all the actions before the competent authorities in order to duly protect our intellectual property. To the said mandatories sufficient power is granted hereby to receive notifications for and on behalf of the Company, to confer powers on those who are to represent the Company judicially and extrajudicially, to answer in court in the matter of all claims and demands that may be presented in connection with the intellectual property rights, and to do everything that may be required before administrative or judicial officers of any class to comply this power, giving them likewise faculty to substitute these presents and revoke such substitutions if necessary.

Done and subscribed in

On the

7th

day of Jan. of 20 11

Signature of the person Issuing the Power

La abajo firmada,

SOLAZYME, INC.domiciliada en
2251 Gateway Boulevard, South San Francisco, California
94080, Estados Unidos

declara por el presente que otorga a

ÁLVARO CORREA-ORDÓÑEZ y/o
RICARDO METKE MÉNDEZ y/o
JUAN PABLO CONCHA-DELGADO y/o
OLGA GEORGETTE OTERO

de la firma Baker & McKenzie Colombia S.A. miembros de Baker & McKenzie, domiciliada en la Av. 82 No. 10-62 Piso 6 en Bogotá D.C., Colombia, poder especial amplio y suficiente, para adelantar ante las autoridades competentes los siguientes asuntos:

Registro y renovación de marcas; depósito y registro de nombres comerciales y enseñanzas comerciales; registro y renovación de nombres de dominio; obtención de patentes de invención, modelos de utilidad, diseños industriales, esquemas de trazado de circuitos integrados y variedades vegetales; inscripción de fusiones, cambios de nombre, traspasos, cambios de domicilio y de dirección; inscripción de contratos de prenda, inscripción de licencias de uso, embargos y otros actos que afecten o limiten los mencionados derechos de propiedad industrial; presentación y tramitación de todos los asuntos relacionados con nombres de dominio; presentación y tramitación de oposiciones, acciones de cancelación, acciones de nulidad, acciones de nulidad y restablecimiento del derecho, medidas cautelares, acciones por infracción de derechos, y demás acciones judiciales, relacionadas con procesos de propiedad industrial, competencia desleal y protección del consumidor,

a cuyo efecto los faculta para presentar solicitudes, peticiones y demandas, interponer recursos, reclamos, pagar impuestos y tasas, recibir, transigir y desistir, incluyendo facultades para entrar en acuerdos amigables, conciliaciones y transacciones, abandonar solicitudes y renunciar a derechos de propiedad industrial, y en general, a dar ante las autoridades competentes todos los pasos necesarios para la debida protección de nuestra propiedad industrial. Asimismo se concede a los expresados mandatarios poder suficiente para recibir notificaciones a nombre de la compañía, otorgar poderes a quienes vayan a representar a la compañía judicial y extrajudicialmente, responder en juicio a todas las reclamaciones o demandas que por motivo de los derechos de propiedad industrial se presentaren y hacer cuanto fuere necesario ante las autoridades administrativas o judiciales de cualquier orden para el cumplimiento de este mandato, dándoles igualmente facultad para sustituir el presente poder y en caso necesario, revocar dichas sustituciones.

Dado y firmado en

A los

días del mes de

de 20

Harrison F. Dillon
President & Chief Technology Officer
Solazyme, Inc.

Firma de la persona que firma el Poder



316

Apostille

(Convention de La Haye du 5 Octobre 1961)

1. Country: United States of America

This public document:

2. has been signed by Jeffrey W. Bullock

3. acting in the capacity of Secretary of State of Delaware

4. bears the seal/stamp of Office of Secretary of State

Certified

5. at Dover, Delaware

6. the seventh day of January, A.D. 2011

7. by Secretary of State, Delaware Department of State

8. No. 0431611

9. Seal/Stamp:

10. Signature:



msar
Secretary of State



Delaware

PAGE 1

The First State

I, JEFFREY W. BULLOCK, SECRETARY OF STATE OF THE STATE OF DELAWARE, DO HEREBY CERTIFY "SOLAZYME, INC." IS DULY INCORPORATED UNDER THE LAWS OF THE STATE OF DELAWARE AND IS IN GOOD STANDING AND HAS A LEGAL CORPORATE EXISTENCE SO FAR AS THE RECORDS OF THIS OFFICE SHOW, AS OF THE SEVENTH DAY OF JANUARY, A.D. 2011.

AND I DO HEREBY FURTHER CERTIFY THAT THE SAID "SOLAZYME, INC." WAS INCORPORATED ON THE THIRTY-FIRST DAY OF MARCH, A.D. 2003.

AND I DO HEREBY FURTHER CERTIFY THAT THE FRANCHISE TAXES HAVE BEEN PAID TO DATE.

AND I DO HEREBY FURTHER CERTIFY THAT THE ANNUAL REPORTS HAVE BEEN FILED TO DATE.



3635550 8300

110025104

You may verify this certificate online at corp.delaware.gov/authver

Jeffrey W. Bullock, Secretary of State
AUTHENTICATION: 8481140

DATE: 01-07-11



318

Secretario de estado

La apostilla o certificado de legalización solo confirma que la firma, sello o estampilla en el documento es auténtica. No significa que el contenido del documento sea correcto. La Apostilla no es valida dentro de los Estados Unidos, sus territorios o posesiones. Puede verificarla en www.sos.ca.gov/business/notary/apostille-search.

APOSTILLA

(Convención De La Haya del 5 octubre de 1961)

1. País Estados Unidos de América
2. Este documento Público ha sido firmado por J. Cohen
3. Quien ejerce funciones de notario público, Estado de California
4. Lleva el sello de J. Cohen notario público Estado de California.
- CERTIFICADO
5. En Sacramento, California.
6. El 21 de enero 2011.
7. Por el secretario de estado suplente- Estado de California
- 8 No. 914898
9. Sello en relieve del estado de California
10. Firma ilegible de Devra Bowen- Secretario de estado

RECONOCIMIENTO DE CALIFORNIA PARA TODO FIN

Estado de California

Condado de San Mateo

El 7 de enero de 2011 ante mí, J. Cohen notario público compareció personalmente Harrison F. Dillon quien demostró sobre una evidencia satisfactoria ser la persona cuyo nombre aparece en el instrumento anexo y reconoció haberlo suscrito en su condición autorizada. y que con su firma en el instrumento, la entidad a nombre de la cual firmó también suscribe el instrumento. Certifico bajo sanción de perjurio bajo las leyes del estado California que el párrafo anterior es verdadero y correcto.

En testimonio estampo mi firma y sello oficial.

Firmado. J. Cohen notario público. Sello de J. Cohen notario público de California.

Condado San Mateo. Comisión N. 1745073. Su comisión vence el 11 junio- 2011.

Marlen Neira
MARLEN NEIRA CASTRO
Traductor e Intérprete Oficial
Inglés - Español - Inglés
Resolución No. 2117 del 7 Oct 1987



Descripción del documento adjunto

Tipo de documento: Poder de abogado

Fecha : 7 enero de 2011. Número de páginas: 1 + certificado

Nombre de signatarios: ninguno

nombre: Dillon- cargo: Presidente. & jefe de tecnología

El signatario Solazyme, Inc.

PODER DE SOLAZYME, INC. .

Bilingüe inglés y español firmado por Harrison F. Dillon. Fechado el 7 enero 2011.

APOSTILLA

(Convención De La Haya del 5 octubre de 1961)

1. País Estados Unidos de América
2. Este documento Público ha sido firmado por Jeffrey W. Bullock
3. Quien ejerce funciones de secretario de estado de Delaware
4. Lleva el sello de secretario de Estado.

CERTIFICADO

5. En Dover Delaware.
6. El 7 de enero 2011.
7. Por el secretario de estado Departamento de Estado
- 8 No. 0431611
9. Sello en relieve del estado de California
10. Firma ilegible

Delaware
El primer Estado

Pag. 1

Yo, por Jeffrey W, Bullock, Secretario del estado de Delaware certifica por el presente que SOLAZYME, INC. Está debidamente constituida conforme a las leyes del Estado de Delaware y está al día, paz y salvo y tiene existencia legal corporativa tal como lo muestran los registros de esta oficina en fecha 7 de enero de 2011.

Certifico además que SOLAZYME, INC. fue constituida el 1 de marzo de 2003.

Certifico además que la compañía ha pagado los impuestos de franquicia hasta la fecha.

Certifico además que la Compañía ha presentado los informes anuales hasta la fecha.

3635550 8300

Jeffrey W, Bullock, Secretario del estado

110025104

Autenticación N. 8481140

Sello de Delaware verificar en corp.delaware.gov/authenticate.shtml Fecha_ 01-07-11

Es traducción conforme Bogota, enero 31- 2011



Marlen Neira C
MARLEN NEIRA CASTRO
Traductor e Intérprete Oficial
Inglés - Español - Inglés
Resolución No. 2117 del 7 Oct 1987



Industria y Comercio
SUPERINTENDENCIA

PLANILLA DE RADICACION - DIGITALIZACION DIARIA RELACION DE ANEXOS

Dependencia:	Fecha:	Recetido:
Tipo de Radicación: Entrada: Salida: Traslado:	Convenio:	
Radicador:		Planilla:

ID	No Radicación	Cone	Item
1			
2			
3			
4			
5			
6	11- 177590 -0		1
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			

Item	Unidad de Conservación
~~1~~	CD
2	Disquetes
3	OVD
4	Fotos
5	Libro
6	Muestra Física
7	Otros
8	Planos
9	Periodico
10	Publicación Periodica
11	Sobre blanco tamaño carta
12	Sobre blanco tamaño oficio
~~13~~	Sobre de manila

Observaciones
ANEXO C.D



No. 11-177590-00000-0000

Fecha: 2011-12-23 12:00:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 379 FASENACIONALII
Act. 411 PRESENTACION Folios: 563

LISTA DE CHEQUEO
ADMISIÓN A TRÁMITE – NUEVAS CREACIONES

PATENTE DE INVENCION ☐ MODELO DE UTILIDAD ☐ Art 33 Decisión 486/00

☐	Indicación que se solicita una patente.
☐	Datos de identificación del solicitante o de la persona que presenta la solicitud
☐	Descripción de la invención
☐	Dibujos de ser estos pertinentes
☐	Comprobante de pago de las tasas establecidas (De ser el caso formato de descuento)
Completa ☐	Incompleta ☐

PATENTE DE INVENCION PCT ☒ MODELO DE UTILIDAD PCT ☐ Art.33 Decisión 486/00, Circular Única

☒	Indicación que se solicita una PCT
☒	Copia de la solicitud en español, tal como fue presentada inicialmente (capítulo descriptivo, reivindicatorio, resumen)
☒	Dibujos de ser estos pertinentes
☒	Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)
Completa ☒	Incompleta ☐

DISEÑO INDUSTRIAL ☐ (Art. 119 Decisión 486/00)

☐	Indicación que se solicita Diseño industrial
☐	Datos de identificación del solicitante o de la persona que presenta la solicitud
☐	Representación gráfica y fotográfica del Diseño industrial o muestra del material que incorpora el diseño
☐	Comprobante de pago de las tasas establecidas
Completa ☐	Incompleta ☐

ESQUEMA DE TRAZADO ☐ (Art. 92 Decisión 486/00)

☐	Indicación que se solicita un esquema de trazado
☐	Datos de identificación del solicitante o de la persona que presenta la solicitud
☐	Representación gráfica de un esquema de trazado
☐	Comprobante de pago de las tasas establecidas
Completa ☐	Incompleta ☐