

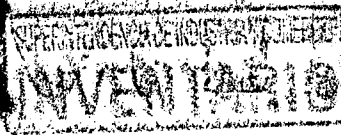
PCT
A
Industria y Comercio

SUPERINTENDENCIA

11-10649

SOLICITUD

PATENTE DE INVENCION



21. EXPEDIENTE No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-010649-00000-0000

Fecha: 2011-01-31 16:12:57 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYI Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 205

54. TÍTULO "MÉTODOS Y COMPOSICIONES PARA MEJORAR LA PRODUCCIÓN DE PRODUCTOS EN MICROORGANISMOS".

51. CLASIFICACIÓN INTERNACIONAL C12P 7/06

71. SOLICITANTE UNIVERSITY OF MASSACHUSETTS Y QTEROS, INC

DOMICILIO BOSTON, ESTADO DE
MASSACHUSETTS, ESTADOS UNIDOS DE AMÉRICA Y
MARLBOROUGH, ESTADO DE MASSACHUSETTS,
ESTADOS UNIDOS DE AMÉRICA.

74. REPRESENTANTE LEGAL ALICIA LLOREDA RICAURTE
T.P. 53.215

22. BOGOTÁ, D.C. 31 ENE 2011

(FORMA P 10)

28-02-2011
N

PETITORIO



SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-010649- -00000-0000

Fecha: 2011-01-31 16:12:57 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 205

FORMULARIO UNICO 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: **UNIVERSITY OF MASSACHUSETTS Y QTEROS, INC.**

Dirección: 225 Franklin Street, 12th Floor, Boston, Massachusetts 02110, Estados Unidos de América y 100 Campus Drive, Marlborough, Massachusetts 01752, Estados Unidos de América.

Nacionalidad o Domicilio: Boston, Estado de Massachusetts, Estados Unidos de América y Marlborough, Estado de Massachusetts, Estados Unidos de América.

Lugar de Constitución: Estado de Massachusetts, Estados Unidos de América.

Teléfono: 3264270 Fax: 6069701
E-mail: jllcco@lloredacamacho.com

IDENTIFICACION

C.C. ☐ NIT ☐
C.E. ☐ Otro ☐

Cual _____

Número _____

③

**REPRESENTANTE
O APODERADO**

Nombre: **ALICIA LLOREDA RICAURTE**

Dirección: Calle 72 No. 5-83 Piso 5º, Bogotá, Colombia

Teléfono: 3264270 Fax: 6069701
E-mail: jllcco@lloredacamacho.com

C.C. ☒ NIT ☐
C.E. ☐ Otro ☐

Cual _____

Número 39.690.713 TP 53.215

④

**INVENTOR(ES)
(72)**

Nombre: BLANCHARD, Jeffrey; LESCHINE, Susan; PETIT, Elsa; FABEL, John y SCHMALISCH, Matthias.

Dirección: 60 Amherst Road, Leverett, MA 01054; 60 Juggler Meadow Road, Leverett, MA 01054; 67 Crescent Street, No. 3, Northampton, MA 01060; 120 Pulpit Hill Road, No. 14, Amherst, MA 01002 y 25 Virginia Street, Unit 2, Somerville, MA 02145.

Nacionalidad o Domicilio: Estados Unidos de América.

⑤

PCT (86)

Solicitud Internacional No. PCT/US2009/051993

Fecha: 28/07/09

Publicación Internacional No. WO 2010/014632 A3

Fecha: 04/02/10

⑥

Título

"MÉTODOS Y COMPOSICIONES PARA MEJORAR LA PRODUCCIÓN DE PRODUCTOS EN MICROORGANISMOS"

⑦

Clasificación Internacional (51) C12P 7/06

⑧

Prioridad

SI ☒ NO ☐

(33) País de Origen

(32) Fecha

(31) Número de Solicitud

US

28/07/08

61/084,233

US

13/07/09

61/225,184

US

27/07/09

61/228,922

⑨

Comprobante de Pago No. 11-001941

Fecha: 12-01-11

10

ANEXOS

- ☒ Comprobante de pago de la tasa de presentación de la solicitud.
- ☒ Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica.
- ☒ Documento de Representación Legal. (University of Massachusetts)
- ☐ 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 reporte preliminar internacional sobre patentabilidad.
- ☐ Traducción del reporte preliminar internacional sobre patentabilidad.
- ☐ 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.
- ☐ Arte Final 12X12.
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionados parcial o totalmente con la invención de esta solicitud.
- ☐ De ser el caso, modificaciones efectuadas a la solicitud internacional.
- ☒ Otros: Reporte de Búsqueda Internacional, Publicación y Opinión escrita de la autoridad de búsqueda Internacional.

11

FIGURA CARACTERISTICA:

12

Solicito la concesión de la patente,

NOMBRE: ALICIA LLOREDA RICAURTE

FIRMA: 

C.C. 39.690.713

T.P. 53.215

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

NIT : 800.176.089-2

- / -



Industria y Comercio

SUPERINTENDENCIA

RECIBO DE CAJA

No. 11 - 001941

Bogotá D.C., 12 de 2011 - 07:38:42

RECIBIDO DE : LLOREDA & CIA. S.A.

NI 830.022.145

*** Soporte del Pago ***

TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	142967230	11/01/2011	31.003.700.00

*** Conceptos Pagados ***

CANT. RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1 50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE INVENCIÓN	571.000.00	571.000.00
			<u>\$571.000.00</u>

SON: **QUINIENTOS SETENTA Y UN MIL PESOS MONEDA CORRIENTE**

Responsable: _____

Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-010649- -00000-0000

Fecha: 2011-01-31 16:12:57 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 205

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

, Enero 12 de 2011 - 07:39:04

ANEXO

El documento que acredita la calidad de REPRESENTANTES LEGALES en Bogotá de los doctores ALICIA LLOREDADA y/o GUSTAVO TAMAYO y/o IGNACIO SANTAMARIA y/o ENRIQUE ALVAREZ para TODOS LOS ASUNTOS RELACIONADOS CON LA PROPIEDAD INDUSTRIAL de la sociedad peticionaria, que faculta a cada uno de dichos abogados para actuar individualmente, se está presentando con esta solicitud. En consecuencia, ruego a usted reconocer la personería de cada uno de dichos Representantes Legales para actuar dentro del presente expediente, a fin de que cualquiera de ellos pueda intervenir en cualquier tiempo respecto de la solicitud en él contenida.


ALICIA LLOREDADA RICAURTE

T.P. No. 53.215

TRADUCCIÓN OFICIAL
DE UN DOCUMENTO ESCRITO EN INGLÉS

PODER

Legalización de un poder otorgado por la **UNIVERSIDAD DE MASSACHUSETTS**, documento escrito en inglés y español.

(firma: ilegible, Nicholas DeCristofaro, Director de Proyectos Comerciales y Propiedad Industrial de la Universidad de Massachusetts)

Certificado Notarial, Legalización y Apostilla de la SECRETARÍA DE ESTADO DEL ESTADO DE MARYLAND, documentos también escritos en inglés.

EJECUCIÓN

Certificado Verificado y Autenticado
en el Condado de Anne Arundel, Maryland, EEUU,
el día 28 de octubre de 2008
(firma: Ahmad Miski)
Notario Público

Esta traducción se realizó a solicitud del interesado a los 5 días del mes de noviembre de 2008.

J. A. Forster

Gerhard Alexander Forster Moreno
C.C. No. 79.392.484 de Bogotá
Traductor e Intérprete Oficial
Alemán e Inglés

**GERHARD ALEXANDER
FORSTER MORENO**
Traductor e Intérprete Oficial
Resolución 0622 - Minjusticia 1995

-- 0 --

El Estado de Maryland

Oficina de la Secretaría de Estado

Apostilla

(Convención de la Haya del 5 de octubre de 1961)

1. País: Estados Unidos de América

Este documento público



J. A. Forster
**GERHARD ALEXANDER
FORSTER MORENO**
Traductor e Intérprete Oficial
Resolución 0622 - Minjusticia 1995

2. ha sido firmado por **Robert P. Duckworth**
3. actuando en su condición de **Secretario de la Corte del Circuito para el Condado de Anne Arundel**
4. está provisto con el sello/estampilla de la **Corte del Circuito para el Condado de Anne Arundel**

Certificado

5. en Annapolis, Maryland
6. el **28° día de octubre de 2008**
7. por La Secretaria de Estado de Maryland

8. Sello

Sello redondo oficial en altorrelieve: SELLO DE LA OFICINA DE LA SECRETARÍA DE ESTADO DE MARYLAND (con escudo)

9. Firma

(fdo.) *ilegible*

Secretario de Estado

No. 226811



G. A. Forster
GERHARD ALEXANDER FORSTER MORENO
Traductor e Interpretre Oficial
Resolución 0622 - Minjusticia 1995

Escudo del Condado de Anne Arundel – Maryland

Estado de Maryland, Condado de Anne Arundel, Sct.

Yo, **ROBERT P. DUCKWORTH**, Secretario de la Corte del Circuito para el Condado de Anne Arundel, Maryland, una corte de registro, certifico por este medio que **Ahmad Miski** era una Notario Público en servicio/nombrado y calificado comenzando el 5º día de mayo de 2005.

En Testimonio de Lo Cual, yo he firmado e impreso el sello de la corte este 28º día de octubre de 2008.

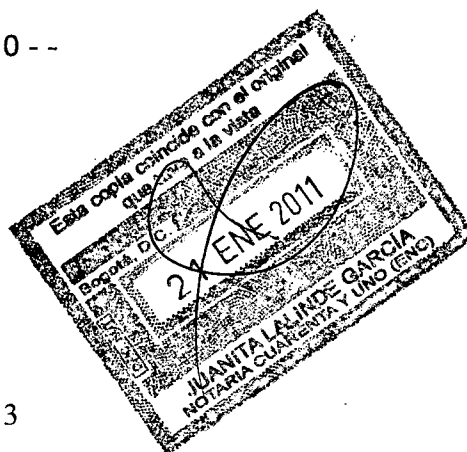
(fdo.) *Robert P. Duckworth*

Robert P. Duckworth, Secretario

Corte del Circuito para el Condado de Anne Arundel, Maryland

Con aposición del sello redondo oficial en altorrelieve con escudo: CORTE DEL CIRCUITO PARA EL CONDADO DE ANNE ARUNDEL – MD

-- 0 --



3

GERHARD ALEXANDER
FORSTER MORENO
Traductor e Interpretre Oficial
Resolución 0622 - Minjusticia 1995

PARA CONOCIMIENTO DE TODOS AQUELLOS QUE VEAN EL PRESENTE

Basado en la información disponible

Legalización de los EEUU

Autentica la firma de

Nicholas DeCristofaro

notarizada por

Lyne D. Laliberte

es genuina y correcta

(fdo.) *Deborah Campbell*

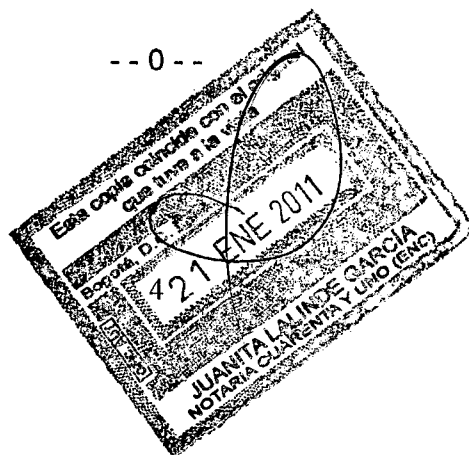
Deborah Campbell

Firma Autorizada

Este día 28 de octubre de 2008, yo Ahmad Miski, Notario Público del Condado de Anne Arundel, Estado de Maryland, certifico por este medio bajo el Juramento de mi Cargo y Sello que la firma de Deborah Campbell es genuina y fiel.

(fdo.) *ilegible*

Mi nombramiento vence el 1º de mayo de 2009



GERHARD ALEXANDER
FORSTER MORENO
Traductor e Interprete Oficial
Resolución 0622 - Minjusticia 1995

CERTIFICADO NOTARIAL

(fdo.) *ilegible*

Notaria Pública

Mi nombramiento vence:

Sello rectangular oficial con escudo:

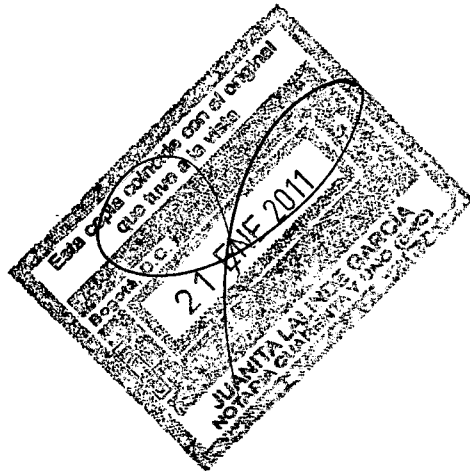
LYNE D. LALIBERTE

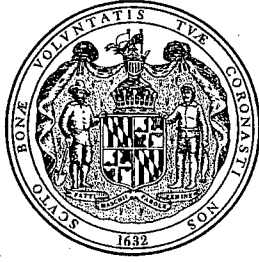
Notaria Pública

Mancomunidad de Massachusetts

Mi Nombramiento Vence el 14 de diciembre de 2012

-- 0 --





The State of Maryland

Office of the Secretary of State

Apostille

(Convention de La Haye du 5 Octobre 1961)

1. Country: United States of America
This public document
2. has been signed by **Robert P. Duckworth**
3. acting in the capacity of **Clerk of the Circuit Court for Anne Arundel County**
4. bears the seal/stamp of the **Circuit Court for Anne Arundel County**

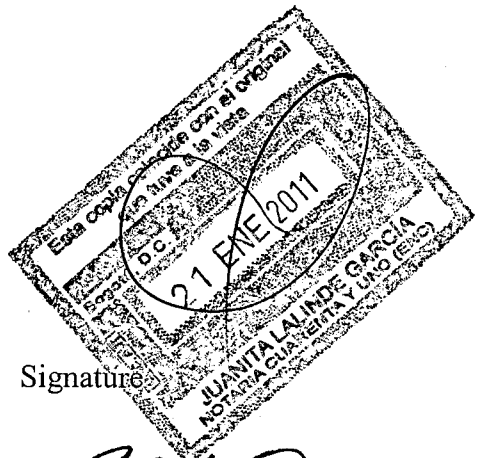
Certified

5. at Annapolis, Maryland
6. the **28th day of October, 2008**
7. by The Secretary of State of Maryland
8. Seal



9. Signature

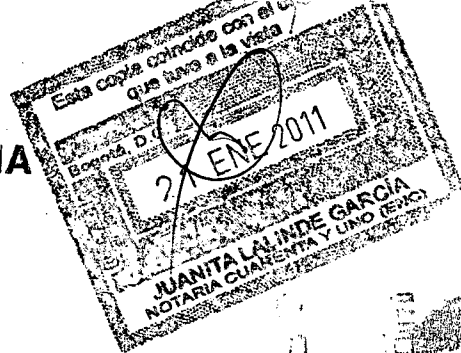
Secretary of State



No. 226811



CONSULADO DE COLOMBIA
WASHINGTON, D.C.



EL CONSULADO DE COLOMBIA

CERTIFICA:

Que es auténtica la firma de Robert P Duckworth
quien el día 28 Octubre 2008
desempeñaba las funciones de Juez del Circuito Maryland
la cual aparece al pie de los documentos anexos relacionados con Documento
Comercial

OCT 30 2008

Dado en Washington, D.C. a los _____

Pagado Impuesto de Timbre 10
Pagado Derecho Consular 15
Total 25

(El Consulado no asume ninguna
responsabilidad por el contenido de los
documentos anexos)

CARLOS ANDRÉS HURTADO VÉLEZ
VICECONSUL



MINISTERIO DE
RELACIONES
EXTERIORES

127814

Calderón

CUYA FIRMA AUTORIZA EN EL
DOCUMENTO DESEMPEÑA
LAS FUNCIONES RELACIONADAS

NO ASUME
RESPONSABILIDAD
DEL TEXTO

2000 NOV -6 P 12:32

WJL
JEFE DE LEGALIZACIÓN
, BOGOTÁ D.C.

To All To Whom These Presents Shall Come
Based upon available information

US Legalization

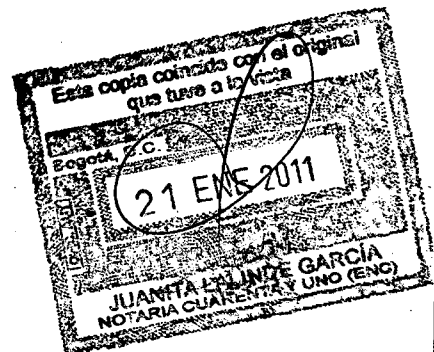
Authenticates the signature of

Nicholas DeCristofaro
notarized by

Lyne D. Laliberte
is true and correct

Deborah Campbell

Deborah Campbell
Authorized Signature



On this day of October 28th, 2008, I Ahmad Miski, Notary Public, Anne Arundel County, State of Maryland, do hereby certify under the Oath of my Office and Seal that the signature of Deborah Campbell is true and accurate.

Ahmad Miski

My commission expires on May 1, 2009

73566

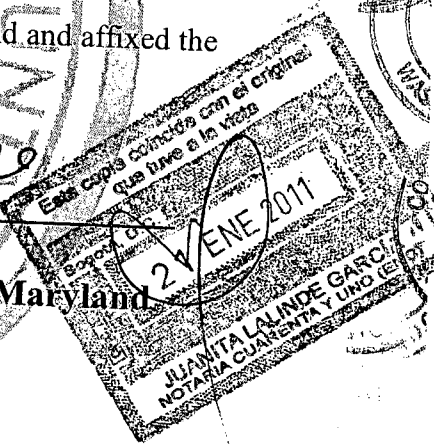
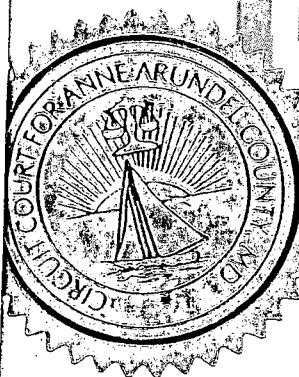
State of Maryland, Anne Arundel County, Sect.

I, ROBERT P. DUCKWORTH, Clerk of the Circuit Court for Anne Arundel County, Maryland, a court of record, do hereby certify that Ahmad Miski was a commissioned/appointed and qualified Notary Public commencing on the 5th day of May, 2005.

In Testimony Whereof, I have hereunto set my hand and affixed the seal of the court this 28th day of October, 2008.

Robert P. Duckworth

Robert P. Duckworth, Clerk
Circuit Court for Anne Arundel County, Maryland



Los abajo firmados, Nicholas DeCristofaro, Director de Empresas Comerciales y Propiedad Intelectual de la Universidad de Massachusetts, domiciliado en 70 Butterfield Terrace, Amherst MA 01003 Actuando en representación de Universidad de Massachusetts, una sociedad organizada y existente bajo las leyes de Massachusetts, con domicilio en 225 Franklin Street, 12th Floor, Boston MA 02110

nombramos por el presente a ALICIA LLOREDA y/o GUSTAVO TAMAYO y/o IGNACIO SANTAMARIA y/o ENRIQUE ALVAREZ, domiciliados en Bogotá, Colombia, con oficinas en dicha ciudad, Calle 72 No. 5-83, piso 6o., nuestros representantes legales y verdaderos en Bogotá, con facultades para recibir notificaciones y nombrar apoderados judiciales o extrajudiciales, con el fin de dar cumplimiento al parágrafo 1o. del artículo 543 del decreto 410 de 1971.

Los doctores ALICIA LLOREDA y/o GUSTAVO TAMAYO y/o IGNACIO SANTAMARIA y/o ENRIQUE ALVAREZ, o los apoderados que cualquiera de ellos designe, quedan especialmente facultados para representarnos ante las autoridades colombianas, administrativas, judiciales o de policía, en todos los trámites relativos a la propiedad industrial, a cuyo efecto les facultamos para dar ante dichas autoridades todos los pasos necesarios al objeto indicado; elevar solicitudes, declaraciones, reclamos, recursos y presentar desistimientos; formular descripciones, enmiendas, limitar el alcance de las solicitudes, presentar oposiciones y apelaciones; instaurar cancelaciones, renunciar total o parcialmente a los registros, abonar todos los impuestos y cuotas y efectuar todo otro pago determinado por la ley; recibir todos los documentos y valores, dando el descargo respectivo; llenar cualesquiera otros requisitos y tomar, en fin, todas las medidas que creyeren conducentes al resguardo de nuestros intereses; y en caso de producirse oposición, pasando los antecedentes a los tribunales, quedan facultados para tomar intervención como demandantes o demandados ante los jueces que sean competentes, pudiendo transar, someter a árbitros, desistir, percibir y apelar, con todas las demás facultades que resulten necesarias, y por el presente declaramos desde ahora válido y bueno cuanto dichos representantes y/o los apoderados que cualquiera de ellos nombre hicieren en nuestro beneficio.

Asimismo, facultamos a los doctores ALICIA LLOREDA y/o GUSTAVO TAMAYO y/o IGNACIO SANTAMARIA y/o ENRIQUE ALVAREZ, para solicitar en la Oficina de Propiedad Industrial la protocolización de este documento y de cualquier poder que cualquiera de ellos otorgue a nombre nuestro.

Dado y firmado



The undersigned, Nicholas DeCristofaro, Director of Commercial Ventures and Intellectual Property at the University of Massachusetts domiciled at 70 Butterfield Terrace, Amherst, MA 01003

Acting on behalf of **UNIVERSITY OF MASSACHUSETTS**, a corporation organized and existing under the laws of Massachusetts, domiciled at 225 Franklin Street, 12th Floor, Boston MA 02110 hereby appoint, ALICIA LLOREDA and/or GUSTAVO TAMAYO and/or IGNACIO SANTAMARIA and/or ENRIQUE ALVAREZ, of Bogota, Colombia, with offices in said town, Calle 72 No. 5-83, Sixth Floor, our true and legal representatives in Bogotá, with power to receive notifications and to appoint judicial or extrajudicial agents, so as to comply with paragraph 1o. of article 543 of decree 410 of 1971.

Doctors ALICIA LLOREDA and/or GUSTAVO TAMAYO and/or IGNACIO SANTAMARIA and/or ENRIQUE ALVAREZ, or the agents by any one of them appointed, are especially empowered to represent us before the administrative, judicial or police authorities of Colombia, in all proceedings pertaining to industrial property, for which purpose we authorize them to take before the said authorities all the necessary steps for the said end; to file petitions, declarations, claims, recourses and withdrawals, to draw up specifications; to file amendments, to limit the scope of applications, file oppositions; file cancellations total or partially waive registrations, to pay all installments and effect all other payments prescribed by law; to receive proper receipts, to comply with all other requirements and, in short, to take all those measures which they may deem fit for the protection of our interests; and in case of opposition, where the proceedings be passed over to the courts, we authorize said agents to act on our behalf, as complainants as well as defendants, before the competent judges, with full powers to make arrangements, to submit to arbitrators, to desist, to collect, to appeal and, in short, with all such further powers as may be required, and we hereby declare from now on as valid and good all that said representatives and/or the agents by any one of them appointed may do for our benefit.

We also authorize doctors ALICIA LLOREDA and/or GUSTAVO TAMAYO and/or IGNACIO SANTAMARIA and/or ENRIQUE ALVAREZ, to apply for protocolization of this document and of any power of attorney by any one of them conferred in our name, before the Office of Industrial Property.

Given and signed this 3rd day of October, 2008.

Signature

And Stamp:

Nicholas DeCristofaro

Date: October 3, 2008

NOTA RY SHOULD EXECUTE OTHER SIDE

CERTIFICADO NOTARIAL (SOCIEDAD)

Estado de _____ {ss
 Condado de _____
 Hoy de _____ de 200 , compareció _____ a quien conozco y
 de quien me consta que firmó el anterior documento, y
 quien debidamente juramentado declaró y dijo que él es _____
 de _____ la compañía nombrada en el
 documento que antecede, una sociedad debidamente
 organizada y existente conforme a las leyes del Estado de _____
 , domiciliada en la ciudad de _____
 , Condado de _____
 Estado de _____
 legalmente constituida para un período de duración que
 aún no ha terminado, que ejerce su objeto social; conforme
 a las leyes del Estado de _____
 ; que el
 declarante está autorizado por dicha Compañía para
 conferir este poder; que el sello estampado por él al pie de
 tal poder, en nombre y en representación de la expresada
 Compañía, es el sello oficial de ella; y que el objeto para el
 cual se confiere este instrumento está comprendido dentro
 del de la sociedad.
 El suscrito Notario ha visto los Estatutos de incorporación
 firmados en _____
 Estado de _____ , con fecha _____ y
 certifica que de conformidad con la legislación de dicho
 Estado la anterior declaración juramentada es prueba
 suficiente de que los hechos sobre que ella versa son
 ciertos o verdaderos, de lo cual da fe.

Notary Public (Notario Público).

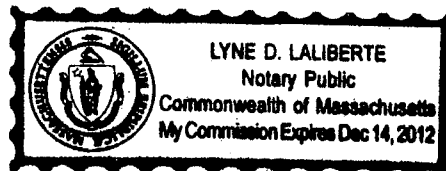


NOTARIAL CERTIFICATE (CORPORATION)

State of Massachusetts {ss
 County of Hampshire
 On this 3rd day of October, 2008 , personally came
 Nicholas DeCristofaro
 to me known, and known to me to be the person who
 signed the foregoing instrument, who being by me duly
 sworn did depose and say that he is the Director of the
 Office of Commercial Ventures and Intellectual Property at
 the University of Massachusetts, Amherst, MA the
 company named within, a Corporation organized and
 existing under the laws of the State of Massachusetts,
 domiciled at the City of Boston, State of Massachusetts,
 legally constituted for a period of tenure not yet expired,
 which exercises its corporate purposes according to the
 laws of the State of Massachusetts, USA ; that the
 deponent is authorized by said Company to confer this
 power of attorney; that the seal affixed hereunto by the
 deponent in the name and on behalf of said Corporation is
 the corporate seal thereof; and that the purpose for which
 this instrument is granted is included within the corporate
 purposes of the Company.
 The undersigned Notary Public has seen the articles of
 Incorporation of said Corporation executed at Boston, State
 of Massachusetts, under date, and hereby certifies that
 under the laws of said State the foregoing acknowledgment
 is sufficient proof that the facts therein contained are true,
 which she attests.

Lyne D. Laliberte
 Notary Public

My commission expires: 12/14/2012



MBIA

MBIA

 Industria y Comercio SUPERINTENDENCIA		SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO DIVISIÓN DE NUEVAS CREACIONES EXTRACTO PARA PUBLICACIÓN SOLICITUD DE PATENTE DE INVENCION PCT	
(21) No. de solicitud:		(51) Int Cl: C12P 7/06	
(22) Fecha de solicitud: 31-01-11		(71) Solicitante(s): UNIVERSITY OF MASSACHUSETTS Y QTEROS, INC.	
(30) Prioridad	(31) No. Prioridad	(32) Fecha	(33) País
	61/084,233	28-07-08	US
	61/225,184	13-07-09	US
	61/228,922	27-07-09	US
		(72) Inventor(es): BLANCHARD, Jeffrey; LESCHINE, Susan; PETIT, Elsa; FABEL, John y SCHMALISCH, Matthias.	
		(74) Apoderado: ALICIA LLOREDA RICAURTE	
(85)	Fecha límite inicio fase nacional: 28-02-11		(87) Publicación internacional
(86)	Datos relativos a la presentación de la solicitud PCT		Fecha: 04-02-10
	Fecha de presentación de la solicitud: 28-07-09		No. publicación: WO 2010/014632 A3
	No. de solicitud: PCT/US2009/051993		
(54) Título: "MÉTODOS Y COMPOSICIONES PARA MEJORAR LA PRODUCCIÓN DE PRODUCTOS EN MICROORGANISMOS"			

Reivindicación número 1-

Un producto para la producción de un biocombustible que comprende: una biomasa lignocelulósica, un microorganismo que es capaz de hidrolizar directamente y fermentar dicha biomasa, en donde dicho microorganismo está modificado para proveer actividad mejorada de una o más celulasas.



Industria y Comercio

SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

DIVISIÓN DE NUEVAS CREACIONES

TARJETA ARCHIVO PROPIETARIOS

PATENTE DE INVENCION ____ PCT X MODELO DE UTILIDAD ____ DISEÑO INDUSTRIAL ____

(21) No. de solicitud:

(22) Fecha de solicitud: 31-01-2011

(51) Clasificación Internacional: C12P 7/06

(71) Solicitante(s): **UNIVERSITY OF MASSACHUSETTS Y QTEROS, INC.**

(72) Inventor(es): BLANCHARD, Jeffrey; LESCHINE, Susan; PETIT, Elsa; FABEL, John y SCHMALISCH, Matthias.

(74) Apoderado: **ALICIA LLOREDA RICAURTE**

(30) Prioridad	(31) No. Prioridad	(32) Fecha	(33) País
	61/084,233	28-07-08	US
	61/225,184	13-07-09	US
	61/228,922	27-07-09	US

(85) Fecha límite inicio fase nacional: 28-02-11

(86) Datos relativos a la presentación de la solicitud PCT

Fecha de presentación de la solicitud: 28-07-09

No. de solicitud: **PCT/US2009/051993**

(87) Publicación internacional

Fecha: **04-02-10**

No. publicación: **WO 2010/014632 A3**

(54) Título: **MÉTODOS Y COMPOSICIONES PARA MEJORAR LA PRODUCCIÓN DE PRODUCTOS EN MICROORGANISMOS.**



Industria y Comercio

SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

DIVISIÓN DE NUEVAS CREACIONES

TARJETA ARCHIVO TEMÁTICO

PATENTE DE INVENCION ____ PCT X MODELO DE UTILIDAD ____ DISEÑO INDUSTRIAL ____

(21) No. de solicitud:

(22) Fecha de solicitud: 31-01-2011

(51) Clasificación Internacional: C12P 7/06

(71) Solicitante(s): **UNIVERSITY OF MASSACHUSETTS Y QTEROS, INC.**

(72) Inventor(es): BLANCHARD, Jeffrey; LESCHINE, Susan; PETIT, Elsa; FABEL, John y SCHMALISCH, Matthias.

(74) Apoderado: **ALICIA LLOREDA RICAURTE**

(30) Prioridad	(31) No. Prioridad	(32) Fecha	(33) Pais
	61/084,233	28-07-08	US
	61/225,184	13-07-09	US
	61/228,922	27-07-09	US

FIGURA CARACTERISTICA
6 X 6 cm

(85) Fecha límite inicio fase nacional: 28-02-11

(86) Datos relativos a la presentación de la solíc. PCT

Fecha de presentación de la solicitud: 28-07-09

No. de solicitud: **PCT/US2009/051993**

(87) Publicación internacional:

Fecha: **04-02-10**

No. Publicación: **WO 2010/014632 A3**

(54) Título: **MÉTODOS Y COMPOSICIONES PARA MEJORAR LA PRODUCCIÓN DE PRODUCTOS EN MICROORGANISMOS.**

(57) Resumen:

Reivindicación número 1:

Un producto para la producción de un biocombustible que comprende: una biomasa lignocelulósica, un microorganismo que es capaz de hidrolizar directamente y fermentar dicha biomasa, en donde dicho microorganismo está modificado para proveer actividad mejorada de una o más celulasas.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 078800150WO1	FOR FURTHER ACTION see Form PCT/ISA/220 as well as, where applicable, item 5 below.	
International application No. PCT/US2009/051993	International filing date (<i>day/month/year</i>) 28 JULY 2009 (28.07.2009)	(Earliest) Priority Date (<i>day/month/year</i>) 28 JULY 2008 (28.07.2008)
Applicant UNIVERSITY OF MASSACHUSETTS et al		

This International search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 5 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

a. With regard to the **language**, the international search was carried out on the basis of :

☒ the international application in the language in which it was filed

☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6bis(a)).

c. ☒ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☒ **Certain claims were found unsearchable** (See Box No. II)

3. ☐ **Unity of invention is lacking** (See Box No. III)

4. With regard to the **title**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

a. the figure of the **drawings** to be published with the abstract is Figure No. 19

☒ as suggested by the applicant.

☐ as selected by this Authority, because the applicant failed to suggest a figure.

☐ as selected by this Authority, because this figure better characterizes the invention.

b. ☐ none of the figure is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2009/051993

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of:

a. a sequence listing filed or furnished

- ☒ on paper
☐ in electronic form

b. time of filing or furnishing

- ☒ contained in the international application as filed
☐ filed together with the international application in electronic form
☐ furnished subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

The applicant failed to comply with PCT/ISA/225 for invitation to furnish to ISA a nucleotide sequence listing in electronic form.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2009/051993

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 13,14,16,20,21
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claims 13, 14, 16, 20, 21 delimit genomic sequences for cellulases or nucleotide sequence itself. However, the exact sequence for the cellulases is unclear from the expression written in claims 13, 14, 16, 20 and there is no electronic form for claim 21. Therefore a meaningful opinion on claims 13, 14, 16, 20, 21 was not made.
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

A. CLASSIFICATION OF SUBJECT MATTER <i>C12P 7/06(2006.01)i, C12N 1/21(2006.01)i, C12N 15/52(2006.01)i, C12R 1/145(2006.01)n</i> According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12P 7/06; C12N 1/15; C12N 1/20; C12N 1/21; C12N 15/00; C12N 5/00 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models (Chinese Patents and application for patent) Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS(KIPO internal) & Keywords: biofuel, lignocellulose, biomass, cellulase, Clostridium phytofermentans, heterologous polynucleotide, ethanol, lactic acid, acetic acid, formic acid, oligomeric carbohydrate, Cphy, cellobiose		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007-0178569 A1 (LESCHINE, S. & WARNICK, T. A.) 02 August 2007 See abstract: paragraphs [0007],[0010],[0023]-[0028],[0050],[0051],[0056],[0069]	1-12, 15, 17-19
A	US 5496725 A1 (YU, I. K.) 05 March 1996 See abstract: claims 1-5	1-12, 15, 17-19
A	Warnick, T. A. et al., Clostridium phytofermentans sp. nov., a cellulolytic mesophile from forest soil, International Journal of Systematic and Evolutionary Microbiology, 2002, Vol.52, pp.1155-1160 See abstract; figure 2; table 1	1-12, 15, 17-19
A	WO 92-16615 A1 (UNIVERSITY OF FLORIDA & BIOENERGY INTERNATIONAL, L.C.) 01 October 1992 See abstract: claims 1, 31-35, 42, 45, 48, 51, 52, 55-60, 64, 65	1-12, 15, 17-19
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 27 APRIL 2010 (27.04.2010)		Date of mailing of the international search report 30 APRIL 2010 (30.04.2010)
Name and mailing address of the ISA/KR  Korean Intellectual Property Office Government Complex-Daejeon, 139 Seonsa-ro, Seo-gu, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140		Authorized officer JEONG, JAE CHEOL  Telephone No. 82-42-481-8385

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2009/051993

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007-0178569 A1	02.08.2007	CA 2640429 A1 EP 1984514 A2 KR 10-2008-0091257 A US 2009-0068714 A1 WO 2007-089677 A2	09.08.2007 29.10.2008 09.10.2008 12.03.2009 09.08.2007
US 5496725 A1	05.03.1996	US 5759845 A1 WO 95-04815 A1	02.06.1998 16.02.1995
WO 92-16615 A1	01.10.1992	CA 2106377 A1 EP 0431047 A1 EP 0560885 A1 EP 0560885 B1 EP 0576621 A1 EP 0576621 A4 EP 0576621 B1 JP 03-457664 B2 JP 03-593125 B2 JP 2004-283176 A JP 3457664 B2 KR 10-0292079 B1 KR 10-1990-0702042 A KR 10-1993-0703441 A US 2003-0008363 A1 US 2005-0158836 A1 US 2007-0172937 A1 US 5028539 A1 US 5424202 A1 US 5482846 A1 US 5487989 A1 US 5554520 A1 US 5821093 A1 US 5916787 A1 US 6107093 A1 US 6849434 B2 US 7192772 B1 WO 90-02193 A1 WO 92-10561 A1 WO 94-06924 A1 WO 95-27064 A1	19.09.1992 12.06.1991 04.06.1997 17.03.1999 05.01.1994 02.11.1994 28.02.2001 01.08.2003 24.11.2004 14.10.2004 20.10.2003 30.03.2006 05.12.1990 30.11.1993 09.01.2003 21.07.2005 26.07.2007 02.07.1991 13.06.1995 09.01.1996 30.01.1996 10.09.1996 13.10.1998 29.06.1999 22.08.2000 01.02.2005 20.03.2007 08.03.1990 25.06.1992 31.03.1994 12.10.1995

(43) International Publication Date
4 February 2010 (04.02.2010)



(10) International Publication Number
WO 2010/014632 A3

- (51) International Patent Classification:
C12P 7/06 (2006.01) *C12N 15/52* (2006.01)
C12N 1/21 (2006.01) *C12R 1/145* (2006.01)
- (21) International Application Number:
PCT/US2009/051993
- (22) International Filing Date:
28 July 2009 (28.07.2009)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/084,233 28 July 2008 (28.07.2008) US
61/225,184 13 July 2009 (13.07.2009) US
61/228,922 27 July 2009 (27.07.2009) US
- (71) Applicants (for all designated States except US): **UNIVERSITY OF MASSACHUSETTS** [US/US]; 225 Franklin Street, 12th Floor, Boston, MA 02110 (US). **QTEROS, INC.** [US/US]; 100 Campus Drive, Marborough, MA 01752 (US).
- (72) Inventors; and
(75) Inventors/Applicants (for US only): **BLANCHARD, Jeffrey** [US/US]; 60 Amherst Road, Leverett, MA 01054 (US). **LESCHINE, Susan** [US/US]; 60 Juggler Meadow Road, Leverett, MA 01054 (US). **PETIT, Elsa** [FR/US]; 67 Crescent Street, No. 3, Northampton, MA 01060 (US). **FABEL, John** [US/US]; 120 Pulpit Hill Road, No. 14, Amherst, MA 01002 (US). **SCHMALISCH, Matthias** [DE/US]; 25 Virginia Street, Unit 2, Somerville, MA 02145 (US).
- (74) Agents: **FASSE, J., Peter** et al.; Fish & Richardson P.C., P.O. Box 1022, Minneapolis, MN 55440-1022 (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,

[Continued on next page]

(54) Title: METHODS AND COMPOSITIONS FOR IMPROVING THE PRODUCTION OF PRODUCTS IN MICROORGANISMS

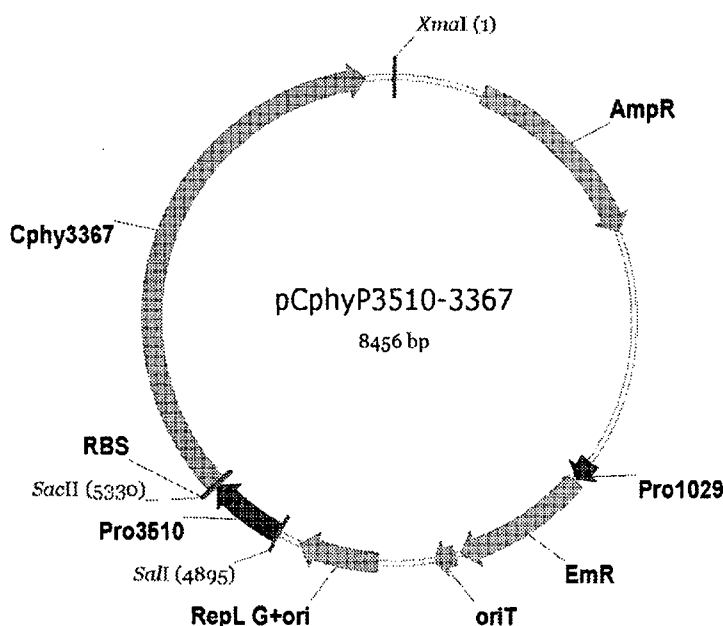


FIG. 19

(57) Abstract: Methods and compositions are provided for improving the production of products, such as fuel products like ethanol, in microorganisms. In particular, methods and compositions are described for improving ethanol production utilizing genes identified in *Clostridium phytofermentans*.



SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

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))

(88) **Date of publication of the international search report:**

8 July 2010

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
FASSE J. PETER

FISH & RICHARDSON P.C. P.O. BOX 1022
MINNEAPOLIS MN 55440-1022 USA

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43bis.1)

Date of mailing
(day/month/year) 30 APRIL 2010 (30.04.2010)

Applicant's or agent's file reference
078800150WO1

FOR FURTHER ACTION
See paragraph 2 below

International application No. PCT/US2009/051993	International filing date (day/month/year) 28 JULY 2009 (28.07.2009)	Priority date(day/month/year) 28 JULY 2008 (28.07.2008)
---	--	--

International Patent Classification (IPC) or both national classification and IPC

C12P 7/06(2006.01)i, C12N 1/21(2006.01)i, C12N 15/52(2006.01)i, C12R 1/145(2006.01)n

Applicant
UNIVERSITY OF MASSACHUSETTS et al

1. This opinion contains indications relating to the following items:
- ☒ Box No. I Basis of the opinion
 - ☐ Box No. II Priority
 - ☒ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
 - ☐ Box No. IV Lack of unity of invention
 - ☒ Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
 - ☐ Box No. VI Certain documents cited
 - ☐ Box No. VII Certain defects in the international application
 - ☐ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**
If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1bis(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.
For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/KR
Korean Intellectual Property Office
Government Complex-Daejeon, 139
Seonsa-ro, Seo-gu, Daejeon 302
-701, Republic of Korea
Facsimile No. 82-42-472-7140



Date of completion of this opinion
27 APRIL 2010 (27.04.2010)

Authorized officer
JEONG, JAE CHEOL
Telephone No.82-42-481-8385



WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.

PCT/US2009/051993

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of :

- ☒ the international application in the language in which it was filed
- ☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43bis.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

a. a sequence listing filed or furnished

- ☒ on paper
- ☐ in electronic form

b. time of filing or furnishing

- ☒ contained in the international application as filed.
- ☐ filed together with the international application in electronic form.
- ☐ furnished subsequently to this Authority for the purposes of search.

4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

The applicant failed to comply with PCT/ISA/225 for invitation to furnish to ISA a nucleotide sequence listing in electronic form.

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.

PCT/US2009/051993

Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application

☒ claims Nos. 13, 14, 16, 20, 21

because:

☐ the said international application, or the said claims Nos. _____
relate to the following subject matter which does not require an international search (*specify*):

☒ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 13, 14, 16, 20, 21
are so unclear that no meaningful opinion could be formed (*specify*):

Claims 13, 14, 16, 20, 21 delimit genomic sequences for cellulases or nucleotide sequence itself. However, the exact sequence for the cellulases is unclear from the expression written in claims 13, 14, 16, 20 and there is no electronic form for claim 21.

☐ the claims, or said claims Nos. _____ are so inadequately supported
by the description that no meaningful opinion could be formed (*specify*):

☒ no international search report has been established for said claims Nos. 13, 14, 16, 20, 21

☒ a meaningful opinion could not be formed without the sequence listing; the applicant did not, within the prescribed time limit:

☐ furnish a sequence listing on paper complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.

☒ furnish a sequence listing in electronic form complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.

☐ pay the required late furnishing fee for the furnishing of a sequence listing in response to an invitation under Rule 13ter.1(a) or (b).

☒ See Supplemental Box for further details.

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.

PCT/US2009/051993

Box No. V Reasoned statement under Rule 43bis.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	1-12,15,17-19	YES
	Claims	NONE	NO
Inventive step (IS)	Claims	NONE	YES
	Claims	1-12,15,17-19	NO
Industrial applicability (IA)	Claims	1-12,15,17-19	YES
	Claims	NONE	NO

2. Citations and explanations :

Reference is made to the following document:

D1: US 20070178569 A1

The present invention relates to methods and compositions for improving the production of products such as fuel products like ethanol in microorganisms.

D1 provides methods for producing biofuels such as ethanol and related materials from renewable and sustainable biomass resources.

1. Novelty & Inventive Step (PCT Article 33(2) & 33(3))

Regarding claims 1-12, 15, 17-19

D1, which is considered to be the closest state of the art to claims 1-12, 15, 17-19 of the present application, provides systems and a method for producing ethanol, hydrogen and organic acid including lactic acid, acetic acid and formic acid from lignocellulosic and carbonaceous biomass by *Clostridium phytofermentans*. The differences between D1 and the above claims lie in that the above claims additionally delimit the microorganism modified to provide enhanced activity of one or more cellulase. However, D1 mentions that *Clostridium phytofermentans* contains exceptionally broad diversity of carbohydrate transport systems well suited to decomposing cellulosic biomass and to be genetically modified strains. Therefore, it is falling within the scope of the customary practice followed by an ordinary skilled person in the art to develop the strains of cellulase activity enhanced. Accordingly, the subject matter of claims 1-12, 15, 17-19 is considered to be novel under PCT Article 33(2), but does not meet the requirement of PCT Article 33(3).

2. Industrial Applicability

Claims 1-12, 15, 17-19 meet the requirements for industrial applicability in PCT Article 33(4).

**WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY**

International application No.

PCT/US2009/051993

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of :

Box No. III.

A sequence listing in electronic form was not furnished to this Authority. Therefore, a meaningful opinion on claims 13, 14, 16, 20, 21 could not be made.

MÉTODOS Y COMPOSICIONES PARA MEJORAR LA PRODUCCIÓN DE PRODUCTOS EN MICROORGANISMOS

RESUMEN – OBJETO Y FINALIDAD DE LA INVENCION

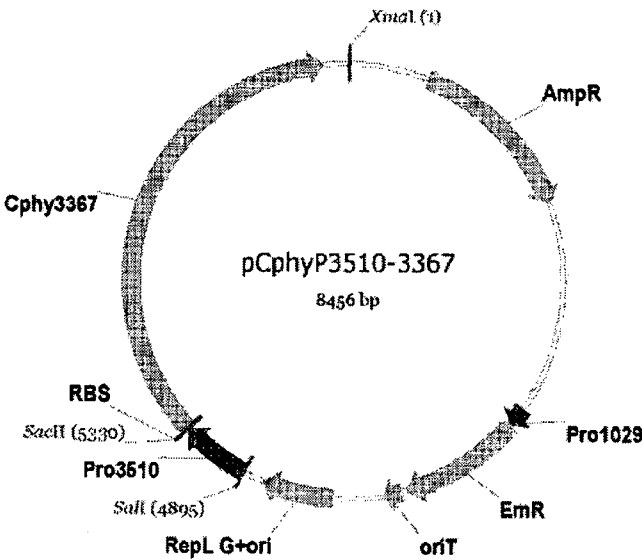


FIG. 19

Se proveen métodos y composiciones para mejorar la producción de productos, tales como productos combustibles como etanol, en microorganismos. En particular, se describen métodos y composiciones para mejorar la producción de etanol utilizando genes identificados en *Clostridium phytofermentans*.

MÉTODOS Y COMPOSICIONES PARA MEJORAR LA PRODUCCIÓN DE PRODUCTOS EN MICROORGANISMOS

REIVINDICACIÓN DE PRIORIDAD

Esta solicitud reivindica el beneficio de prioridad de la Solicitud Provisional de Patente de E.U. Serial No. 61/084,233 presentada en Julio 28, 2008, Solicitud Provisional de Patente de E.U. Serial No. 61/225,184 presentada en Julio 13, 2009, y Solicitud Provisional de Patente de E.U. Serial No. 61/228,922, presentada en Julio 17, 2009, y los contenidos enteros de las mismas que se incorporan aquí por referencia.

CAMPO DE LA INVENCION

La presente invención se relaciona con el campo de la microbiología, biología molecular y biotecnología. Más específicamente, la presente invención se relaciona con métodos y composiciones para mejorar la producción de productos, tales como etanol e hidrógeno, en microorganismos.

ANTECEDENTES

Existe un interés en desarrollar métodos y productos con el propósito de encontrar una fuente alternativa de energía a partir de fuentes renovables y sostenidas resultado de la biomasa. Las manifestaciones inherentes de recursos energéticos contenidos en la biomasa - tal como lo son los carbohidratos - resultarían ser un desperdicio de esta y de los cultivos de los cuales se obtienen estos productos, tales como los cereales maíz y trigo, o pasturas como el pasto forrajero.

Un desafío actual es crear estrategias viables y redituables para la conversión de de carbohidratos en energía utilizable. Las estrategias para derivar energía aprovechable a partir de los carbohidratos incluyen la elaboración de etanol y otros alcoholes conversión de carbohidratos en oxígeno y la conversión directa de carbohidratos en energía eléctrica por intermedio de celdas de combustible.

Ejemplos de procesos para derivar etanol de distintos tipos de biomasa se encuentran disponibles en: *Journal of Outlook for Biomass Ethanol Production and Demand (EIA Forecasts)*, 2002; Sheehan, *Biotechnology Progress*, 15:8179, 1999; Martin, *Enzyme Microbes Technology*, 31 :274, 2002; Greer, *BioCycle*, 61-65, April 2005; Lynd, *Microbiology and Molecular Biology Reviews*, 66:3, 506-577, 2002; and Lynd et al. in "Consolidated Bioprocessing of Cellulosic Biomass: An Update," *Current Opinion in Biotechnology*, 16:577-583, 2005.

RESUMEN

Esta aplicación se sustenta, entre otros aspectos, en la identificación de los genes encriptados dentro del *Clostridium phytofermentans* que pueden estar potencialmente relacionados con el crecimiento de sustrato útiles para la producción de compuestos tales como combustibles, e.g., etanol e hidrógeno. Los genes identificados dentro de la presente pesquisa pueden expresarse heterológamente en otros micro-organismos para proveer nuevas o mejoradas funciones. También, los genes pueden ser expresados en el *Clostridium phytofermentans*, e.g., de un ácido nucleico introducido exogénicamente para proveer el aumento del desempeño de las funciones.

Algunas realizaciones incluyen poli nucleótidos contiene un nucleótido asilado en el que se halla encriptado como mínimo una hidrolasa identificada en la *Clostridium phytofermentans*. En tales versiones el ácido nucleico aislado puede ser seleccionado de la tabla 6. En algunas de estas, tanto más específicas, la hidrolasa es seleccionada del grupo consistente en Cphy3367, Cphy3368, Cphy0430, Cphy3854, Cphy0857, Cphy0694, and Cphy1929. La designación Cphy3367 representa el número JGI, el cual hace referencia al Centro Nacional para la Información en Bio-tecnología (NCBI, por sus iniciales en inglés) a la marca de posición genética en el registro de GenBank para *Clostridium phytofermentans*. En presentaciones venideras los poli-nucleótidos podrán llegar a presentar la característica de una secuencia regulatoria operativa asociada al ácido nucleico aislado encriptado al interior de la hidrolasa.

Algunas realizaciones incluyen poli nucleótidos que contienen un ácido nucleico en que se codifica como mínimo un transportador con un unión ATP en casete (*ABC*) identificado en la *Clostridium phytofermentans*. En tales versiones el ácido nucleico aislado puede ser escogido de la tabla 7. En presentaciones específicas el transportador *ABC* es seleccionado del grupo conformado por Cphy3854, Cphy3855, Cphy3857, Cphy3858, Cphy3859, Cphy3860, Cphy3861, and Cphy3862. En próximas realizaciones le será posible al poli nucleótido contener una secuencia regulatoria operativa asociada al ácido nucleico codificando al interior del transportador *ABC*.

Algunas realizaciones incluyen poli-nucleótidos que contienen un ácido nucleico codificando al menos un regulador transcripcional identificado en la *Clostridium phytofermentans*. En tales presentaciones el ácido nucleico puede ser seleccionado de la tabla 8. En futuras presentaciones los poli-nucleótidos pueden llegar a incluir una secuencia reguladora operativa asociada al ácido nucleico aislado codificando el regulador transcripcional.

Algunas realizaciones incluyen casetes que contienen cualquier combinación de los ácidos nucleicos codificando las hidrolasas, los transportadores *ABC* y los reguladores transcripcionales descritos aquí. En una presentación un casete de poli-nucleótido puede contener un ácido nucleico aislado codificando al menos un transportador *ABC*. En otra presentación un casete de poli-nucleótido puede contener un ácido nucleico codificando al menos una hidrolasa y un ácido nucleico codificando un transportador *ABC*. En otra presentación un casete puede contener un ácido nucleico codificando al menos a un regulador transcripcional. En otra presentación un casete poli-nucleótido puede contener un ácido nucleico codificando al menos una hidrolasa y un ácido nucleico codificando al menos a un regulador transcripcional. En otra presentación un casete poli-nucleótido puede contener un ácido nucleico aislado codificando al menos un transportador *ABC* y un ácido nucleico codificando al menos un regulador transcripcional. En otra presentación un casete poli-nucleótido puede albergar a un ácido nucleico codificando al menos un transportador *ABC* y un ácido nucleico codificando un regulador transcripcional. E incluso, en otra presentación, un casete poli-nucleótido puede contener al menos una hidrolasa y un ácido nucleótido aislado con al menos un transportador *ABC* y un ácido nucleico codificando como mínimo un regulador transcripcional.

Algunas realizaciones incluyen la expresión de los casetes que contengan a cualquier poli-nucleótido descrito en la presente descripción y una secuencia regulatoria operativa asociada al casete poli-nucleótido.

Algunas realizaciones incluyen organismos recombinantes que contienen algún poli-nucleótido, poli nucleótido casete y / o sus expresión casete descrita aquí. En presentaciones particulares, los organismos recombinantes pueden ser seleccionados del grupo constituido por el *Clostridium cellulovorans*, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium josui*, *Clostridium papyrosolvens*, *Clostridium cellobioparum*, *Clostridium hungatei*, *Clostridium cellulosi*, *Clostridium stercorarium*, *Clostridium termitidis*, *Clostridium thermocopriae*, *Clostridium celerecrescens*, *Clostridium polysaccharolyticum*, *Clostridium populeti*, *Clostridium lentocellum*, *Clostridium chartatabidum*, *Clostridium aldrichii*, *Clostridium herbivorans*, *Acetivibrio cellulolyticus*, *Bacteroides cellulosolvens*, *Caldicellulosiruptor saccharolyticum*, *Ruminococcus albus*, *Ruminococcus flavefaciens*, *Fibrobacter succinogenes*, *Eubacterium cellulosolvens*, *Butyrivibrio fibrisolvens*, *Anaerocellum thermophilum*, *Halocella cellulolytica*, *Thermoanaerobacterium thermosaccharolyticum* y *Thermoanaerobacterium saccharolyticum*.

Algunas realizaciones incluyen proteínas aisladas que con una hidrolasa codificada en la *Clostridium phytofermentans*. En Algunas realizaciones se incluyen los métodos para la producción de etanol. Tales métodos comprenden el cultivo de un micro-organismo, proveer el sustrato y cualquier proteína descrita aquí.

Algunas realizaciones incluyen poli-nucleótidos casete aislados que incluyen uno o más, dos o más o incluso tres de: a.) una secuencia que codifica la hidrolasa del *Clostridium phytofermentans*, una secuencia codificando al transportador del *Clostridium phytofermentans* ATP unión casete (ABC) y una secuencia codificando al regulador transcripcional de la *Clostridium phytofermentans*. En Algunas realizaciones la hidrolasa es seleccionada de del grupo conformado por: *Cphy3368*, *Cphy3367*, *Cphy1799*, *Cphy1800*, *Cphy2105*, *Cphy1071*, *Cphy0430*, *Cphy1163*, *Cphy3854*, *Cphy1929*, *Cphy2108*, *Cphy3158*, *Cphy3207*, *Cphy3009*, *Cphy3010*, *Cphy2632*, *Cphy3586*, *Cphy0218*, *Cphy0220*, *Cphy1720*, *Cphy3160*, *Cphy2276*,

Cphyl714, *CphyO694*, *Cphy3202*, *Cphy3862*, *CphyO858*, *Cphyl510*, *Cphy2128*, *Cphyl 169*, *Cphyl888*, *Cphy2919*, y *Cphyl612*. En Algunas realizaciones el transportador ABC es seleccionado del grupo integrado por: *Cphyl529*, *Cphyl530*, *Cphyl531*, *Cphy3858*, *Cphy3859*, *Cphy3860*, *Cphy2569*, *Cphy2570*, *Cphy2571*, *Cphy2654*, *Cphy2655*, *Cphy2656*, *Cphy3588*, *Cphy3589*, *Cphy3590*, *Cphy3210*, *Cphy3209*, *Cphy3208*, *Cphy2274*, *Cphy2273*, *Cphy2272*, *Cphy2268*, *Cphy2267*, *Cphy2266*, *Cphy2265*, *Cphy2012*, *Cphy2011*, *Cphy2010*, *Cphy2009*, *Cphyl717*, *Cphyl716*, *Cphyl715* *Cphyl451*, *Cphyl450*, *Cphyl449*, *Cphyl448*, *Cphyl 134*, *Cphyl 133* y *Cphyl 132*.

Algunas realizaciones incluyen micro-organismos recombinantes que incluyen el ácido nucleico revelado aquí, e.g., uno o más, dos o más, o todos los tres de: a.) un ácido nucleico exógeno codificante la hidrolasa del *Clostridium phytofermentans*, un ácido nucleico exógeno codificante de un transportador para hidrolasa de *Clostridium phytofermentans* de unión ATP de tipo casete (ABC) y un regulador transcripcional de un ácido nucleico exógeno codificante del *Clostridium phytofermentans*. En Algunas realizaciones la hidrolasa es seleccionada del grupo conformado por: *Cphy3368*, *Cphy3367*, *Cphyl799*, *Cphyl800*, *Cphy2105*, *CphylO71*, *Cphy0430*, *Cphyl 163*, *Cphy3854*, *Cphyl929*, *Cphy2108*, *Cphy3158*, *Cphy3207*, *Cphy3009*, *Cphy3010*, *Cphy2632*, *Cphy3586*, *CphyO218*, *Cphy0220*, *Cphyl720*, *Cphy3160*, *Cphy2276*, *Cphyl714*, *CphyO694*, *Cphy3202*, *Cphy3862*, *CphyO858*, *Cphyl510*, *Cphy2128*, *Cphyl 169*, *Cphyl888*, *Cphy2919*, and *Cphyl612*. En Algunas realizaciones el transportador ABC es seleccionado del grupo integrado por: *Cphyl529*, *Cphyl530*, *Cphyl531*, *Cphy3858*, *Cphy3859*, *Cphy3860*, *Cphy2569*, *Cphy2570*, *Cphy2571*, *Cphy2654*, *Cphy2655*, *Cphy2656*, *Cphy3588*, *Cphy3589*, *Cphy3590*, *Cphy3210*, *Cphy3209*, *Cphy3208*, *Cphy2274*, *Cphy2273*, *Cphy2272*, *Cphy2268*, *Cphy2267*, *Cphy2266*, *Cphy2265*, *Cphy2012*, *Cphy2011*, *Cphy2010*, *Cphy2009*, *Cphyl717*, *Cphyl716*, *Cphyl715* *Cphyl451*, *Cphyl450*, *Cphyl449*, *Cphyl448*, *Cphyl 134*, *Cphyl 133*, y *Cphyl 132*.

En Algunas realizaciones, los micro-organismos se seleccionaron de un grupo consistente de: *Clostridium cellulovorans*, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium josui*, *Clostridium papyrosolvens*, *Clostridium cellobioparum*, *Clostridium hungatei*, *Clostridium cellulosi*, *Clostridium stercorarium*, *Clostridium termitidis*, *Clostridium thermocopriae*, *Clostridium celerecrescens*, *Clostridium polysaccharolyticum*, *Clostridium populeti*, *Clostridium lentocellum*,

Clostridium chartatabidum, *Clostridium aldrichii*, *Clostridium herbivorans*, *Acetivibrio cellulolyticus*, *Bacteroides cellulosolvens*, *Caldicellulosiruptor saccharolyticum*, *Ruminococcus albus*, *Ruminococcus flavefaciens*, *Fibrobacter succinogenes*, *Eubacterium cellulosolvens*, *Butyrivibrio βbrisolvens*, *Anaerocellum thermophilum*, *Halocella cellulolytica*, *Thermoanaerobacterium thermosaccharolyticum* y *Thermoanaerobacterium saccharolyticum*.

Algunas realizaciones incluyen los métodos para producir etanol que incluyen el cultivo de al menos un micro-organismo recombinante descrito aquí. Tales presentaciones pueden proveer un sustrato para el micro-organismo. En particular versiones en las que el sustrato puede seleccionarse de entre un grupo integrado por aserrín, harina de madera, pulpa de madera, pulpa de papel, residuos de troncos de la madera de la pulpa para papel, forrajes como los del pasto forrajero, biomasa de plantas y cultivos tales como los del crambe, algas, cascajo de arroz, bagazo, hojas de yute, materia de macro algas, materia de micro algas, cortes de hierba, rastrojo de maíz, tusas de mazorca, grano de maíz, maíz trillado, granos de los que se pueda destilar y pectina. En ciertas presentaciones la sustancia puede ser pectina.

Algunas realizaciones incluyen métodos para procesar un sustrato de una hidrolasa que incluye el suministro de un micro-organismo que exogénicamente exprese una hidrolasa de *Clostridium phytofermentans* y provea el sustrato de la hidrolasa para el micro-organismo, tal como es el sustrato es procesado para conformar un producto. En Algunas realizaciones los micro-organismo exogénicamente expresan un transportador ATP de *Clostridium phytofermentans* de unión de casete (ABC) que exporta o importa el producto.

Algunas realizaciones incluyen un compuesto para la producción de un bio-combustible que contiene biomasa lignino-celulósica y un micro-organismo que es capaz de inducir hidrólisis directa y fermentación de la biomasa mencionada, al interior de la cual el micro-organismo se modifica para proveer una actividad mejorada de una o más celulasas (e.g., una o más celulasas reveladas aquí, e.g., Cphy3367, Cphy3368, CphyO218, Cphy3207, Cphy2058, y Cphyl 163). En Algunas realizaciones es capas de la directa fermentación de carbono de cinco y seis azúcares. En Algunas realizaciones el micro-organismo es una bacteria, e.g., una especie de *Clostridium*, e.g., *Clostridium phytofermentans*. En Algunas realizaciones

el micro-organismo comprende uno o más poli-nucleótidos heterólogos que realzan la actividad de una o más celulasas.

Algunas realizaciones incluyen un producto para la producción de un bio-combustible que incluye una biomasa carbonácea y un micro-organismo que es capaz de la hidrólisis directa y la fermentación de dicha bio-masa, en la que dicho micro-organismo es modificado para proveer actividad reforzada de una o más celulasas. (e.g., una o más celulasas reveladas aquí, e.g., Cphy3367, Cphy3368, CphyO218, Cphy3207, Cphy2058, y Cphyl 163). En Algunas realizaciones el micro-organismo es capaz de producir fermentados acabados. En Algunas realizaciones una parte sustancial de de los productos fermentados terminados incluyen ácido láctico, ácido acético y / o ácido fórmico. En Algunas realizaciones el micro-organismo es capaz de absorber una o más carbohidratos complejos. En Algunas realizaciones la biomasa tiene una concentración alta de carbohidratos oligo-méricos comparado con los carbohidratos mono-méricos.

Algunas realizaciones incluyen un proceso para la producción de bio-combustible que incluye:

- a.) Entrar en contacto con biomasa carbonácea con un micro-organismo que es capaz de llevar a la hidrólisis directa de dicha biomasa. De aquí que el micro-organismo modificado incrementa la actividad de una o más enzimas de celulasa (e.g., una o más celulasas reveladas aquí, e.g., Cphy3367, Cphy3368, CphyO218, Cphy3207, Cphy2058, and Cphyl 163); y
- b.) Permitir el tiempo necesario para que se efectúen la hidrólisis y la fermentación para la creación del bio-combustible.

En Algunas realizaciones el micro-organismo es capaz de absorber de uno más carbohidratos complejos. En Algunas realizaciones la biomasa tiene un más alto nivel de de concentración de carbohidratos oligo-méricos en comparación con los carbohidratos mono-mericos. En Algunas realizaciones la hidrólisis deviene en una gran concentración de celubiosa y / o carbohidratos oligo-meros más largos en comparación de los mono-mericos.

La sección de los encabezados usados aquí es para propósitos organizacionales exclusivamente y no están diseñados para limitar el objeto referido

de modo alguno. Toda la literatura y materiales de índole afín citados en esta aplicación (patentes, aplicaciones de las patentes, artículos, libros, tratados y páginas web) han sido expresamente incorporados por referencia en su totalidad para cualquier propósito. Cuando se refiere a definiciones de términos aparecen y difieren de las definiciones ofrecidas en las enseñanzas presentes la definición dada en las actuales instrucción debe prevalecer. Será ponderada en su preeminencia "a propósito" de las escalas de temperatura, concentraciones y de tiempos discutidos en las descripciones anteriores, tales como las ligeras e insustanciales desviaciones que se encuentran contempladas en el trabajo previo. En esta aplicación el uso del número gramatical en singular significa específicamente plural, a menos que se especifique lo contrario. También el uso de "comprender" y "contener" no están apuntados para ser considerados como limitaciones específicas. Se da por entendido que ambos anteceden a una descripción general y la explicación ulterior es didáctica y explicativa exclusivamente y no restringen la invención. Los artículos "un" y "uno" son utilizados aquí para referirse a uno o más de uno (*i.e., al menos uno de...*) del objeto gramatical del artículo. Tal como sirve de ejemplo "*un elemento*" que se interpreta como un elemento o más de un elemento.

A menos que de otro modo se defina términos científicos y técnicos usados en conexión con la invención descrita aquí tendrán los significados que son usualmente entendidos por consuetudinarios dentro del oficio. Más aún, a menos que de modo expreso y manifiesto sea indispensable para figura el contexto términos en singular incluirán a los plurales y viceversa. De manera general nomenclaturas utilizadas vinculadas con , y técnicas de, cultivo de células y de tejidos, biología molecular, química de proteínas y oligo- o poli-nucleótidos e hibridación descritas aquí son de dominio público al interior de los profesionales del campo. Las técnicas usadas son estandarizadas como, por ejemplo, para la purificación y preparación de ácidos nucleicos, análisis químico, recombinación de ácido nucleico y síntesis oligonucleotida. Las reacciones enzimáticas y las técnicas de purificación se han desarrollado conforme a las especificaciones, o tal como de manera usual se efectúan en la usanza descrita aquí. Las técnicas y procedimientos aquí descritos son generalmente ejecutados conforme a los métodos convencionales bien conocidos en el oficio y tal como son descritos en las referencia generales y las más específicas consignadas en la siguiente referencia. Ver. e.g., *Sambrook et al, Molecular Cloning: A Laboratory Manual (Third ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 2000)*.

Las nomenclaturas empleadas en conjunto con los procedimientos y técnicas de laboratorio aquí especificados son bien conocidas y mayormente comunes a todos los especialistas del ramo.

Como se utiliza en concordancia con las realizaciones provistas aquí, las siguientes definiciones, a menos que se indique lo contrario, serán entendidas con los siguientes significados:

"Nucleótido" se referirá a un fosfato-éster de un nucleosido, como una unidad monómera o dentro de un ácido nucleico. "Nucleótido 5'-trifosfato" se refiere a un nucleótido con un grupo trifosfato ester en la posición 5' y es en ocasiones denotado como " " o "dNTP" y "ddNTP" para señalar de modo más concreto las características estructurales del azúcar ribosa. El grupo ester trifosfato puede incluir sustituciones de sulfuros por varios oxígenos, e.g. α -dio-nucleótido 5'-trifosfato. Para una reseña de la química de ácidos nucleicos se sugiere ver: *Shabarova, Z. and Bogdanov, A. Advanced Organic Chemistry of Nucleic Acids, VCH, New York, 1994.*

El término "ácido nucleico" y "molécula de ácido nucleico" se refieren a las secuencias naturales de ácido nucleico tales como el ADN (ácido desoxiribunocleico) y ARN (ácido ribunocleico) ácidos nucleicos artificiales, análogos de los mismos y combinaciones a partir de ellos.

Tal como se usado aquí, los términos "poli-nucleótido" y "oligonucleótido" son usados de forma intercambiable y significan polímeros mono-catenarios y bi-catenarios de los monómeros de nucleótidos (ácidos nucleicos) incluyendo, pero no limitados a, 2' de-oxi-ribu-nucleótidos (ácido nucleico) y ribuno-cleotidos (ARN) vinculado a fosfo-di-éster inter-nucleótido con enlaces e.g. 3'-5' y 2'-5', vínculos invertidos como el 5'-5', estructuras ramificadas o ácidos nucleicos análogos. Los poli-nucleótidos se han alineado con contra tales como H^+ , NH_4^+ , tri-alquila-amonio, Mg^{2+} , Na^+ y afines. Un poli-nucleótido puede ser compuesto por completo de de-oxi-ribunocleotidos, enteramente de ribunocleotidos o de mixturas quiméricas a partir de ambas. Los poli-nucleótidos pueden ser compuestos de nucleo-base y análogos al

azúcar. Los poli-nucleótidos típicamente se ubican en un tamaño de unas pocas unidades monoméricas , como por ejemplo 5-40 cuando estas son más frecuentemente asociadas a la criba del oficio como oligo-nucleótidos, lo cual es extensivo a miles de unidades de nucleótidos monoméricos. A menos que se indique lo contrario, cuando la secuencia de un poli-nucleótido sea representada por esta se entenderá que los nucleótidos están en el orden de 5' a 3' de izquierda a derecha y que "a" señala de-oxi-adenosina, "C" indica de-oxi-citidina, "G" es por de-oxi-guanosina y "T" implica timidina.

"Combustibles y / u otros químicos" son usados aquí para referirse a compuestos idóneos como combustibles líquidos o gaseosos incluyendo, aunque no limitándose, a hidrocarburos, hidrógeno, metano, compuestos hidróxidos tales como alcoholes (e.g. etanol, butanol, propanol, metanol, etc.), compuestos de carbonilos tales como aldehídos y cetonas (e.g. acetona, formo-aldehído, 1-propanal, etc.), ácidos orgánicos , derivados de ácidos orgánicos tales como esterres (e.g. esterres de cera , glicéridos, etc.) y otros compuestos funcionales incluyendo, pero no limitándose a 1, 2-propanediol, 1, 3-propanediol, ácido lácteo, ácido fórmico, ácido acético, ácido succínico y ácido pirúvico producido por enzimas tales como la celulasa, polisacáridos, lipasas, proteasas, ligninas y hemi-celulasas.

El término "plásmido" se refiere al vector ácido circular nucleico. Generalmente, los plásmidos contienen un origen de replicación que permite muchas copias del plásmido para ser producido en una célula bacteriana, aunque en ocasiones podría ser eucariota, sin que esta se integre al plásmido dentro de la célula anfitriona de ADN.

El término "constructo" es usado aquí para referirse a la secuencia de nucleótidos recombinante, generalmente una molécula de ácido que ha sido generada con el propósito de la expresión de una secuencia específica de nucleótido(s) o para ser usada en la construcción de otra secuencia de nucleótido recombinante. En general "constructo" es empleado aquí para referirse a una molécula de ácido recombinante.

Una "casete de expresión" se refiere a un ensamble de elementos poli-

nucleótidos que permiten la transcripción de un poli-nucleótido a una célula anfitriona. Típicamente la expresión casete incluye un promotor y un heterólogo o una secuencia poli-nucleotida nativa que sea transcrita. Las expresiones casete o constructos también pueden incluir, e.g., señales de finalización de la transcripción, señales de poli-adenilación y elementos mejoradores.

Por "vector de expresión" se entiende al vector que permite la expresión del poli-nucleótido en el interior de la célula. La expresión del poli-nucleótido dentro de la célula. La expresión del poli-nucleótido incluye a los eventos transcripcionales y / o post-transcripcionales. Una "expresión constructo" es una expresión vector dentro de la cual una secuencia de nucleótido de interés ha sido insertada de manera tal que sea operativa asociada a la expresión de las secuencias presentes en el vector de expresión.

Un "operón" se refiere a una unidad de elementos de poli-nucleótidos que produce un mensajero ARN (mARN). Típicamente el operón incluye a un promotor y uno o más genes estructurales. Típicamente el operón contiene uno o más genes estructurales los cuales son transcritos al interior de un mARN polisitrónico. Una molécula de mARN sencilla que codifica más de una proteína. En algunas realizaciones un operón puede contener un operador que regula la actividad de los genes estructurales de un operón.

El término "célula hospedera" como es usado aquí se refiere a la célula que es transformada mediante los métodos y composiciones de la invención. En general las células anfitrionas como se emplean aquí significan que una célula de un micro-organismo dentro de la cual un ácido nucleico de interés es inoculado.

El término "transformación" como es usado aquí se refiere a la cambio genético total o transitorio, e.g., un cambio genético permanente inducido en una célula seguida de la incorporación de las secuencias de ácido no anfitrionas.

El término "célula transformada" como es usado aquí se refiere a una célula al interior de la cual (o al interior de la predecesora de esta) se ha introducido por

medio de las técnicas de recombinación de ácido nucleico, una molécula de ácido codificante de un gen producto de interés, tal como por ejemplo ARN y / O una proteína.

El término "gen" como se usa aquí se refiere a alguna y a todas las regiones codificantes de un genoma hospedero, o regiones en las que se codifica solo un ARN funcional (e.g., tARN, rARN, ARN's regulatorios tales como los ribosomas) e incluye regiones no codificantes y regiones regulatorias. El término "gen" incluye dentro de su alcance la lectura abierta de marcos codificantes de poli-péptidos específicos, intrones y y nucleótidos no adyacentes 5' y 3' secuencias no codificantes relacionadas con la regulación de la expresión. En este apartado un gen puede con el tiempo conformar señales de control tales como promotores, potenciadores y / o señales de terminación que están naturalmente asociadas con un gen dado, o señales de control heterólogo. Una secuencia de gen puede ser cADN o de ácido nucleico genómico o un fragmento obtenido a partir de aquellos. Un gen puede ser introducido dentro de un vector conveniente para el mantenimiento extra-cromo somático o para la integración del hospedero.

El término "gen de interés", "secuencia de nucleótido de interés", "poli-nucleótido de interés" o "ácido nucleico de interés" como son usados aquí se refieren a cualquier nucleótido ácido nucleico que codifique una proteína u otra molécula que sea deseable para ser expresada en su célula anfitriona (e.g., para la producción de proteína u otra molécula biológica e.g., un producto ANR) como objetivo de la célula. La secuencia del nucleótido de interés puede ser operativamente asociada a otras secuencias dentro de las cuales se facilitará la expresión, e.g., un promotor.

El término "promotor" es usado aquí para referirse a una secuencia mínima nucleica apenas suficiente para dirigir la transcripción de un ácido nucleico que pueda asociar y vincularse. El término "promotor inducible" tal como se ha usado aquí se refiere a un promotor que es transcripcionalmente activo cuando se lía a un activador transcripcional, el cual reaccionará bajo circunstancias específicas e.g., en la presencia de una señal química particular o combinación de señales químicas que afectan la unión del activador transcripcional al promotor y / o afectan la función del activador transcripcional mismo.

Los términos "operador", "secuencias de control" o "secuencia regulatoria" tal como se han usado aquí se refieren a secuencias de ácido nucleico que regulan la expresión de un vínculo codificante operativo de una secuencia en una célula anfitriona. Las secuencias de control que son convenientes para células procariotas, como por ejemplo incluyendo a un promotor, opcionalmente un operador de secuencia y una posición de unión ribosomática.

Por "conexión operativa" o "vínculo operativo" y afines se quiere dar a entender un enlazamiento de los elementos poli-nucleótidos en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor o un estimulante está operativamente vinculado a una secuencia codificante si esta llega a afectar la transcripción de la secuencia codificante. En Algunas realizaciones operablemente vinculadas significa que las secuencias de ácido nucleico que se encuentran asociadas se dan de manera típicamente contigua y donde se hace necesario unir proteínas a regiones codificantes contiguas y en plano de lectura. Una secuencia codificante está "operativamente asociada a" otra secuencia cuando el ARN polimerasa se transcriba a las dos secuencias dentro de una mRNA sencilla, la cual será traducida en un polipéptido simple teniendo aminoácidos derivados de ambas secuencias codificantes. Las secuencias codificantes no necesitan de estar contiguas una de la otra mientras las secuencias llegan finalmente a procesarse dando así como resultado la proteína deseada.

"Conectándose operativamente" un promotor transcribe poli-nucleótidos para transcribir por los medios de los poli-nucleótidos bajo el control regulatorio del promotor, el cual controla la transcripción y opcionalmente la traducción de aquel poli-nucleótido. En la construcción de combinaciones de genes promotores / estructurales heterólogos es usual posicionar un promotor o una variante de aquel a una distancia del lugar exacto donde comienza la transcripción del poli-nucleótido transcribible, el que aproximadamente está a la misma distancia entre el promotor y el gen que controla en su posición natural; específicamente el gen del que el promotor es derivado. Como es conocido en el oficio, algunas variaciones en esta distancia pueden acomodarse sin perder la función. Similarmente el posicionamiento típico de un elemento regulatorio de la secuencia tal como un operador o estimulante con respecto al poli-nucleótido transcribible se emplazará bajo el control del

anterior definido por el posicionamiento del elemento en su posición natural. De modo más concreto, los genes de los cuales se ha derivado.

"Cultivar" significa incubar una célula u organismo bajo condiciones en las cuales la célula u organismo pueda soportar parte, cuando no todos, sus procesos biológicos. Como ejemplo, una célula que es cultivada puede crecer o reproducirse, o puede no ser biológicamente viable aunque manifieste actividad bio-química o procesos tales como replicación, transcripción, translación etc.

Por "organismo transgénico" se aduce a un organismo no humano (e.g., organismos unicelulares, micro-organismos) mamíferos, no mamíferos e.g., nematodos o drosfila) que no tengan una secuencia endógena de ácido presente en una porción de sus células o estén notablemente integradas a su línea de ácido nucleico.

El término "biomasa" como es usado aquí se refiere a una masa de material vivo o biológicamente activo e incluye tanto natural como procesado, tanto como materiales orgánicas de modo muy amplio.

"Recombinante" se refiere a los poli nucleótidos sintetizados o manipulados de otro modo in vitro (poli-nucleótidos recombinados) y los métodos para usar los recombinantes poli-nucleótidos y a los métodos en los que se usan los poli-nucleótidos recombinados para producir productos genéticos codificados por aquellos en células u otros sistemas biológicos disponibles. Como por ejemplo una poli-nucleótido clonado que puede bien ser insertado en un vector de expresión adecuado, tal como un plásmido bacteriano y el plásmido puede ser empleado para transformarse en una célula huésped adecuado. Un huésped celular que comprenda los poli nucleótidos recombinantes es denominado como "recombinante hospedero celular" o un "recombinante bacteriano" El gen es entonces expresado en el poli nucleótido recombinado para producir e.g., una "proteína recombinada" Además, un poli nucleótido recombinado puede servir en la función de no-codificante, como por ejemplo a modo de promotor, origen de replicación o emplazamiento de unión ribosomática.

El término "recombinación homóloga" se refiere al proceso de recombinación entre dos moléculas de ácido nucleico basadas en una secuencia de ácido nucleico similar. El término comprende recombinaciones recíprocas y no recíprocas también calificada como conversión de genes) Además de la recombinación puede ser el resultado de eventos con equivalencia o no-equivalentes. Intersecciones equivalentes tienen espacio entre dos secuencias equivalentes o regiones cromosómicas donde la yuxtaposición sucede entre idénticos (o sustancialmente idénticos) segmentos de secuencias no equivalentes de las regiones de los cromosomas. Yuxtaposiciones inequivalentes típicamente son el resultado de duplicaciones de genes y deleciones. Para una descripción de las enzimas y los mecanismos involucrados en las re-combinaciones homólogas se sugiere consultar: *Watson et al., Molecular Biology of the Gene pp 313-327, The Benjamin/Cummings Publishing Co. 4th ed. (1987).*

El término "no – homólogos o de integración aleatoria" se refiere a cualquier proceso por el cual el ácido es integrado al interior del genoma que no involucra recombinaciones homólogas. Este es en apariencia un proceso aleatorio en el que la incorporación puede ocurrir en cualquier momento en un amplio número de posiciones genómicas.

Una "secuencia de poli-nucleótido heterólogo" o un "ácido nucleico heterólogo" resultan ser términos relativos para un poli-nucleótido que es funcionalmente relacionado con otro poli-nucleótido, tal como una secuencia promotora, de manera tal que las dos secuencias de poli-nucleótidos no están acopladas en la misma relación mutua de manera efectiva. Secuencias de poli nucleótidos heterólogos, e.g., un promotor operablemente vinculado a un ácido nucleico heterólogo y un poli-nucleótido incluyendo su promotor nativo que ha sido insertado en el interior de vector heterólogo para conseguir la transformación en una célula anfitriona recombinada. Secuencias de poli- nucleótidos heterólogos son consideradas "exógenas" debido a que fueron introducidos en la célula anfitriona por medio de vías de transformación. No obstante, los poli nucleótidos pueden originarse de desde una fuente ajena de la misma referencia. La modificación de secuencias de poli nucleótidos heterólogos pueden ocurrir, e.g., por el tratamiento de un poli nucleótido que puede ser operablemente asociado a un elemento regulatorio: la Modificación

puede también suceder por técnicas tales como la muta génesis direccionada en locación.

El término “expresado endogénicamente” se refiere a poli nucleótidos que son nativos de la célula anfitriona y se expresan naturalmente en dicha unidad.

“Competente para expresarse” se refiere a una célula anfitriona que provee de suficiente ambiente celular para la expresión de poli nucleótidos endógenos y / O exógenos.

Esta solicitud se relaciona con la Solicitud Provisional de E.U. Serial No. 61/032,048, presentada en Febrero 27, 2008; Solicitud Internacional Serial No. PCT/US2009/35597, presentada en Febrero 27, 2009; Solicitud de E.U. Serial No. 12/419,211, presentada en Abril 6, 2009; Solicitud Provisional de E.U. Serial No. 61/060,620, presentada en Junio 11, 2008; y Solicitud de E.U. Serial No. presentada en Junio 11, 2009, cada una de las cuales se incorpora aquí por referencia en su totalidad para cualquier propósito.

Las siguientes figuras, descripción y ejemplos ilustran ciertas realizaciones de la presente invención en detalle. Aquellos versados en el arte reconocerán que hay numerosas variaciones y modificaciones que están abarcadas por su alcance. En concordancia, la descripción de ciertas realizaciones no debe estimarse como limitante del alcance de la presente invención.

BREVE DESCRIPCIÓN DE LAS FIGURAS

La FIG.1 es una serie de diagramas de ejemplos de combinaciones de genes para polinucleótidos. R representa una secuencia reguladora transcripcional; A, B y C representan secuencias que codifican un transportador del casete de unión a ATP (ABC); GH representa una secuencia que codifica una glicósido hidrolasa; y S representa la secuencia señal.

La FIG. 2 es una serie de diagramas de ejemplos específicos de combinaciones de genes en *C. phytofermentans*. Los números representan la

localización de secuencias específicas en el cromosoma de *C. phytofermentans*.

La FIG.3 es un diagrama del diseño del microarreglo Affymetrix de *C. phytofermentans*. Los guiones representan sondas de 24 bases sintetizadas sobre el microarreglo. Las cajas representan los marcos de lectura abiertos predichos, por ejemplo, las regiones de codificación de las proteínas. Se utilizan once sondas de 24 bases para medir el nivel de cada marco de lectura abierto (ORF, por sus siglas en inglés). Las regiones intergénicas se cubren en ambos lados del ADN mediante sondas de 24 bases separadas por una sola base de ADN.

La FIG. 4 es un diagrama del método de determinación de los límites del transcripto de ARNm. Un transcripto hipotético de ARNm incluye regiones que no codifican que se extienden 5' y 3' del correspondiente ORF predicho. Las sondas se representan con guiones. En este ejemplo, tres sondas a la izquierda (5') del ORF y dos sondas a la derecha (3') del ORF indicarían los límites del transcripto de ARNm.

La FIG. 5 es una representación del cromosoma de *C. phytofermentans*.

La FIG. 6 es una gráfica que muestra el contenido GC de segmentos del genoma de 1 kb como función de la distancia a lo largo del genoma de *C. phytofermentans*. Se numeraron seis islas genómicas con contenido GC > 50%. Estas seis regiones consisten de un total de dieciséis regiones de 1kb.

La FIG. 7 es un árbol de similaridad de la cepa *C. phytofermentans* y taxas relacionadas dentro de la clase Clostridia con base en secuencias génicas de ARNr de 16S. El agrupamiento I comprende la enfermedad que causa Clostridia, el agrupamiento III comprende Clostridia celulítica y el agrupamiento XIVa comprende microbios intestinales y las secuencias metagenómicas están en el género Clostridium. Los números en los nodos son los niveles de soporte del análisis de muestreo (porcentajes) basados en el análisis de vecindad de 1000 conjuntos de datos re muestreados. Se utilizó *Bacillus subtilis* como un outgroup. Barra, 4 sustituciones de nucleótidos por posición.

La FIG. 8 es una gráfica circular que muestra la cantidad de mejores coincidencias (corte en el valor e de 0.01) de *C. phytofermentans*. ISDg CDSs en otros genomas bacterianos secuenciados en la clase Clostridia.

Las FIG. 9A y 9B son gráficas circulares que muestran una comparación de los genes que codifican glicósido hidrolasa (GA) (9A) y todos los genes en diferentes organismos (9B) utilizando BLASTP.

La FIG. 10 es un árbol de similaridad que muestra la filogenia molecular de los

dominios de la familia glicósido hidrolasa GH9.

La FIG. 11 es un árbol de similaridad que muestra la filogenia molecular de los dominios de la familia glicósido hidrolasa GH5.

La FIG. 12 es un diagrama esquemático que muestra el ejemplo de hidrolasas putativas. Algunas hidrolasas pueden ser extracelulares o unidas a la membrana. GH: glicósido hidrolasas; CBM: dominio de unión a carbohidratos.

La FIG. 13 es una descripción de la ingesta y metabolismo de xilosa en *C. phytofermentans*.

La FIG. 14 es una descripción de la ingesta y metabolismo de fucosa en *C. phytofermentans*.

La FIG. 15 es una descripción de la ingesta y metabolismo de ramnosa en *C. phytofermentans*.

La FIG. 16 es una descripción de la regulación, ingesta y metabolismo de laminarina en *C. phytofermentans*.

La FIG. 17 es una descripción de la ingesta y metabolismo de celobiosa en *C. phytofermentans*.

La FIG. 18 es una descripción del mapa de plásmidos para pIMP-Cphy.

La FIG. 19 es una descripción del mapa de plásmidos para pCphyP3510-3367.

DESCRIPCIÓN DETALLADA

Varias realizaciones reveladas aquí están dirigidas generalmente hacia composiciones y métodos para hacer microorganismos recombinantes que son capaces de producir un combustible cuando crece bajo una variedad de condiciones de fermentación. Generalmente, un microorganismo recombinante puede producir de forma eficiente y estable un combustible, tal como etanol, y compuestos relacionados, de modo que se provee un alto rendimiento a partir de materia prima de biomasa relativamente económica tal como celulosa. En algunas realizaciones, un microorganismo recombinante puede catalizar de forma eficiente y estable la conversión de materia prima de biomasa relativamente económica tal como

lignocelulosa, para producir sacáridos y polisacáridos, y compuestos relacionados.

En el presente, existe un número limitado de técnicas para utilizar organismos recombinantes que son capaces de producir un combustible. Las técnicas varias tienen frecuentemente problemas que pueden llevar a una baja producción de combustible, alto costo, y subproductos indeseables. Por ejemplo, algunas técnicas conocidas utilizan grano de maíz y otros cereales como materia prima. Sin embargo, competir por las demandas de alimentación y alimento en el suministro de granos y los precios puede limitar eventualmente la expansión de la producción de etanol a partir de maíz y otros cereales. Otras fuentes de materia prima incluyen lignocelulosa a partir de la cual se puede producir etanol vía sacarificación y fermentación (Lynd, L. R., Cushman, J. H., Nichols, R. J. & Wyman, C. E. Fuel ethanol from cellulosic biomass," Science 251, 1318-1323 (1991)). Ya que la lignocelulosa es el componente primario de la biomasa y el material biológico mas abundante en la tierra, los combustibles derivados de la biomasa de tipo lignocelulosa son por consiguiente alternativas de energía renovable que tienen el potencial para sostener la economía, energía y el ambiente en todo el mundo. Sin embargo, la producción de etanol lignocelulosico convencional requiere un proceso multipasos costoso y complejo que incluye la producción de y el pretratamiento de material lignocelulosico con enzimas sacarolíticas exógenas, la hidrólisis de polisacáridos presentes en la biomasa pretratada, y la fermentación separada de los azúcares de hexosas y pentosas.

En una realización, los métodos y las composiciones de la invención comprenden modificar o reconfigurar genéticamente un microorganismo para mejorar la actividad enzimática de una o más enzimas, incluyendo pero no limitado a celulasa(s). Ejemplos de dichas modificaciones incluyen modificar elementos endógenos reguladores de ácidos nucleicos para incrementar la expresión de una o más enzimas (e.g. vincular de forma operable un gen que codifica una enzima blanco con un promotor fuerte), introduciendo en un microorganismo copias adicionales de moléculas de ácido nucleico para proveer actividad mejorada de una enzima, vinculando de forma operable genes que codifican una o más enzimas a un promotor inducible o una combinación de los mismos.

Varios microorganismos de la invención pueden modificarse para mejorar la actividad de una o más celulasas, o enzimas asociadas con el procesamiento de la

celulosa. La clasificación de las celulasas se basa usualmente en agrupar enzimas para formar una familia con actividad idéntica o similar, pero no necesariamente la misma especificidad de sustrato. Una de estas clasificaciones es el sistema CAZY (CAZY por Enzimas Activas a Carbohidratos, en inglés), por ejemplo, en donde existen 115 Glicósido Hidrolasas diferentes (GH) listadas, nombradas GH1 a GH155. Cada una de las diferentes familias de proteínas usualmente tiene una correspondiente actividad enzimática.

Esta base de datos incluye las enzimas activas celulosa y hemicelulasa. Adicionalmente, el genoma entero anotado de *Clostridium phytofermentans* está disponible en la internet www.ncbi.nlm.nih.gov/sites/entrez.

Algunas realizaciones descritas aquí simplifican el proceso multipasos convencional de la producción de etanol lignocelulósico suministrando métodos y composiciones en donde la biomasa lignocelulósica se puede fermentar a etanol en un solo paso. Esto es conocido como bioprocesamiento consolidado (CBP, por sus siglas en inglés). Ya que el CBP agiliza todo el proceso de conversión, y reduce costos y desperdicio de energía, se prevé como el único bioproceso de etanol celulósico económica y ambientalmente sostenible.

En algunas realizaciones, se proveen los polinucleótidos y los casetes de expresión para un sistema eficiente de producción de combustible. Los polinucleótidos y los casetes de expresión se pueden utilizar para preparar vectores de expresión para transformar los microorganismos para conferirles a los microorganismos transformados la capacidad de producir productos eficientes, tales como combustible, en cantidades útiles.

En algunas realizaciones, el metabolismo de un microorganismo se puede modificar mediante la introducción y expresión de varios genes. De acuerdo con algunas realizaciones de la presente invención, los microorganismos recombinantes pueden utilizar genes de *Clostridium phytofermentans* (ISDgT, 700394T de American Type Culture Collection) como biocatalizador para la conversión mejorada de, por ejemplo, celulosa, a un combustible, tal como etanol e hidrógeno.

En algunas realizaciones, *C. phytofermentans* (700394^T de American Type Culture Collection) se puede definir con base en las características fenotípicas y genotípicas de una cepa cultivada, ISDg^T (Warnick et al, International Journal of Systematic and Evolutionary Microbiology, 52:1155-60, 2002). el genoma entero anotado de *Clostridium phytofermentans* está disponible en la internet en www.ncbi.nlm.nih.gov/sites/entrez. Varias realizaciones generalmente se relacionan con sistemas, y métodos y composiciones para producir combustibles y/o otros productos orgánicos útiles que involucran la cepa ISDg^T y/o cualquier otra cepa de la especie *C. phytofermentans*, que se puede derivar de la cepa ISDg^T o aislar separadamente. La especie se puede definir utilizando consideraciones taxonómicas estándar (Stackebrandt and Goebel, International Journal of Systematic Bacteriology, 44:846-9, 1994): las cepas con valores de homología de la secuencia de ARNr 16S de 97% y mayores como se compara con la cepa tipo (ISDg^T) se consideran cepas de *C. phytofermentans*, a menos que hayan mostrado tener valores de reasociación a ADN de menos de 70%. Existe considerable evidencia que indica que los microbios que tienen valores de reasociación de 70% o mayores también tienen al menos 96% de la identidad de la secuencia de ADN y comparten rasgos fenotípicos que definen una especie. Los análisis de la secuencia genómica de la cepa ISDg^T de *C. phytofermentans* indican la presencia de grandes cantidades de genes y loci genéticos que tienen tendencia a estar involucrados en mecanismos y vías para la fermentación de polisacáridos de plantas, dando lugar a las propiedades inusuales de fermentación de este microbio. Con base en las consideraciones taxonómicas mencionadas anteriormente, todas o casi todas las cepas de las especies de *C. phytofermentans* también poseerían estas propiedades de fermentación. Las cepas de *C. phytofermentans* pueden ser aislados naturales, o cepas genéticamente modificadas.

Se pueden introducir varios vectores de expresión en un microorganismo hospedero de modo que el microorganismo transformado puede producir grandes cantidades de combustible en varias condiciones de fermentación. Los microorganismos recombinantes se pueden modificar de modo que se produce un combustible estable con alto rendimiento cuando se hace crecer en un medio que comprende, por ejemplo, celulosa.

C. phytofermentans, solo o en combinación con uno o más microbios diferentes, puede fermentar a gran escala un material de biomasa celulósica en un biocombustible combustible, tal como etanol, propanol, y/o hidrógeno (ver, e.g. Solicitud de Patente de E.U. No. 2007/0178569; Warnick et al, Int J Syst Evol Microbiol (2002), 52 1155-1160, cada una de las cuales se incorpora aquí por referencia en su totalidad).

Los polinucleótidos, casetes de expresión y vectores de expresión revelados aquí se pueden utilizar con muchos microorganismos hospederos para la producción de combustibles tales como etanol e hidrógeno. Por ejemplo, además de *Clostridium phytofermentans*, microorganismos celulolíticos tales como *Clostridium cellulovorans*, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium josui*, *Clostridium papyrosolvens*, *Clostridium cellobioparum*, *Clostridium hungatei*, *Clostridium cellulosi*, *Clostridium stercorarium*, *Clostridium termitidis*, *Clostridium thermocopriae*, *Clostridium thermocellum*, *Clostridium celerecrescens*, *Clostridium polysaccharolyticum*, *Clostridium populeti*, *Clostridium lentocellum*, *Clostridium chartatabidum*, *Clostridium aldrichii*, *Clostridium herbivorans*, *Acetivibrio cellulolyticus*, *Bacteroides cellulosolvens*, *Caldicellulosiruptor saccharolyticum*, *Ruminococcus albus*, *Ruminococcus flavefaciens*, *Fibrobacter succinogenes*, *Eubacterium cellulosolvens*, *Butyrivibrio β brisolvens*, *Anaerocellum thermophilum*, y *Halocella cellulolytica* son hospederos particularmente atractivos, puesto que son capaces de hidrolizar celulosa. Otros microorganismos que se pueden utilizar incluyen, por ejemplo, microbios sacarolíticos tales como *Thermoanaerobacterium thermosaccharolyticum* y *Thermoanaerobacterium saccharolyticum*. Hospederos potenciales adicionales incluyen otras bacterias, levaduras, algas, hongos y células eucarióticas.

En varias realizaciones, los polinucleótidos, casetes de expresión y vectores de expresión revelados aquí se pueden utilizar con *C. phytofermentans* u otras especies de *Clostridia* para incrementar la producción de combustible tal como etanol e hidrógeno.

Como apreciará la persona versada en la materia, la habilidad para producir organismos recombinantes que pueden producir combustibles que tienen gran beneficio, especialmente para una producción de combustible eficiente, rentable y

ambientalmente amigable.

Realizaciones Ejemplares

La siguiente descripción y ejemplos ilustran algunas realizaciones de la presente invención en detalle. Aquellas personas versadas en la materia reconocerán que existen numerosas variaciones y modificaciones que están abarcadas por su alcance. En concordancia, la descripción de una realización preferida no debería estimarse como limitante del alcance de la presente invención.

Varias realizaciones de la invención ofrecen beneficios relacionados con la producción de combustibles utilizando microorganismos recombinantes. Los polinucleótidos, casetes de expresión, vectores de expresión y microorganismos recombinantes para la optimización de la producción de combustible se revelan de acuerdo con algunas realizaciones de la presente invención.

Hidrolasas

Algunas realizaciones descritas aquí se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* como hidrolasas codificantes. Algunas realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* como hidrolasas codificantes. Las ventajas para utilizar ácidos nucleicos que codifican hidrolasas incluyen mejorar las capacidades y desempeño de microorganismos para hidrolizar polímeros, por ejemplo, polisacáridos y polipéptidos.

Las hidrolasas pueden incluir enzimas que degradan polímeros tales como disacáridos, trisacáridos y polisacáridos, polipéptidos, y proteínas. Los polímeros también pueden incluir, por ejemplo, celulosas, hemicelulosas, pectinas, ligninas, y proteoglicanos. Ejemplos de enzimas y actividades enzimáticas que degradan polisacáridos pueden incluir, pero no están limitadas a, glicosido hidrolasas (GH), glicosil transferasas (GT), polisacárido liasas (PL), carbohidrato esterasas (CE), y proteínas que contienen módulos de unión a carbohidratos (CBM)- (disponible en la internet en "cazy.org"; Coutinho, P.M. & Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In "Recent Advances in Carbohydrate

Bioengineering," H.J. Gilbert, G. Davies, B. Henrissat and B. Svensson eds., The Royal Society of Chemistry, Cambridge, pp. 3-12).

En algunas realizaciones, GH, GT, PL, CE, y CMB pueden ser enzimas individuales con diferentes actividades. En otras realizaciones, GH, GT, PL, CE, y CMB pueden ser dominios enzimáticos con una actividad catalítica particular. Por ejemplo, una enzima con actividades múltiples puede tener múltiples dominios enzimáticos, incluyendo por ejemplo los dominios catalíticos de GH, GT, PL, CE, y/o CMB.

Las o-glicosil hidrolasas son un amplio grupo de enzimas que hidrolizan el enlace glicosídico entre dos o más carbohidratos, o entre un fragmento carbohidrato y uno no carbohidrato. Un sistema de clasificación para las glicosil hidrolasas, basado en la similaridad de la secuencia, ha llevado a la definición de 85 familias diferentes PUBMED: 7624375, PUBMED:8535779, PUBMED. Esta clasificación está disponible en el sitio web PUBMED CAZy (C Arbohydrate- Active EnZymes). Dado que el plegamiento de las proteínas se conserva mejor que sus secuencias, algunas de las familias se pueden agrupar en "clanes".

La familia glicósido hidrolasa 9 comprende enzimas con varias actividades conocidas, tales como endoglucanasa y celobiohidrolasa. En *C. phytofermentans*, una celulasa GH9 ejemplar es ABX43720.

Cualquier enzima hidrolítica se puede seleccionar a partir del genoma anotado de *C. phytofermentans* para utilización en productos y procesos de invención. Los ejemplos incluyen enzimas tales como una o más endoglucanasa, quitinasa, celbiohidrolasa o celulasas endoprocésivas (con extremo reductor o no).

Adicionalmente, un microorganismo, tal como *C. phytofermentans* se puede modificar para mejorar la producción de una o más enzimas celulasa o hidrolasa o una o más de dichas enzimas puede estar expresada de forma heteróloga en un hospedero diferente (e.g. otras bacterias o levaduras). Para la expresión heteróloga, las bacterias o levaduras se pueden modificar a través de tecnología recombinante (e.g., Brat et al. Appl. Env. Microbiol. 29; 75:2304-2311, revelando la expresión de la

xilosa isomerasa en *Saccharomyces cerevisiae*).

Se pueden hacer otras modificaciones para mejorar la producción del producto final (e.g. etanol) en un microorganismo recombinante de la invención. Por ejemplo, el hospedero puede comprender un segmento de ADN heterólogo adicional, cuyo producto de expresión es una proteína involucrada en el transporte de mono- y/o oligosacáridos en el hospedero recombinante. Asimismo, genes adicionales de la vía glicolítica se pueden incorporar al hospedero. De esta forma, se puede lograr una tasa mejorada de producción de etanol.

Una de las características más impactantes e inesperada del genoma de *C. phytofermentans* es la cantidad y diversidad de genes que codifican enzimas activas a carbohidratos. Esta diversidad no es equiparable en organismos relacionados con *C. phytofermentans*. La tabla 1 ilustra la diversidad de genes de carbohidratos en relación con otros organismos.

TABLA 1: Número y diversidad de genes activos a carbohidratos

Organismo	Cantidad de genes de glicósido hidrolasa	Familias glicósido hidrolasa
<i>C. phytofermentans</i>	109	39
<i>C.beijerinckii</i>	75	25
<i>C.botulinum</i>	23	10
<i>C.perfringens</i>	38	21
<i>C.thermocellum</i>	70	23
<i>Caldicellulosiruptor saccharolyticus</i>	61	31
<i>Thermoanaerobacter ethanolicus</i>	15	26

El genoma de *C. phytofermentans* incluye un rango diverso de genes GH, PL,

CE, y CBM con un amplio rango de funciones putativas predichas utilizando los métodos descritos aquí y los métodos bien conocidos en el arte. Las tablas 2 y 5 muestran ejemplos de algunas de las actividades conocidas de algunos de los miembros de la familia GH, PL, CE, y CBM, que se predice están presentes en *C. phytofermentans*, respectivamente. Las actividades conocidas se listan por actividad y correspondiente número EC según lo determinado por la Unión Internacional de Bioquímica y Biología Molecular.

TABLA 2: Actividades conocidas de los miembros de la familia glicósido hidrolasa

Familia Glicósido Hidrolasa	Actividades conocidas	Número de dominios predichos en <i>C. phytofermentans</i> .
1	beta-glucosidasa (EC 3.2.1.21); beta-galactosidasa (EC 3.2.1.23); beta-manosidasa (EC 3.2.1.25); beta-glucoronidasa (EC 3.2.1.31); beta-D-fucosidasa (EC 3.2.1.38); florizin-hidrolasa (EC 3.2.1.62); 6-fosfogalactosidasa (EC 3.2.1.85); 6-fosfo-beta-glucosidasa (EC 3.2.1.86); estrictosidinbeta-glucosidasa (EC 3.2.1.105); lactasa (EC 3.2.1.108); amigdalina-beta-glucosidasa (EC 3.2.1.117); prunasin-beta-glucosidasa (EC 3.2.1.118); raucafricina-beta-glucosidasa (EC 3.2.1.125); tioglucosidasa (EC 3.2.1.147); beta-primeverosidasa (EC 3.2.1.149); isoflavonoide 7-O-beta-apiosil-glucosidasa (EC 3.2.1.161); hidroxisourato hidrolasa (EC 3.-.-.-); beta-glicosidasa (EC 3.2.1.-)	1
2	beta-galactosidasa (EC 3.2.1.23); beta-manosidasa (EC 3.2.1.25); beta-glucoronidasa (EC 3.2.1.31); manosilglicoproteínaendo-beta-manosidasa (EC 3.2.1.152); exobeta-glucosaminidasa (EC 3.2.1.-)	5
3	beta-glucosidasa (EC 3.2.1.21); xilan 1,4-beta-xilosidasa (EC 3.2.1.37); beta-N-acetilhexosaminidasa (EC 3.2.1.52); glucan 1,3 beta-glucosidasa (EC 3.2.1.58); glucan 1,4 beta-glucosidasa (EC 3.2.1.74); exo-1,3-1,4-glucanasa (EC 3.2.1.-); alfa-L-arabinofuranosidasa (EC 3.2.1.55).	8
4	maltosa-6-fosfato glucosidasa (EC 3.2.1.122); alfa-glucosidasa (EC 3.2.1.20); alfa-galactosidasa (EC 3.2.1.22); 6-fosfo-beta-glucosidasa (EC 3.2.1.86); alfa-glucoronidasa (EC 3.2.1.139).	3
5	quitosanasa (EC 3.2.1.132); beta-manosidasa (EC 3.2.1.25); celulasa (EC 3.2.1.4); glucan 1,3-beta-glucosidasa (EC 3.2.1.58); liqueninasa (EC 3.2.1.73); glucan endo-16,-beta-	3

	glucosidasa (EC 3.2.1.75); manan endo-1,4-beta—manosidasa (EC 3.2.1.78); endo-1,4-beta-xilanasa (EC 3.2.1.8); celulosa 1,4-beta-celobiosidasa (E.C. 3.2.1.91); endo-1,6-beta-galactanasa (EC 3.2.1.-); bet-1,3-mananasa (EC 3.2.1.-); endo-beta-1,4-glucanasa específica a xiloglucano (EC 3.2.1.151)	
8	quitosanasa (EC 3.2.1.132); celulasa (EC 3.2.1.4); liqueninasa (EC 3.2.1.73); endo-1,4-beta-xilanasa (EC 3.2.1.8); exo-oligoxilanasa liberadora del extremo reductor de xilosa (EC 3.2.1.156)	1
9	endoglucanasa (EC 3.2.1.156); celobiohidrolasa (E.C. 3.2.1.91); beta-glucosidasa (EC 3.2.1.21)	1
10	xilanasa (EC 3.2.1.8); endo-1,3-beta-xilanasa (EC 3.2.1.32)	6
11	xilanasa (EC 3.2.1.8)	1
12	endoglucanasa (EC 3.2.1.4); xiloglucan hidrolasa (EC 3.2.1.151); beta-1,3-1,4-glucanasa (EC 3.2.1.73); xiloglucano endotransglicosilasa (EC 2.4.1.207)	1
13	Alfa-amilasa (EC 3.2.1.1); pululanasa (EC 3.2.1.41); ciclomaltodextrin glucanotransferasa (EC 2.4.1.19); ciclomaltodextrinasa (EC 3.2.1.54); trehalosa-6-fosfato hidrolasa (EC 3.2.1.93); oligo-alfa-glucosidasa (EC 3.2.1.10); amilasa maltogénica (EC 3.2.1.133); neopululanasa (EC 3.2.1.135); alfa-glucosidasa (EC 3.2.1.20); alfa-amilasa formadora de maltotetraosa (EC 3.2.1.60); isoamilasa (EC 3.2.1.68); glucodextranasa (EC 3.2.1.70); alfa-amilasa formadora de maltohexaosa (EC 3.2.1.98); enzima ramificadora (EC 2.4.1.18); trehalosa sintasa (EC 5.4.99.16); 4-glucanotransferasa (EC 2.4.1.25); amilasa formadora de maltopentosa (EC 3.2.1.-); amilosacarasa (EC 2.4.1.4); sacarosa fosforilasa (EC 2.4.1.7); maltooligosiltrehalosa trehalohidrolasa (EC 3.2.1.141); isomaltulosa sintasa (EC 5.4.99.11)	7
16	xiloglucan:xiloglucosiltransferasa (EC 2.4.1.207); queratan-sulfato-endo-1,4-beta-galactosidasa (EC 3.2.1.103); glucan-endo-1,3-beta-D- glucosidasa (EC 3.2.1.39); endo-1,3(4)-beta-glucanasa (EC 3.2.1.6); liqueninasa (EC 3.2.1.73); agarasa (EC 3.2.1.81); beta-carragenasa (EC 3.2.1.83); xiloglucanasa (EC 3.2.1.151)	1
18	quitinasa (EC 3.2.1.14); endo-beta-N-acetilglucosaminidasa (EC 3.2.1.96); proteínas no catalíticas; inhibidores de xilanasa;	6

	concanavalina B; narbonina	
19	quitinasa (EC 3.2.1.14)	2
20	beta-hexosaminidasa (EC 3.2.1.52); lacto-N-biosidasa (EC 3.2.1.140); -1,6-N-acetilglucosaminidasa (EC 3.2.1.-);	3
25	lisozima (EC 3.2.1.17)	1
26	beta-mananasa (EC 3.2.1.78); beta-1,3-xilanasa (EC 3.2.1.32)	3
28	poligalacturonasa (EC 3.2.15); exo- poligalacturonasa (EC 3.2.67); exo-poligalacturonosidasa (EC 3.2.82); ramnogalacturonasa (EC 3.2.1.-); endo- xilogalacturonano hidrolasa (EC 3.2.1.-); ramnogalacturonan alfa-L-ramnopirano hidrolasa (EC 3.2.1.40)	5
29	alfa-L-fucosidasa (EC 3.2.1.51)	3
30	glucosilceramidasa (EC 3.2.1.45); beta-1,6-glucanasa (EC 3.2.1.75); beta-xilosidasa (EC 3.2.1.37)	2
31	alfa-glucosidasa (EC 3.2.1.20); alfa-1,3-glucosidasa (EC 3.2.1.84); sacarasa-isomaltasa (EC 3.2.1.48) (EC3.2.1.10); alfa-xilosidasa (EC 3.2.1.-); alfa glucanliasa (EC 4.2.2.13); isomaltosiltransferasa (EC 2.4.1.-)	3
36	alfa-galactosidasa (EC 3.2.1.22); alfa-N-acetilgalactosaminidasa (EC 3.2.1.49); estaquiosa sintasa (EC 2.4.1.67); rafinosa sintasa (EC 2.4.1.82)	2
38	alfa-manosidasa (EC 3.2.1.24); alfa-manosidasa (EC 3.2.1.114);	1
43	beta-xilosidasa (EC 3.2.1.37); beta-1,3-xilosidasa (EC 3.2.1.-); alfa-L-arabinofuranosidasa (EC 3.2.1.55); arabinanasa (EC 3.2.1.99); xilanasa (EC 3.2.1.8); galactan 1,3-beta-galactosidasa (EC 3.2.1.145)	8
48	endoglucanasa (EC 3.2.1.4); quitinasa (EC 3.2.1.14); celobiohidrolasas: se ha reportado que algunas celobiohidrolasas de esta familia actúan a partir de los extremos reductores de la celulosa (EC 3.2.1.-), mientras que se ha reportado que otras operan a partir de los extremos no reductores para liberar celobiosa o celotriosa o celotetraosa (EC 3.2.1.-). Esta familia también contiene celulasas endo-procesivas (EC 3.2.1.-), cuya actividad es difícil de distinguir de aquella de las celobiohidrolasas.	1

51	alfa-L-arabinofuranosidasa (EC 3.2.1.55); endoglucanasa (EC 3.2.1.4)	1
65	trehalasa (EC 3.2.1.28); maltosa fosforilasa (EC 2.4.1.8); trehalosa fosforilasa (EC 2.4.1.64); kojibiosa fosforilasa (EC 2.4.1.230)	4
67	alfa-glucuronidasa (EC 3.2.1.139); xilan alfa-1,2-glucuronidasa (EC 3.2.1.131)	1
73	peptidoglican hidrolasas con especificidad endo-beta-N-acetilglucosaminidasa (EC 3.2.1.-); sólo hay un reporte, no confirmado de la actividad de beta-1,4-N-acetilmuramoilhidrolasa (EC 3.2.1.17)	1
77	amilomaltasa o 4-alfa-glucanotransferasa (EC 2.4.1.25)	1
85	endo-beta-N-acetilglucosaminidasa (EC 3.2.1.96)	1
87	micodextranasa (EC 3.2.1.61); alfa-1,3-glucanasa (EC 3.2.1.59)	3
88	d-4,5 beta glucoronil hidrolasa insaturada (EC 3.2.1.-)	4
94	celobiosa fosforilasa (EC 2.4.1.20); celodextrina fosforilasa (EC 2.4.1.49); quitobiosa fosforilasa (EC 2.4.1.-); beta-1,2-glucan sintasa cíclica (EC 2.4.1.-)	5
95	alfa-1,2-L-fucosidasa (EC 3.2.1.63); alfa-L-fucosidasa (EC 3.2.1.51)	2
105	ramnogalacturonil hidrolasa insaturada (EC 3.2.1.-)	3
106	alfa-L-ramnosidasa (EC 3.2.1.40)	1
112	lacto-N-biosa fosforilasa o galacto-N-biosa fosforilasa (EC 2.4.1.211)	3

TABLA 3: Actividades conocidas de los miembros de la familia polisacárido liasa

Familia polisacárido liasa	Actividades conocidas	Número de dominios predichos en <i>C. phytofermentans</i> .
1	pectato liasa (EC 4.2.2.2); exo- pectato liasa (EC 4.2.2.9); pectin liasa (EC 4.2.2.10)	1
7	alginato liasa (EC 4.2.2.3); L-guluronato liasa (EC	1

	4.2.2.11)	
9	pectato liasa (EC 4.2.2.2); exopoligalacturonato liasa (EC 4.2.2.9)	4
11	pectato liasa (EC 4.2.2.2); exopoligalacturonato liasa (EC 4.2.2.9)	1
12	heparinsulfato liasa (EC 4.2.2.8)	1
15	oligoalginato liasa (EC 4.2.2.8)	1
17	alginato liasa (EC 4.2.2.3)	1

TABLA 4: Actividades conocidas de los miembros de la familia carbohidrato esterasa

Familia carbohidrato esterasa	Actividades conocidas	Número de dominios predichos en <i>C. phytofermentans</i> .
2	acetil xilan esterasa (EC 3.1.1.72)	2
4	acetil xilan esterasa (EC 3.1.1.72); quitin diacetilasa (EC 3.5.1.41); quitooligosacárido deacetilasa (EC 3.5.1.-); peptidoglican GlcNAc deacetilasa (EC 3.5.1.-); peptidoglican ácido N-acetilmurámico deacetilasa (EC 3.5.1.-)	8
8	pectin metilesterasa (EC 4.2.2.3)	1
9	N-acetilglucosamina 6-fosfato deacetilasa (EC 3.5.1.25), N-acetilgalactosamina-6-fosfato deacetilasa (EC 3.5.1.80).	2
12	pectin acetilesterasa (EC 3.1.1.-); ramnogalacturonan acetilesterasa (EC 3.1.1.-); acetil xilanesterasa (EC 3.1.1.72)	1
15	4-O-metil-glucoronilesterasa (EC 3.1.1.-)	1

TABLA 5: Actividades conocidas de los miembros de la familia de módulos de unión a carbohidratos

Familia CBM	Actividades conocidas	Número de dominios predichos en <i>C. phytofermentans</i> .
2	Módulos de aprox. 100 residuos encontrados en muchas enzimas bacterianas con actividades de unión a celulosa putativa, quitina y/o xilano.	1
3	Módulos de aprox. 150 residuos encontrados en enzimas bacterianas. La función de unión a celulosa ha sido demostrada en muchos casos. En una instancia la unión a quitina se ha reportado.	5
4	Módulos de aprox. 150 residuos encontrados en enzimas bacterianas. La función de estos módulos ha sido demostrada con xilano, -1,3-glucano, -1,4-1,4-glucano, -1,6-glucano y celulosa amorfa pero no con celulosa cristalina	4
5	Módulos de aprox. 60 residuos encontrados en enzimas bacterianas. Relacionados de forma distante con la familia CBM12.	1
6	Módulos de aprox. 120 residuos. La función de unión a celulosa ha sido demostrada en un caso sobre celulosa amorfa y xilano. Algunos de estos módulos también unen -1,3—glucano.	1
12	Módulos de aprox. 40-60 residuos. La mayoría de estos módulos se encuentra entre quitinasas en donde la función es unión de quitina. Relacionados de forma distante con la familia CBM5.	2
13	Módulos de aprox. 150 residuos que frecuentemente aparecen como una repetición interna triple, una excepción incluye, xilanasas II de <i>Actinomadura</i> sp. FC7 (GenBank U08894). Estos módulos fueron identificados primero en varias lectinas de plantas tales como ricino o aglutinina de <i>Ricinus communis</i> que se unen a residuos de galactosa. La estructura tridimensional de una lectina vegetal ha sido determinada y muestra una simetría pseudo-triple de acuerdo con la repetición triple de la secuencia observada. Estos módulos se han encontrado desde entonces en una cantidad de proteínas diferentes de varias funciones incluyendo glicósido hidrolasas y glicosiltransferasas. Mientras que en las lectinas vegetales este módulo se une a manosa, la unión a xilano se ha demostrado en la xilanasas A y la arabinofuranosidasa B de <i>Streptomyces lividans</i> . Se ha mostrado la unión a GalNAc	1

	para el correspondiente módulo de GalNAc transferasa 4. Para las otras proteínas, la especificidad de unión de estos módulos no se ha establecido. La simetría pseudo-triple del módulo CBM13 se ha confirmado ahora en la estructura 3D de la xilanasa intacta de dos dominios de <i>Streptomyces olivaceoviridis</i> .	
22	Una función de unión a xilano se ha demostrado en varios casos y la afinidad con glucanos -1,3/-1,4- mezclados en uno. En varios casos se ha observado también un efecto termoestabilizador.	1
32	Se ha demostrado unión a galactosa y lactosa para el módulo de <i>Micromonospora viridifaciens</i> sialidase (PMID: 16239725); se ha mostrado unión a ácido poligalacturónico para un miembro de <i>Yersinia</i> (PMID: 17292916); se ha mostrado la unión a LacNAc (-D-galactosil-1,4-D-N-acetilglucosamina) para una N-acetilglucosaminidasa de <i>Clostridium perfringens</i> (PMID: 16990278)	5
35	Módulos de aprox. 130 residuos. Un módulo que se conserva en tres enzimas degradadoras de xilano de <i>Cellvibrio</i> se une al xilano y la interacción es dependiente de calcio, mientras que un módulo de una <i>Cellvibrio mananasa</i> se une a mananos solubles decoradas y manooligosacáridos. Un módulo en una galactan 1,3-galactosidasa de <i>Phanerochaete chrysosporium</i> se une a galactano.	4
36	Módulos de aprox. 130 residuos. Un módulo que se conserva en tres enzimas degradadoras de xilano de <i>Cellvibrio</i> se une al xilano y la interacción es dependiente de calcio, mientras que un módulo de una <i>Cellvibrio mananasa</i> se une a mananos solubles decoradas y manooligosacáridos. Un módulo en una galactan 1,3-galactosidasa de <i>Phanerochaete chrysosporium</i> se une a galactano.	1
41	Módulos de aprox. 100 residuos encontrados primero en pululanases bacterianas. El módulo N terminal de <i>Thermotoga marítima</i> Pul13 ha mostrado unirse a los glucanos amilosa, amilopectina, pululano, y fragmentos de oligosacáridos derivados de estos polisacáridos.	1
46	Módulos de aprox. 100 residuos, encontrados en el extremo C de varias celulasas GH5. La función de unión a celulosa se demostró en un caso.	1
48	Módulos de aprox. 100 residuos con función de unión a	2

	glicógeno, anexados a módulos GH13. También encontrados en la subunidad beta (unión a glicógeno) de proteínas quinasa activadas por AMP (AMPK)	
50	Módulos de aprox. 50 residuos que se encuentran unidos a varias enzimas de las familias GH18, GH19, GH23, GH24, GH25 Y GH73, i.e. enzimas que fragmentan o quitina o peptidoglicano. La unión a quitopentaosa demostró en el caso de <i>Pteris ryukyuensis</i> quitinasa A [Ohnuma T et al.; PMID: 18083709]. Los módulos de CBM50 también se encuentran en una multitud de otras enzimas que se dirigen al peptidoglicano tales como peptidasas y amidasas.	4

Algunas realizaciones incluyen genes que codifican hidrolasas se muestran en la Tabla 6. El número JGI se refiere a la etiqueta NCBI del locus en el registro de GenBank.

TABLA 6: Hidrolasas predichas en *C. phytofermentans*.

JGI No.	GH	Arquitectura del módulo
Cphy0191	GH43	
Cphy0203	GH105	
Cphy0218	GH31	
Cphy0220	GH3	
Cphy0288	GH88	
Cphy0430	GH94	
Cphy0530	GH2	
Cphy0531	GH43	
Cphy0607	GH20	
Cphy0662	GH3	
Cphy0666	GH106	
Cphy0694	GH94	
Cphy0699	GH3	
Cphy0711	GH2	
Cphy0769	GH4	
Cphy0776	GH88	
Cphy0857	GH94	
Cphy0858	GH30	
Cphy0874	GH95	

Cphy0875	GH43	
Cphy0934	GH88	
Cphy1019	GH65	
Cphy1071	GH26	CBM35-GH26-CBM3
Cphy1125	GH3	
Cphy1163	GH5	
Cphy1169	GH51	
Cphy1308	GH87	
Cphy1395	GH95	
Cphy1435	GH19	
Cphy1510	GH10	
Cphy1596	GH3	
Cphy1612		
Cphy1640	GH12	
Cphy1652	GH18	
Cphy1688	GH*	
Cphy1711	GH28	
Cphy1714	GH85	GH85-CBM32
Cphy1720	GH38	
Cphy1750	GH105	
Cphy1775	GH*	SLH-GH*-CBM32-CBM32
Cphy1799	GH18	CBM12-GH18
Cphy1800	GH18	GH18-CBM12
Cphy1815	GH18	GH18-LRR
Cphy1873	GH87	CBM35-CBM6-GH87
Cphy1874	GH65	
Cphy1877	GH31	
Cphy1882	GH87	GH87-SORT
Cphy1888		
Cphy1919	GH105	
Cphy1929	GH94	
Cphy1934	GH13	
Cphy1936	GH36	
Cphy1937	GH1	
Cphy1943	GH19	CBM5-GH19
Cphy2025	GH2	

Cphy2028	GH43	
Cphy2058	GH5	
Cphy2105	GH11	
Cphy2108	GH10	CBM22-GH10-SORT
Cphy2128	GH26	CBM35-GH26-X2-X2-CBM3
Cphy2190	GH29	
Cphy2276	GH26	CBM35-GH26
Cphy2304	GH13	CBM41-CMB48-GH13-SORT
Cphy2331	GH13	CBM48-GH13
Cphy2332	GH3	
Cphy2341	GH13	
Cphy2342	GH13	
Cphy2344	GH13	
Cphy2349	GH77	
Cphy2350	GH13	
Cphy2567	GH28	
Cphy2572	GH18	
Cphy2632	GH43	
Cphy2736	GH28	
Cphy2848	GH4	
Cphy2919		
Cphy3009	GH3	
Cphy3010	GH10	
Cphy3011	GH43	
Cphy3023	GH29	
Cphy3028	GH29	
Cphy3029	GH88	
Cphy3056	GH36	
Cphy3081	GH2	
Cphy3109	GH25	
Cphy3158	GH67	
Cphy3160	GH2	
Cphy3202	GH5	Gh5-x2-cbm46-cbm2
Cphy3207	GH8	
Cphy3217	GH28	
Cphy3239	GH20	

Cphy3310	GH28	
Cphy3313	GH65	
Cphy3314	GH65	
Cphy3329	GH3	
Cphy3367	GH9	GH9-CBM3-X2-X2-CBM3
Cphy3368	GH48	GH48-X2-CBM3
Cphy3388	GH16	GH16-CBM4-CBM4-CBM4-CBM4

En algunas realizaciones, las enzimas que degradan polisacáridos pueden incluir enzimas que degradan celulosa, es decir, celulasas. Algunas celulasas, incluyendo las endocelulasas (EC 3.2.1.4) y exocelulasas (EC 3.2.1.91) hidrolizan enlaces beta-1,4-glicosídicos.

Ejemplos de endocelulasas predichas en *C. phytofermentans* pueden incluir genes con la familia GH5, tales como, Cphy3368; Cphyl 163, y Cphy2058; la familia GH8, tal como Cphy3207; y la familia GH9, tal como Cphy3367. Ejemplos de exocelulasas en *C. phytofermentans* pueden incluir genes dentro de la familia GH48, tales como Cphy3368. Algunas exocelulasas hidrolizan polisacáridos para producir de 2 a 4 oligosacáridos unitarios de glucosa, lo que resulta en disacáridos celodextrinas (celobiosa), trisacáridos (celotriosa), o tetrasacáridos (celotetraosa). Los miembros de las familias GH5, GH9 y GH48 pueden tener tanto actividad exo como endocelulasa.

En algunas realizaciones, las enzimas que degradan polisacáridos pueden incluir enzimas que pueden tener la habilidad para degradar hemicelulosa, es decir, hemicelulasas (Leschine, S.B. en Handbook on Clostridia (ed. Dürre, P.) (CRC Press, Boca Raton, 2005)). La hemicelulosa puede ser un componente principal de la biomasa de las plantas y puede contener una mezcla de pentosas y hexosas, por ejemplo, D-xilopiranosas, L-arabinofuranosa, D-manopiranosas, D-glucopiranosas, D-galactopiranosas, ácido D-glucopiranosilurónico y otros azúcares (Aspinall, G. O. The Biochemistry of Plants 473, 1980; Han, J. S. & Rowell, J. S. en Paper and composites from agro-based resources 83, 1997). En ciertas realizaciones, las hemicelulasas predichas identificadas en *C. phytofermentans* pueden incluir enzimas activas sobre la estructura principal de la hemicelulosa, por ejemplo, endo-beta-1,4-D-xilanasas (EC 3.2.1.8), tales como los miembros de las familias GH5, GH10, GH11,

y GH43; 1,4-beta-D-xilósido xilohidrolasa (EC 3.2.1.37), tales como los miembros de las familias GH30, GH43, y GH3; y beta-mananasa (EC 3.2.1.78), tales como los miembros de la familia GH26. (Ver Tabla 6).

En algunas realizaciones, las hemicelulasas predichas identificadas en *C. phytofermentans* pueden incluir enzimas activas en los grupos laterales y sustituyentes de hemicelulosa, por ejemplo, alfa-L-arabinofuranosidasa (EC 3.2.1.55), tales como los miembros de las familias GH3, GH43, y GH51; alfa-xilosidasa, tales como los miembros de la familia GH31; alfa-fucosidasa (EC 3.2.1.51), tales como los miembros de las familias GH95 y GH29; galactosidasa, tales como los miembros de las familias GH1, GH2, GH4, GH36, GH43; y acetil-xilanesterasa (EC 3.1.1.72), tales como CE2 y CE4. (Ver Tabla 6).

En algunas realizaciones, las enzimas que degradan polisacáridos pueden incluir enzimas que tienen la habilidad de degradar pectina, es decir, pectinasas. En las paredes celulares de las plantas, la red entrecruzada de celulosa se puede incrustar en una matriz de pectinas que puede estar entrecruzada de forma covalente a xiloglicanos y ciertas proteínas estructurales. La pectina puede comprender homogalacturonano (HG) o ramnogalacturonano (RH).

En otras realizaciones, las pectinasas identificadas en *C. phytofermentans* pueden hidrolizar HG. El HG puede estar compuesto de unidades de ácido D-galacturónico (D-galA), que pueden estar acetiladas y metiladas. Las enzimas que hidrolizan HG pueden incluir, por ejemplo, 1,4-alfa-D galacturonanliasa (EC 4.2.2.2), tales como los miembros de las familias PL1, PL9, y PL11; glucuronil hidrolasa, tales como los miembros de las familias GH88 y GH105; pectina acetilesterasa tales como los miembros de la familia CE 12; y pectina metilesterasa, tales como los miembros de la familia CE8. (Ver Tabla 6).

Incluso en algunas realizaciones, las pectinasas identificadas en *C. phytofermentans* pueden hidrolizar RH. Los RH pueden tener una estructura principal compuesta de residuos alternantes de 1,2-alfa-L-ramnosa (L-Rha) y 1,4-alfa-D-galacturónico (Lau, J. M., McNeil M., Darvill A. G. & Albersheim P. Structure of the backbone of ramnogalacturonan I, a pectic polysaccharide in the primary cell walls of

plants. Carbohydrate research 137, 111 (1985)). Los residuos de ramnosa de las estructuras principales pueden tener galactano, arabinano, o arabinogalactano enlazados a C4 como cadenas laterales. Las enzimas que hidrolizan HG pueden incluir, por ejemplo, endo- ramnogalacturonasa, tal como los miembros de la familia GH28; y ramnogalacturonano liasa, tales como los miembros de la familia PL11. (Ver Tabla 6).

Algunas realizaciones incluyen enzimas que pueden hidrolizar almidón. *C. phytofermentans* puede degradar almidón y quitina ((Warnick, T. A., Methe, B. A. & Leschine, S. B. *Clostridium phytofermentans* sp. nov., a cellulolytic mesophile from forest soil. Int. J. Syst. Evol. Microbiol. 52, 1155-1160 (2002); Leschine, S. B. en Handbook on Clostridia (ed Dürre, P.) (CRC Press, Boca Raton, 2005); Reguera, G. & Leschine, S. B. Chitin degradation by cellulolytic anaerobes and facultative aerobes from soils and sediments. FEMS Microbiol. Lett. 204, 367-374 (2001)). Las enzimas que hidrolizan almidón incluyen alfa-amilasa, glucoamilasa, beta-amilasa, exo-alfa-1,4-glucanasa, y pululanasa. Ejemplos de las enzimas predichas identificadas en *C. phytofermentans* involucradas en la hidrólisis del almidón incluyen los miembros de la familia GH13. (Ver Tabla 6).

En otras realizaciones, las hidrolasas pueden incluir enzimas que hidrolizan quitina. Ejemplos de enzimas que pueden hidrolizan quitina incluyen los miembros de la familia GH18 y GH19. (Ver Tabla 6).

Incluso en algunas realizaciones, las hidrolasas pueden incluir enzimas que hidrolizan liquen, es decir, liquenasa, por ejemplo, los miembros de la familia GH16, tales como Cphy3388.

En algunas realizaciones, las hidrolasas pueden incluir miembros de la familia CBM. Sin desear vincularse con ninguna teoría, los dominios de CBM pueden funcionar para localizar complejos enzimáticos para sustratos particulares. Ejemplos de familias de CBM predichas identificadas en *C. phytofermentans* que pueden unir celulosa incluyen miembros de las familias CBM2, CBM3, CBM4, CBM6, y CBM46. Ejemplos de familias de CBM predichas identificadas en *C. phytofermentans* que pueden unir xilano incluyen miembros de las familias CBM2, CBM4, CBM6, CBM13,

CBM22, CBM35, y CBM36. (Ver tabla 6). En otras realizaciones, los miembros de la familia del dominio CBM pueden funcionar para estabilizar un complejo enzimático.

Algunas realizaciones incluyen polinucleótidos que codifican al menos una hidrolasa predicha identificada en *C. phytofermentans*.

Transportadores del Casete de Unión a ATP

Algunas realizaciones descritas aquí se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican transportadores del casete de unión a ATP (transportadores ABC). Algunas realizaciones se relacionan con métodos para producir combustible utilizando estos polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican transportadores ABC. Las ventajas de utilizar ácidos nucleicos que codifican transportadores ABC incluyen incrementar la capacidad de los organismos transformados para transportar compuestos dentro del organismo y utilizar dichos compuestos en las vías bioquímicas para producir combustible, y por consiguiente mejorar la producción de combustible. Ejemplos de dichos compuestos incluyen los productos de la hidrólisis de polímeros.

Las proteínas del transportador ABC utilizan la hidrólisis de ATP para transportar una amplia variedad de sustancias a lo largo de la membrana plasmática. Dichas sustancias pueden incluir azúcares y aminoácidos. Los transportadores ABC se pueden identificar utilizando los métodos descritos aquí y los métodos bien conocidos en el arte. Los transportadores ABC comprenden al menos dos tipos de dominios, dominios transmembranales y dominios de unión a nucleótidos (e.g. ATP). Algunos transportadores ABC también incluyen un dominio de unión a solutos que asiste en la mediación del transporte de solutos. Estos dominios pueden estar presentes en la misma cadena de polipéptido o en múltiples cadenas de polipéptidos. Algunos miembros de la familia de transportadores ABC comprenden el dominio ABC tran (pfam00005): Más miembros de la familia de transportadores ABC pueden comprender cuatro dominios dentro de dos mitades simétricas que están unidas

mediante una región larga cargada y un segmento altamente hidrofóbico (Hyde et al, Nature, 346:362-365 (1990); Luciani et al, Genomics, 21 : 150-159 (1994)).

En realizaciones más ejemplares, los casetes de polinucleótidos, casetes de expresión, vectores de expresión, y organismos que comprenden transportadores ABC se identifican en *C. phytofermentans*. Dichos agrupamientos de genes se pueden identificar utilizando los métodos descritos aquí y los métodos bien conocidos en el arte. En algunas realizaciones, los genes y los agrupamientos de genes se pueden identificar por el grado de homología entre los agrupamientos de grupos ortólogos de proteínas (COG, por sus siglas en inglés). Dichos genes y agrupamientos de genes se pueden incluir en casetes o expresar juntos. Los ejemplos pueden incluir los transportadores ABC predichos y los dominios de los transportadores de ABC mostrados en la Tabla 7. La columna “No.” representa agrupamientos putativos. Los dominios del transportador ABC pueden incluir dominios de transducción de señal.

TABLA 7: Transportadores ABC predichos y otras proteínas/dominios en *C. phytofermentans*.

No.	JGI No.	Localización	COG	Descripción de COG
1	Cphy0110	147354-148388	CDA1	Xilanasa/quitina deacetilasa predicha
	Cphy0111	148828-149403	Pth	Peptidil-ARNt hidrolasa
	Cphy0112	149444-152983	Mfd	Factor de acoplamiento de reparación de la transcripción (helicasa de la superfamilia II)
	Cphy0113	153051-154184		No hay coincidencia COG
	Cphy0114	154273-155049	CcmA	Sistema de transporte multifármaco tipo ABC
	Cphy0115	155051-156268	NatB	Bomba de entrada de Na ⁺ de tipo ABC, componente permeasa
	Cphy0116	156368-156760	ArsR	Reguladores transcripcionales predichos
	Cphy0117	156810-157454	COG0490	Proteína de unión a ligando, reguladora putativa relacionada con los dominios

				C terminales de los canales de K ⁺
	Cphy0118	158002-159138	OpuBA	Sistemas de transporte prolina/glicina betaína de tipo ABC
	Cphy0119	159135-160706	OpuBC	(lipo)proteína de unión a glicina betaina/colina periplásmica de un sistema de transporte tipo ABC (proteína de unión osmoprotectora)
2	Cphy0191	236617-238155	XynB	Beta-xilosidasa
	Cphy0192	238298-240091	SalX	Sistema de transporte de péptidos antimicrobianos de tipo ABC, componente ATPasa
	Cphy0193	240094-242037	SalX	Sistema de transporte de péptidos antimicrobianos de tipo ABC, componente ATPasa
	Cphy0194	242049-242735		No hay coincidencia COG
	Cphy0195	242919-244052	BaeS	Histidina quinasa de transducción de señal
	Cphy0196	244049-244729	OmpR	Reguladores de respuesta que consisten de un receptor tipo CheY y un dominio de unión a ADN de hélice alada
3	Cphy0288	358277-359431	COG1331	Proteína altamente conservada que tiene un dominio tioredoxina
	Cphy0289	359443-360273	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0290	360288-361181	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy0291	361234-362583	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0292	362867-364387	LytS	Regulador putativo de la autólisis celular
	Cphy0293	364406-366004	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
4	Cphy0337	423685-425433	MdIB	Sistema de transporte de

				multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy0338	425701..427869	Tex	Proteína accesoria transcripcional
	Cphy0339	428277..429278	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy0340	429392..430300	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0341	430405..432018	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0342	432371..434707	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy0343	434803..437214		No hay coincidencia COG
5	Cphy0430	547250..549745	COG3459	Celobiosa fosforilasa
	Cphy0431	550144..551712	DpdA	Sistema de transporte de dipéptidos de tipo ABC, componente periplásmico
	Cphy0432	551801..552742	DppB	Sistemas de transporte de dipéptido/oligopéptido/níquel de tipo ABC, componentes de permeasa
	Cphy0433	552764..553552	DppC	Sistemas de transporte de dipéptido/oligopéptido/níquel tipo ABC, componentes de permeasa
	Cphy0434	553633..555201	COG1123	Componentes ATPasa de varios sistemas de transporte de de tipo ABC, contiene ATPasa duplicada
6	Cphy0484	612448..613455	PurR	Reguladores transcripcionales
	Cphy0485	613468..614448	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy0486	614460..615353	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0487	615427..617103	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0488	617267..619552		No hay coincidencia COG
	Cphy0489	619572..621440		No hay coincidencia COG
	Cphy0490	621481..622278	FabG	Deshidrogenasas con

				diferentes especificidades (relacionadas con deshidrogenasas de cadena corta)
	Cphy0491	622395..623030	RpiB	Ribosa 5-fosfato isomerasa RpiB
	Cphy0492	623046..627716	COG3858	Glicosil hidrolasa predicha
	Cphy0493	627945..630125	Tar	Proteína quimiotáctica aceptadora de metilo
	Cphy0494	630357..631754	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0495	631901..633715	COG2972	Proteína de transducción de señal predicha con un domino ATPasa C terminal
	Cphy0496	633810..635441	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy0497	635568..638291	Markorf3686	Glicosiltransferasa, familia 51: proteína de unión a penicilina, precursor de transpeptidasa
	Cphy0498	638612..641344	MrcA	Proteína de unoin carboxipeptidasa/penicilina de membrana
7	Cphy0525	665552..667120	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy0526	667117..668901	COG2972	Proteína de transducción de señal predicha con un domino ATPasa C terminal
	Cphy0527	669186..670076	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy0528	670088..670933	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0529	671066..672460	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0530	672748..674523	Lacz	Beta-galactosidasa/beta-glucuronidasa
	Cphy0531	674706..676121	XynB	Beta-xilosidasa
8	Cphy0615	800570..801004		No hay coincidencia COG
	Cphy0616	801007..801984		No hay coincidencia COG
	Cphy0617	802351..804156	LytS	Regulador putativo de la autólisis celular
	Cphy0618	804146..804922	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio

				de unión a ADN de tipo AraC.
	Cphy0619	805147..806538	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0620	806729..807646	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy0621	807636..808505	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0622	808851..809189	MutS	ATPasa reparación de malapareamiento (familia MutS)
	Cphy0623	809475..810185		No hay coincidencia COG
	Cphy0624	810523..812793	XynA	Beta-1,4-xilanasa
	Cphy0625	813011..813778	COG0627	Esterasa predicha
	Cphy0626	813769..814512	COG0627	Esterasa predicha
9	Cphy0662	862473..864716	BglX	Glicosidasas relacionadas con beta glucosidasa
	Cphy0663	864876..866546	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0664	866795..867769	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy0665	867787..868674	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0666	868759..871452		No hay coincidencia COG
	Cphy0667	871581..872183	AcrR	Regulador transcripcional
	Cphy0668	873253..873825		No hay coincidencia COG
	Cphy0669	873964..874098	GalE	UDP-glucosa 4-epimerasa
10	Cphy0694	896904..900245	COG3459	Celobiosa fosforilasa
	Cphy0695	900644..901510	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy0696	901526..902437	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0697	902506..904050	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0698	904124..905161	PurR	Reguladores transcripcionales
	Cphy0699	905298..907529	BglX	Glicosidasas relacionadas con beta glucosidasa
11	Cphy0764	9869994..988166	Med	Sistema de transporte tipo ABC no caracterizado, componente periplásmico/lipoproteína de

				superficie
	Cphy0765	988605..990140	COG3845	Sistemas de transporte tipo ABC no caracterizados, componentes de ATPasa
	Cphy0766	990133..991233	COG4603	Sistemas de transporte tipo ABC no caracterizados, componentes de permeasa
	Cphy0767	991233..992192	COG1079	Sistemas de transporte tipo ABC no caracterizados, componentes de permeasa
	Cphy0768	992375..993160	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy0769	993285..994607	CelF	Alfa-galactosidasas/6-fosfobetaglucosidasas, familia 4 de glicosil hidrolasas
12	Cphy0770	994704..995759		No hay coincidencia COG
	Cphy0771	995811..996878	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy0772	996904..998826	COG2972	Proteína de transducción de señal predicha con un domino ATPasa C terminal
	Cphy0773	999105..1000040	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy0774	1000095.. 1000940	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0775	1001005..1002576	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0776	1002623..1003717	COG1331	Proteína altamente conservada que tiene un dominio tioredoxina
13	Cphy0854	1089598..1090047	MarR	Reguladores transcripcionales
	Cphy0855	1090050..1092317	MdIB	Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy0856	1092314..1094194	MdIB	Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy0857	1094446..1097148	COG3459	Celobiosa fosforilasa

	Cphy0858	1097657..1098994	COG5520	O-glicosilhidrolasa
	Cphy0859	1099142..1099768		No hay coincidencia COG
	Cphy0860	1099884..1100825	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy0861	1100837..1101745	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0862	1101768..1103474	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0863	1103665..1105464	COG2972	Proteína de transducción de señal predicha con un dominio ATPasa C terminal
	Cphy0864	1105480..1106991	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
14	Cphy0928	1182272..1184572	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy0929	1184776..1185768	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy0930	1185782..1186666	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy0931	1186734..1188368	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy0932	1188484..1190736		No hay coincidencia COG
	Cphy0933	1190798..1191850		No hay coincidencia COG
	Cphy0934	1191923..1193098		No hay coincidencia COG
15	Cphy1010	1276704..1277816	COG1476	Reguladores transcripcionales predichos
	Cphy1011	1278019..1278399		No hay coincidencia COG
	Cphy1012	1278635..1279747	PurR	Reguladores transcripcionales
	Cphy1013	1280007..1281491	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1014	1281508..1282416	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy1015	1282416..1283258	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
16	Cphy1071	1354865..1357051	ManB	Beta-mananasa
	Cphy1074	1358682..1360004	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1075	1360064..1360906	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy1076	1360906..1361769	UgpE	Sistema de transporte de azúcar tipo ABC, componente

				de permeasa
	Cphy1077	1361911..1362948	PurR	Reguladores transcripcionales
17	Cphy1118	1410194..1411819	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1119	1411890..1412846	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy1120	1412857..1413726	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy1121	1413748..1414641	COG2103	Fosfato isomerasa de azúcar predicha
	Cphy1122	1414707..1415564	RpiR	Reguladores transcripcionales
	Cphy1123	1415584..1416768	COG2377	Chaperona molecular predicha relacionada de forma distante con las metaloproteinasas de pliegue HSP70
	Cphy1124	1416904..1418049	COG3876	Proteína no caracterizada conservada en bacterias
	Cphy1125	1418251..1419804	BglX	Glicosidasas relacionadas con beta glucosidasa
18	Cphy1390	1729845..1730771	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy1391	1730803..1731633	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy1392	1731696..1733045	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1393	1733168..1734928	COG2972	Proteína de transducción de señal predicha con un dominio ATPasa C terminal
	Cphy1394	1734977..1736548	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy1395	1736693..1738969		No hay coincidencia COG
19	Cphy1679	2059856 - 2060530	OmpR	Reguladores de respuesta que consisten de un receptor tipo CheY y un dominio de unión a ADN de hélice alada
	Cphy1680	2060527 - 2061597	BaeS	Histidina quinasa de transducción de señal
	Cpyh1681	2061728 - 2063485	MdiB	Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy1682	2063482 - 2065323	MdiB	

Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y

				ATPasa
	Cphy1683	2065548 - 2066213	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy1685	2067701 - 2068645		No hay coincidencia COG
	Cphy1686	2068969 - 2070048	COG2706	Ciclase 3-carboximuconato
	Cphy1687	2070592 - 2071602	CDA1	Xilanasas/quitinasa deacetilasa predicha
	Cphy1688	2071762 - 2073555	COG3292	dominio de sensor de unión a ligando periplásmico predicho
20	Cphy1706	2092890 - 2095226	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy1707	2095571 - 2097463	DraC	ADP-ribosilglicohidrolasa
	Cphy1708	2097665 - 2099221	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1709	2099418 - 2100341	LplB	Sistema de transporte de polisacárido tipo ABC, componente de permeasa
	Cphy1710	2100360 - 2101262	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy1711	2101420 - 2103282	PGU1	Endopigalactonurasa
	Cphy1712	2103332 - 2104336	Aes	Esterasa/lipasa
	Cphy1713	2104536 - 2105900		No hay coincidencia COG
	Cphy1714	2106148 - 2109003	COG4724	Endo-beta-N-acetilglucosaminidasa D
	Cphy1715	2109418 - 2110368	LplB	Sistema de transporte de polisacárido tipo ABC, componente de permeasa
	Cphy1716	2110380 - 2111312	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy1717	2111371 - 2112876	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1720			Familia 38 de glicósido hidrolasa
	Cphy1719			proteína hipotética
	Cphy1718			Glicosidasa relacionada con PH1107
21	Cphy1873	2302530 - 2305805	COG1572	Proteína conservada no caracterizada
	Cphy1874	2305913 - 2308174	ATH1	Hidrolasas trihalosa y maltosa (posibles fosforilasas)
	Cphy1875	2308280 - 2308927	COG0637	Fosfatasa/fosfohexomutasa predicha
	Cphy1876	2309264 - 2310301	PurR	Reguladores transcripcionales
	Cphy1877	2310355 - 2312748	COG1501	Alfa-glucosidasas, familia 31 de glicosil hidrolasas
	Cphy1878	2312810 - 2314963	Tar	Proteína quimiotáctica aceptadora de metilo
	Cphy1879	2315244 - 2316557	MalE	
	Cphy1880	2316885 - 2318273	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy1881	2318273 - 2319553	MalG	Sistema de transporte de maltosa tipo ABC, componente de permeasa

	Cphy1882	2319649 - 2323881		No hay coincidencia COG
	Cphy1883	2324145 - 2325200	PurR	Reguladores transcripcionales
22	Cphy1915	2362169 - 2364496	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy1916	2364818 - 2365780	LplB	Sistema de transporte de polisacárido tipo ABC, componente de permeasa
	Cphy1917	2365793 - 2366710	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy1918	2366768 - 2368441	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1919	2369079 - 2370116	COG24225	Glucoronil hidrolasa insaturada predicha, involucrada en la regulación de las propiedades de la superficie bacterial y proteínas relacionadas
23	Cphy2187	2700406 - 2702664	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy2188	2702785 - 2703105		No hay coincidencia COG
	Cphy2189	2703144 - 2704676	CcdA	Proteína de la biogénesis de citocromo c
	Cphy2190	2704764 - 2706071	COG3669	Alfa-L-fucosidasa
	Cphy2191	2706212 - 2707117	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2192	2707130 - 2708053	LplB	Sistema de transporte de polisacárido tipo ABC, componente de permeasa
	Cphy2193	2708129 - 2709802	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
24	Cphy2265	2791741 - 2793189	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2266	2793236 - 2795134		No hay coincidencia COG
	Cphy2267	2795272 - 2796264	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2268	2696278 - 2797195	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy2269	2797161 - 2799770		No hay coincidencia COG
	Cphy2272	2801915 - 2802787	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2273	2802804 - 2803739	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy2274	2803758 - 2806796	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2275	2806789 - 2807427	COG5578	Proteína integral de membrana predicha
	Cphy2276	2807484 - 2809082	ManB	Beta-mananasa
	Cphy2277	2809660 - 2810001	markof1779	proteína hipotética

	Cphy2277	2809660 - 2810001	markof1779	proteína hipotética
	Cphy2278	2810437 - 2811483	PurR	Reguladores transcripcionales
25			PulA	Vía de secreción Tipo II, pululanasa PulA y glicosidasas relacionadas
	Cphy2304	2833695 - 2840126		No hay coincidencia COG
	Cphy2305	2840111 - 2840968		
	Cphy2306	2840949 - 2841791	MalG	Sistema de transporte de maltosa tipo ABC, componente de permeasa
	Cphy2307	2841802 - 2843130	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy2308	2843299 - 2844621	MalE	
26	Cphy2338	2882581 - 2883465	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy2339	2883550 - 2884353	LoIE	Sistema de transporte Tipo ABC, involucrado en la liberación de lipoproteínas, componente de permeasa
	Cphy2340	2884437 - 2885246	LoIE	Sistema de transporte Tipo ABC, involucrado en la liberación de lipoproteínas, componente de permeasa
	Cphy2341	2885548 - 2887212	AmyA	Glicosidasas
	Cphy2342	2887378 - 2888994	AmyA	Glicosidasas
	Cphy2343	2889298 - 2891253	GDB1	Enzima degradante de glicógeno
	Cphy2344	2891543 - 2893270	AmyA	Glicosidasas
	Cphy2345	2893664 - 2894530	MalG	Sistema de transporte de maltosa tipo ABC, componente de permeasa
	Cphy2346	2894530 - 2895924	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy2567	3133582 - 3135006	PGU1	Endopigalactonurasa
	Cphy2568	3135310 - 3135972	COG1600	Proteína Fe-S no caracterizada
27	Cphy2569	3136132 - 3137703	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2570	3137758 - 3138693	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2571	3138707 - 3139672	LplB	Sistema de transporte de polisacárido tipo ABC, componente de permeasa
	Cphy2572	3140149 - 3141216	ChiA	Chitinasa
	Cphy2731	3321178 - 3322029	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2732	3322042 - 3322992	LplB	Sistema de transporte de polisacárido tipo ABC, componente de permeasa
28	Cphy2733	3323144 - 3324745	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2734	3325192 - 3327486	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC

	Cphy2735	3327576 - 3328295	TesA	Lisofosfolipasa L1 y esterasas relacionadas
	Cphy2736	3328548 - 3330104	PGU1	Endopigalactonurasa
29	Cphy3009	3672467 - 3674620	BglX	Glicosidasas relacionadas con beta-glucosidasa
	Cphy3010	3674634 - 3675599	XynA	Beta 1.4-xilanasas
	Cphy3011	3676460 - 3678076	XynB	Beta xilosidasa
	Cphy3012	3678276 - 3679193	SufB	Sistema de transporte Tipo ABC, involucrado en el ensamblaje agrupamiento Fe-S, componente permeasa
	Cphy3013	3679197 - 3679919	SufC	Sistema de transporte Tipo ABC, involucrado en el ensamblaje agrupamiento Fe-S, componente ATPasa
30	Cphy3066	3748299 - 3750080	MdiB	Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy3067	3750058 - 3751908	MdiB	Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy3068	3751920 - 3752993		No hay coincidencia conocida
	Cphy3069	3753565 - 3754962	CDA1	Xilanasas/quitina deacetilasa predicha
	Cphy3102	3792846 - 3793784	COG1216	Glicosiltransferasas predichas
31	Cphy3103	3793820 - 3794788	COG1215	Glicosiltransferasas, probablemente involucradas en la biogénesis de la pared celular
	Cphy3104	3794965 - 3796023	COG1123	Componentes ATPasa de varios sistemas de transporte Tipo ABC, contienen ATPasa duplicada
	Cphy3105	3796156 - 3797211	DppD	Sistema de transporte dipéptido/oligopéptido/niquel Tipo ABC, componente de ATPasa
	Cphy3106	3797424 - 3798398	DppC	Sistema de transporte dipéptido/oligopéptido/niquel Tipo ABC, componente de permeasa
	Cphy3107	3798399 - 3799337	DppB	Sistema de transporte dipéptido/oligopéptido/niquel Tipo ABC, componentes de permeasa
	Cphy3108	3799364 - 3801211	OppA	Sistema de transporte de oligopéptido, componente periplásmico
	Cphy3109	3802071 - 3803957	Acm	Lizozima M1 (1,4-beta-N-acetilmuramidasa)
	Cphy3207	3910130 - 3911275	CelA	Endoglucanasa Y
	Cphy3208	3911468 - 3912373	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
32	Cphy3209	3912465 - 3913424	LplB	Sistema de transporte de polisacárido tipo ABC, componente de permeasa
	Cphy3239	3943604..3945511	Chb	N-acetil-beta-hexosaminidasa
33	Cphy3240	3945691..3947259	UgpB	Sistema de transporte de

				azúcar tipo ABC, componente periplásmico
	Cphy3241	3947325..3948227	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3242	3948227..3949159	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
34	Cphy3309	4028258..4028884		No hay coincidencia COG
	Cphy3310	4028993..4030549	PGU1	Endopigalacturonasa
	Cphy3311	4030600..4031262	COG0637	Fosfatasa/fosfohexomutasa predicha
	Cphy3312	4031558..4032568	PurR	Reguladores transcripcionales
	Cphy3313	4032640..4034886	ATH1	Hidrolasas trihalosa y maltosa (posibles fosforilasas)
	Cphy3314	4035015..4037432	ATH1	Hidrolasas trihalosa y maltosa (posibles fosforilasas)
	Cphy3315	4037436..4038314	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3316	4038325..4039221	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy3317	4039364..4040482	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy3318	4041326..4041751	MarR	Reguladores transcripcionales
			COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
35	Cphy3327	4049796..4050236		No hay coincidencia COG
	Cphy3328	4050701..4050982		No hay coincidencia COG
	Cphy3329	4051097..4054015	BglX	Glicosidasas relacionadas con beta glucosidasa
	Cphy3330	4054141..4054431		No hay coincidencia COG
	Cphy3331	4055168..4057156	SalY	Sistema de transporte de péptidos antimicrobianos de tipo ABC, componente ATPasa
	Cphy3332	4057149..4057928	SalX	Sistema de transporte de péptidos antimicrobianos de tipo ABC, componente ATPasa
	Cphy3333	4058214..4059266	BaeS	Histidina quinasa de transducción de señal
	Cphy3334	4059285..4059956	OmpR	Reguladores de respuesta que consisten de un receptor tipo CheY y un dominio de unión a ADN de hélice alada
	Cphy3395	4154158..4154739	AcrR	Regulador transcripcional
	Cphy3396	4154932..4156332	CelF	Alfa-galactosidasas/6-fosfobetaglucosidasas, familia 4 de glicosil hidrolasas
36	Cphy3397	4156675..4157571	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC

	Cphy3398	4157639..4159183	XynB	Beta-xilosidasa
	Cphy3399	4159229..4161427	ATH1	Hidrolasas trihalosa y maltosa (posibles fosforilasas)
	Cphy3400	4161559..4162410	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3401	4162423..4163340	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy3402	4163470..4164825	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy3403	4165352..4166407		No hay coincidencia COG
37	Cphy3404	4166683..4168008	COG5520	O-glicosilhidrolasa
	Cphy3405	4168051..4168812	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy3406	4168817..4170628	COG2972	Proteína de transducción de señal predicha con un domino ATPasa C terminal
	Cphy3407	4170711..4171556	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3408	4171564..4172445	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy3409	4172661..4174016	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
38	Cphy3568	4405037..4406263	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3569	4406276..4407250	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy3570	4407401..4409101	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy3571	4409134..4411029	Chb	N-acetil-beta-hexosaminidasa
39	Cphy3585	4427384..4428406	PurR	Reguladores transcripcionales
	Cphy3586	4428652..4430196	COG3687	Arabinogalactan endo-1,4-beta galactosidasa
	Cphy3587	4430612..4432444		No hay coincidencia COG
	Cphy3588	4432750..4433562	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3589	4433555..4434439	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy3590	4434607..4435923	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy3591	4436346..4438502	Tar	Proteína quimiotáctica aceptadora de metilo
	Cphy3778	4630973..4631749	COG3694	Sistema de transporte no caracterizado de tipo ABC, componente de permeasa
40	Cphy3779	4631756..462532	COG4587	Sistema de transporte no caracterizado de tipo ABC,

41				componente de permeasa
	Cphy3779	4631756..462532	COG4587	Sistema de transporte no caracterizado de tipo ABC, componente de permeasa
	Cphy3780	4632622..4633419	COG4586	Sistema de transporte no caracterizado de tipo ABC, componente de ATPasa
	Cphy3781	4633787..4635553	MdlB	Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy3782	4635540..4637324	MdlB	Sistema de transporte de multifármaco tipo ABC, componentes de permeasa y ATPasa
	Cphy3783	4637694..4638797	LytR	Regulador transcripcional
	Cphy3784	4638930..4639193		No hay coincidencia COG
	Cphy3785	4639618..4641696	COG1501	Alfa-glucosidasas, familia 31 de glicosil hidrolasas
	Cphy3854	4724145..4726538	COG3459	Celobiosa fosforilasa
	Cphy3855	4726828..4728252	{ManB}	Fosfomanomutasa
	Cphy3857	4730021..4731766	LytS	Regulador putativo de la autólisis celular
	Cphy3858	4731867..4733216	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy3859	4733354..4734235	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy3860	4734248..4735123	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3861	4735380..4736159	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy3862	4736925..4744298	XynA	Beta-1,4-xilanasa

Ciertas realizaciones incluyen el uso de ácidos nucleicos que codifican los transportadores de ABC predichos que transportan cualquier producto de la hidrólisis de polímeros. Dichos productos de hidrólisis pueden incluir monosacáridos, por ejemplo, glucosa, manosa, fucosa, galactosa, arabinosa, ramnosa, y xilosa; disacáridos, por ejemplo, trehalosa, maltosa, lactosa, sacarosa, celobiosa; xilobiosa, y oligosacáridos, por ejemplo, celotriosa, celotetraosa, xilotriosa, xilotetraosa, inulina, rafínosa, y melecitosa.

Ciertas realizaciones incluyen transportadores ABC predichos que transportan

celobiosa, por ejemplo, transportadores ABC predichos codificados por Cphy2464, Cphy2465, y Cphy2466.

Reguladores transcripcionales

Algunas realizaciones descritas aqui se relacionan con polinucleótidos, casetes de polinucleótidos casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican reguladores transcripcionales. Otras realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican reguladores transcripcionales.

Los reguladores transcripcionales identificados en *C. phytofermentans* incluyen miembros de las familias AraC y PurR. Los reguladores de AraC pueden incluir activadores transcripcionales de genes involucrados en el metabolismo de carbono (Gallegos M. T. et al. AraC/XylS Family of Transcriptional Regulators. Microbiol. Mol. Biol. Rev. 61, 393-410 (1997)). Los Reguladores de PurR pueden incluir miembros de la familia represora de lactosa (Ramos, J. L. et al. The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev. 69, 326-356 (2005)).

Algunas realizaciones incluyen los reguladores transcripcionales predichos mostrados en la Tabla 8.

Tabla 8: Reguladores transcripcionales predichos en *C. phytofermentans*.

JGI No.	COG	Descripción COG
Cphy0029	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0171	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0342	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0385	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0464	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0572	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC

Cphy0176	AcrR	Regulador transcripcional
Cphy0461	AcrR	Regulador transcripcional
Cphy0674	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0709	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0730	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0768	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0928	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0971	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0610	AcrR	Regulador transcripcional
Cphy1148	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy1165	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy1168	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0667	AcrR	Regulador transcripcional
Cphy0672	AcrR	Regulador transcripcional
Cphy1364	AcrR	Regulador transcripcional
Cphy1472	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy1367	AcrR	Regulador transcripcional
Cphy1513	AcrR	Regulador transcripcional
Cphy1528	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy1546	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy1633	AcrR	Regulador transcripcional
Cphy1683	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy1706	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy1762	AcrR	Regulador transcripcional
Cphy1837	AcrR	Regulador transcripcional
Cphy1838	AcrR	Regulador transcripcional
Cphy1856	AcrR	Regulador transcripcional
Cphy1864	AcrR	Regulador transcripcional
Cphy1910	AcrR	Regulador transcripcional
Cphy2667	AcrR	Regulador transcripcional
Cphy3395	AcrR	Regulador transcripcional
Cphy3621	AcrR	Regulador transcripcional
Cphy1915	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy2187	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy2230	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy2239	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC

Cphy2338	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy2461	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy2556	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0989	ARO8	Reguladores transcripcionales que contienen un dominio HTH de unión a ADN y un dominio aminotransferasa (familia MocR) y sus ortólogos eucarióticos
Cphy1088	ARO8	Reguladores transcripcionales que contienen un dominio HTH de unión a ADN y un dominio aminotransferasa (familia MocR) y sus ortólogos eucarióticos
Cphy2734	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy2228	ARO8	Reguladores transcripcionales que contienen un dominio HTH de unión a ADN y un dominio aminotransferasa (familia MocR) y sus ortólogos eucarióticos
Cphy1297	ARO8	Reguladores transcripcionales que contienen un dominio HTH de unión a ADN y un dominio aminotransferasa (familia MocR) y sus ortólogos eucarióticos
Cphy1446	ARO8	Reguladores transcripcionales que contienen un dominio HTH de unión a ADN y un dominio aminotransferasa (familia MocR) y sus ortólogos eucarióticos
Cphy3132	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy3142	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy3156	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy3159	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy3181	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy3256	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0116	ArsR	Reguladores transcripcionales predichos
Cphy3397	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
Cphy0179	ArsR	Reguladores transcripcionales predichos
Cphy1004	ArsR	Reguladores transcripcionales predichos
Cphy1359	ArsR	Reguladores transcripcionales predichos
Cphy2129	ArsR	Reguladores transcripcionales predichos
Cphy2151	ArsR	Reguladores transcripcionales predichos
Cphy2725	COG1327	Regulador transcripcional predicho, consiste de dominios de una cinta de Zn y cono de ATP
Cphy2664	COG1329	Reguladores transcripcionales, similar a <i>M. xanthus</i> CarD
Cphy2583	COG1386	Regulador transcripcional predicho, que contiene el dominio HTH

Cphy0065	COG1476	Reguladores transcripcionales predichos
Cphy0169	COG1476	Reguladores transcripcionales predichos
Cphy0954	COG1476	Reguladores transcripcionales predichos
Cphy1010	COG1476	Reguladores transcripcionales predichos
Cphy1967	COG1476	Reguladores transcripcionales predichos
Cphy2111	COG1476	Reguladores transcripcionales predichos
Cphy2424	COG1476	Reguladores transcripcionales predichos
Cphy0424	COG1521	Regulador transcripcional putativo, factor accesorio homólogo de Bvg
Cphy1270	COG1695	Reguladores transcripcionales predichos
Cphy1963	COG1695	Reguladores transcripcionales predichos
Cphy2018	COG1695	Reguladores transcripcionales predichos
Cphy2071	COG1695	Reguladores transcripcionales predichos
Cphy2526	COG1695	Reguladores transcripcionales predichos
Cphy3164	COG1695	Reguladores transcripcionales predichos
Cphy3562	COG1695	Reguladores transcripcionales predichos
Cphy0073	COG1725	Reguladores transcripcionales predichos
Cphy0185	COG1725	Reguladores transcripcionales predichos
Cphy1279	COG1725	Reguladores transcripcionales predichos
Cphy2235	COG1725	Reguladores transcripcionales predichos
Cphy2319	COG1725	Reguladores transcripcionales predichos
Cphy3464	COG1725	Reguladores transcripcionales predichos
Cphy3903	COG1725	Reguladores transcripcionales predichos
Cphy1405	COG1733	Reguladores transcripcionales predichos
Cphy1661	COG1733	Reguladores transcripcionales predichos
Cphy1850	COG1733	Reguladores transcripcionales predichos
Cphy0009	COG1959	Regulador transcripcional predicho
Cphy1824	COG1959	Regulador transcripcional predicho
Cphy0991	COG2378	Regulador transcripcional predicho
Cphy1647	COG2378	Regulador transcripcional predicho
Cphy2042	COG2378	Regulador transcripcional predicho
Cphy3341	COG2378	Regulador transcripcional predicho
Cphy0512	COG3437	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio HD-GYP
Cphy1859	COG3655	Regulador transcripcional predicho
Cphy2069	COG3682	Regulador transcripcional predicho

Cphy2324	COG3682	Regulador transcripcional predicho
Cphy2354	COG3682	Regulador transcripcional predicho
Cphy3800	COG4109	Regulador transcripcional predicho que contiene dominios CBS
Cphy0293	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy0496	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy0525	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy0579	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy0618	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy0771	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy0864	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy1394	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy1583	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy1722	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy2007	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy2141	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy2253	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy3034	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy3211	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy3282	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy3327	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC

Cphy3405	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy3697	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy3861	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy3887	COG4753	Regulador de respuesta que contiene un dominio receptor de tipo CheY y un dominio de unión a ADN tipo AraC
Cphy0782	COG4800	Regulador transcripcional predicho, con un dominio HTH
Cphy2167	COG4977	Regulador transcripcional que contiene un dominio amidasa y un dominio HTH de unión a ADN tipo AraC
Cphy1848	COG4978	Regulador transcripcional, dominio/componente de unión efector
Cphy1313	FadR	Reguladores transcripcionales
Cphy1561	FadR	Reguladores transcripcionales
Cphy3829	FadR	Reguladores transcripcionales
Cphy1187	GlpR	Reguladores transcripcionales del metabolismo de azúcar
Cphy2030	GlpR	Reguladores transcripcionales del metabolismo de azúcar
Cphy0655	GntR	Reguladores transcripcionales
Cphy3095	GntR	Reguladores transcripcionales
Cphy3764	GntR	Reguladores transcripcionales
Cphy0097	HipB	Reguladores transcripcionales predichos
Cphy0275	HipB	Reguladores transcripcionales predichos
Cphy0510	HipB	Reguladores transcripcionales predichos
Cphy2912	HipB	Reguladores transcripcionales predichos
Cphy2313	HrcA	Regulador transcripcional del gen de choque térmico
Cphy0149	Lrp	Reguladores transcripcionales
Cphy1811	Lrp	Reguladores transcripcionales
Cphy1102	LysR	Regulador transcripcional
Cphy1229	LysR	Regulador transcripcional
Cphy1477	LysR	Regulador transcripcional
Cphy1757	LysR	Regulador transcripcional
Cphy1783	LysR	Regulador transcripcional
Cphy1902	LysR	Regulador transcripcional
Cphy2431	LysR	Regulador transcripcional
Cphy3040	LysR	Regulador transcripcional
Cphy1293	LysR	Regulador transcripcional

Cphy2156	LysR	Regulador transcripcional
Cphy3352	LysR	Regulador transcripcional
Cphy3361	LysR	Regulador transcripcional
Cphy2557	LytR	Regulador transcripcional
Cphy2794	LytR	Regulador transcripcional
Cphy2795	LytR	Regulador transcripcional
Cphy3783	LytR	Regulador transcripcional
Cphy3892	LytR	Regulador transcripcional
Cphy0854	MarR	Reguladores transcripcionales
Cphy1696	MarR	Reguladores transcripcionales
Cphy1755	MarR	Reguladores transcripcionales
Cphy1844	MarR	Reguladores transcripcionales
Cphy1979	MarR	Reguladores transcripcionales
Cphy2138	MarR	Reguladores transcripcionales
Cphy2555	MarR	Reguladores transcripcionales
Cphy2561	MarR	Reguladores transcripcionales
Cphy2661	MarR	Reguladores transcripcionales
Cphy3318	MarR	Reguladores transcripcionales
Cphy3835	MarR	Reguladores transcripcionales
Cphy3246	NagC	Regulador transcripcional/ quinasa de azúcar
Cphy0329	NagC	Regulador transcripcional/ quinasa de azúcar
Cphy1578	NagC	Regulador transcripcional/ quinasa de azúcar
Cphy3420	NagC	Regulador transcripcional/ quinasa de azúcar
Cphy3573	NagC	Regulador transcripcional/ quinasa de azúcar
Cphy1273	PspC	Regulador transcripcional que responde a estrés putativo
Cphy0484	PurR	Reguladores transcripcionales
Cphy0568	PurR	Reguladores transcripcionales
Cphy0595	PurR	Reguladores transcripcionales
Cphy0698	PurR	Reguladores transcripcionales
Cphy1077	PurR	Reguladores transcripcionales
Cphy1821	PurR	Reguladores transcripcionales
Cphy1876	PurR	Reguladores transcripcionales
Cphy1883	PurR	Reguladores transcripcionales
Cphy2278	PurR	Reguladores transcripcionales
Cphy2351	PurR	Reguladores transcripcionales
Cphy3312	PurR	Reguladores transcripcionales

Cphy3585	PurR	Reguladores transcripcionales
Cphy1454	PurR	Reguladores transcripcionales
Cphy1012	PurR	Reguladores transcripcionales
Cphy0590	PurR	Reguladores transcripcionales
Cphy2353	PurR	Reguladores transcripcionales
Cphy2467	PurR	Reguladores transcripcionales
Cphy2742	PurR	Reguladores transcripcionales
Cphy3700	PurR	Reguladores transcripcionales
Cphy1265	RocR	Regulador transcripcional que contiene PAS, ATPasa tipo AAA, y dominios de unión a ADN.
Cphy1122	RpiR	Reguladores transcripcionales
Cphy3564	RpiR	Reguladores transcripcionales
Cphy0098	SoxR	Reguladores transcripcionales predichos
Cphy0276	SoxR	Reguladores transcripcionales predichos
Cphy0738	SoxR	Reguladores transcripcionales predichos
Cphy1008	SoxR	Reguladores transcripcionales predichos
Cphy1410	SoxR	Reguladores transcripcionales predichos
Cphy1458	SoxR	Reguladores transcripcionales predichos
Cphy1609	SoxR	Reguladores transcripcionales predichos
Cphy1613	SoxR	Reguladores transcripcionales predichos
Cphy2039	SoxR	Reguladores transcripcionales predichos
Cphy3049	SoxR	Reguladores transcripcionales predichos
Cphy3623	SoxR	Reguladores transcripcionales predichos
Cphy3713	SoxR	Reguladores transcripcionales predichos
Cphy3755	SoxR	Reguladores transcripcionales predichos
Cphy3934	Spo0J	Reguladores transcripcionales predichos
Cphy1191	TroR	Regulador transcripcional dependiente de Mn

Ciertas realizaciones incluyen un regulador transcripcional predicho codificado por Cphy2467.

Combinaciones

Algunas realizaciones descritas aquí se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y organismos que comprenden más de un, e.g. dos o más genes identificados en C.

phytofermentans. Algunas realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos casetes de expresión, vectores de expresión, y microorganismos que comprenden más de un gen, e.g. dos o más genes, identificados en *C. phytofermentans*.

Las combinaciones pueden incluir casetes de polinucleótidos que contienen más de un gen identificado en *C. phytofermentans*. En dichas realizaciones, cualquier gen descrito aquí se puede utilizar en combinación con cualquier otro gen descrito aquí. Por ejemplo, cualquier ácido nucleico identificado en *C. phytofermentans* que codifica una hidrolasa se puede utilizar en combinación con cualquier ácido nucleico identificado en *C. phytofermentans* que codifica un transportador ABC. En realizaciones posteriores, cualquier ácido nucleico que codifica una hidrolasa identificada en *C. phytofermentans* se puede utilizar en combinación con un ácido nucleico que codifica un transportador ABC competente, identificado en *C. phytofermentans*, tal como un ácido nucleico que codifica una xilanasa combinada con un ácido nucleico que codifica un transportador de xilosa.

Como se usa aquí, competente puede referirse a al menos dos genes asociados con una vía bioquímica particular. Por ejemplo, competente se puede referir a al menos dos genes en donde el producto del primer gen puede ser el sustrato para el segundo gen, y así sucesivamente. Las ventajas de utilizar genes competentes incluyen la habilidad para engendrar un organismo recombinante con múltiples actividades codificadas por un casete de polinucleótido, por ejemplo, un organismo transformado con un casete de polinucleótidos que comprende una hidrolasa y el transportador ABC competente puede hidrolizar el polímero del sustrato particular para la hidrolasa, y transportar el producto hidrolizado en la célula vía el transportador ABC competente. Una persona versada en la materia puede identificar ejemplos de genes competentes descritos aquí.

En otras realizaciones, cualquier ácido nucleico identificado en *C. phytofermentans* que codifica una hidrolasa se puede utilizar en combinación con cualquier ácido nucleico identificado en *C. phytofermentans* que codifica un regulador transcripcional. En realizaciones adicionales, cualquier ácido nucleico que codifica una hidrolasa identificada en *C. phytofermentans* se puede utilizar en combinación con un ácido nucleico que codifica un regulador transcripcional competente

identificado en *C. phytofermentans*.

En realizaciones particulares, cualquier ácido nucleico identificado en *C. phytofermentans* que codifica un transportador ABC se puede utilizar en combinación con cualquier ácido nucleico identificado en *C. phytofermentans* que codifica un regulador transcripcional. En realizaciones adicionales, cualquier ácido nucleico que codifica un transportador ABC identificado en *C. phytofermentans* se puede utilizar en combinación con un ácido nucleico que codifica un regulador transcripcional competente identificado en *C. phytofermentans*.

En algunas realizaciones, cualquier ácido nucleico identificado en *C. phytofermentans* que codifica una hidrolasa se puede utilizar en combinación con cualquier ácido nucleico identificado en *C. phytofermentans* que codifica un transportador ABC, y cualquier ácido nucleico adicional identificado en *C. phytofermentans* que codifica un regulador transcripcional. En realizaciones adicionales, cualquier ácido nucleico que codifica una hidrolasa identificada en *C. phytofermentans* se puede utilizar en combinación con cualquier ácido nucleico que codifica un transportador ABC competente identificado en *C. phytofermentans*, y cualquier ácido nucleico que codifica un regulador transcripcional competente identificado en *C. phytofermentans*.

En algunas realizaciones, las combinaciones pueden incluir el uso secuencial de más de un gen identificado en *C. phytofermentans*. Por ejemplo, un organismo puede transformarse con un polinucleótido que comprende cualquier gen descrito aquí, y posteriormente transformar con al menos un gen diferente descrito aquí.

Las realizaciones ejemplares de los casetes de polinucleótidos que comprenden, o consisten esencialmente de, combinaciones de al menos dos genes se muestran en la Figura 1. En una realización, la hidrolasa predicha codificada mediante Cphy2276 se puede combinar con los dominios de transportador ABC competente predichos codificados mediante Cphy2272, Cphy2273, y Cphy2274. En otra realización, la hidrolasa predicha codificada mediante Cphy3207 se puede combinar con los dominios del transportador ABC competente predichos codificados mediante Cphy3210, Cphy3209, y Cphy3208, y el regulador transcripcional

competente predicho codificado mediante Cphy3211, y la proteína de transducción de señal codificada por Cphy3212. En otra realización, los dominios del transportador ABC predichos codificados mediante Cphy0862, Cphy0861, y Cphy0860 se pueden combinar con el regulador transcripcional predicho codificado mediante Cphy0864, y la proteína de transducción de señal predicha codificada por Cphy0863. En otra realización, los dominios del transportador ABC codificados por Cphy2466, Cphy2465, y Cphy2464 se pueden combinar con el regulador transcripcional predicho codificado por Cphy2467. En otra realización, la hidrolasa predicha codificada por Cphy1877 se puede combinar con el regulador transcripcional predicho codificado por Cphy1876.

En más realizaciones ejemplares, los casetes de polinucleótidos, casetes de expresión, vectores de expresión, y organismos que comprenden más de un gen pueden comprender agrupamientos de genes identificados en *C. phytofermentans*. Dichos agrupamientos de genes se pueden identificar utilizando los métodos descritos aquí y los métodos bien conocidos en el arte. En algunas realizaciones, los genes y agrupamientos de genes se pueden identificar mediante el grado de homología entre agrupamientos de grupos ortólogos de proteínas (COG, por sus siglas en inglés). Dichos genes y agrupamientos de genes se pueden incluir en casetes o expresar juntos. Ejemplos de agrupamientos de genes identificados en *C. phytofermentans* se muestran en la Tabla 9.

Tabla 9: Agrupamientos de genes identificados en *C. phytofermentans*.

No.	JGI No.	Localización	COG	Descripción de COG
1	Cphy1799	2214443..2216098	COG3469	Quitinasa
	Cphy1800	2216331..2218289	ChiA	quitinasa
2	Cphy1528	1877364..11878218	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy1529	1878477..1879796	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1530	1879890..1880777	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy1531	1880788..1881615	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy1532	1881755..1882096	COG5646	Proteína conservada no caracterizada
3	Cphy3206	3907719..3909908	Tar	Proteína quimiotáctica aceptadora de metilo

	Cphy3207	3910130..3911275	CelA	Endoglucanasa Y
	Cphy3208	3911468..3912373	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3209	3912465..3913424	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
	Cphy3210	3913601..3915310	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy3211	3915499..3915145	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy3212	3917186..3918976	COG2972	Proteína de transducción de señal predicha con un dominio ATPasa C terminal
4	Cphy3367	4103996..4106953	--	--
	Cphy3368	4107033..4109792	--	--
5	Cphy3858	4731867..4733216	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy3859	4733354..4134235	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy3860	4734248..4735123	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy3861	4735380..4736159	COG4753	Regulador de la respuesta que contiene el dominio receptor tipo CheY y dominio de unión a ADN de tipo AraC.
	Cphy3862	4736925..4744298	XynA	Beta-1,4-xilanasa
6	Cphy2272	2801915..2802787	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2273	2802804..2803739	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy2274	2803758..2806796	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2275	2806789..2807427	COG5578	Proteína integral de membrana predicha
	Cphy2276	2807484..2809082	ManB	Beta-mananasa
7	Cphy2464	3025234..3026121	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2465	3026126..3027100	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy2466	3027334..3028755	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2467	3028826..3029881	PurR	Reguladores transcripcionales
8	Cphy1448	1788560..1789609	PhnD	Sistema de transporte de

				fosfato/fosfonato tipo ABC, componente periplásmico
	Cphy1449	1789748..1790512	COG3638	Sistema de transporte de fosfato/fosfonato tipo ABC, componente ATPasa
	Cphy1450	1790509..1791330	COG3639	Sistema de transporte de fosfato/fosfonato tipo ABC, componente permeasa
	Cphy1451	1791345..1792154	COG3639	Sistema de transporte de fosfato/fosfonato tipo ABC, componente permeasa
	Cphy1452	1792326..1793861	UshA	Fosfodiesterasa 5'-nucleotidasa/2',3'-cíclica y esterasas relacionadas
9	Cphy1071	1354865..1357051	ManB	Beta-mananasa
	Cphy1074	1358682..1360004	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1075	1360064..1360906	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy1076	1360906..1361769	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
10	Cphy1132	1424925..1425929	RbsB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy1133	1426063..1427142	AraH	Sistemas de transporte de ribosa/xilosa/arabinosa/galactósido tipo ABC, componentes de permeasa
	Cphy1134	1427155..1428675	MglA	Sistema de transporte de azúcar tipo ABC, componente ATPasa
11	Cphy1694	2078054..2081500	--	--
	Cphy1695	2081574..2082656	WcaA	Glicosiltransferasas involucradas en la biogénesis de la pared celular
12	Cphy1876	2309264..2310301	PurR	Reguladores transcripcionales
	Cphy1877	2310355..2312748	COG1501	Alfa-glucosidasas, familia 31 de glicosil hidrolasas
13	Cphy2105	2602117..2602755	--	--
	Cphy2106	2603026..2603856	COG1262	Proteína conservada no caracterizada
	Cphy2107	2604341..2605282	COG3708	Proteína no caracterizada conservada en bacterias
	Cphy2108	2605495..2607915	XynA	Beta-1,4-xilanasa
14	Cphy2237	2756774..2758021	Galk	galactoquinasa
	Cphy2238	2758041..2758286	--	--
	Cphy2239	2758623..2759510	AraC	Proteínas que contienen el dominio de unión a ADN de tipo AraC
	Cphy2240	2759556..2760221	--	--
	Cphy2241	2760521..2761603	MglC	Sistema de transporte de glucosa/galactosa tipo ABC, componente permeasa
	Cphy2242	2761619..2763118	MglA	Sistema de transporte de azúcar

				tipo ABC, componente ATPasa
	Cphy2243	2763191..2764294	RbsB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
15	Cphy2262	2788055..2789236	COG2942	N-acil-D-glucosamina 2-epimerasa
	Cphy2263	2789493..2790608	TesA	Lisofosfolipasa L1 y esterasas relacionadas
	Cphy2264	2790617..2791639	COG2152	Glicosilasa predicha
	Cphy2265	2791741..2793189	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2266	2793236..2795134	--	--
	Cphy2267	2795272..2796264	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2268	2796278..2797195	UgpA	Sistema de transporte de azúcar tipo ABC, componentes de permeasa
	Cphy2269	2797161..2799770	--	--
	Cphy2270			
	Cphy2271			
	Cphy2272			Componente membranal interno de sistemas de transporte dependientes de proteínas de unión
	Cphy2273			Componente membranal interno de sistemas de transporte dependientes de proteínas de unión
	Cphy2274			Familia de proteínas de unión a soluto extracelular 1
	Cphy2275			Proteína hipotética
	Cphy2276			Manan endo 1,4-beta manosidasa
16	Cphy2569	3136132..3137703	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2570	3137758..3138693	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2571	3138707..3139672	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
17	Cphy2654	3239628..3241388	UgpB	Sistema de transporte de azúcar tipo ABC, componente periplásmico
	Cphy2655	3241527..3242447	UgpE	Sistema de transporte de azúcar tipo ABC, componente de permeasa
	Cphy2656	3242462..3243409	LplB	Sistema de transporte de polisacáridos tipo ABC, componente de permeasa
18	Cphy2807	3420322..3421581	COG1216	Glicosiltransferasas predichas
	Cphy2808	3421710..3422930	--	--
	Cphy2809	3423037..3429723	Smc	ATPasas de segregación de cromosomas
	Cphy2810	3429863..3430933	WcaA	Glicosiltransferasas involucradas en la biogénesis de la pared celular
	Cphy2811	3430994..3433558	RfaG	Glicosiltransferasa

	Cphy2812	3433803..3434108	--	--
	Cphy2813	3434217..3435191	RfaG	Glicosiltransferasa
	Cphy2814	3435346..3436719	COG1216	Glicosiltransferasas predichas
	Cphy2815	3437022..3437582	COG1633	Proteína conservada no caracterizada
	Cphy2816	3437827..3438198	--	--
	Cphy2817	3438599..3440215	--	--
	Cphy2818	3440301..3440876	AmiC	N-acetilmuramoil-L-alanina amidasa
19	Cphy3009	3472467..3674620	BglX	Glicosidasas relacionadas con beta-glucosidasa
	Cphy3010	3674634..3675599	XynA	Beta-1,4-xilanasas
20	Cphy3419	4198367..4199833	XylB	Quinasas para azúcares (pentosa y hexulosa)
	Cphy3420	4200152..4201297	NagC	Regulador transcripcional/quinasa de azúcar
21	Cphy3854	4724145..4726538	COG3459	Celobiosa fosforilasa
	Cphy3855	4726828..4728252	ManB	Fosfomanomutasa
	Cphy3857	4730021..4731766	LytS	Regulador putativo de la autólisis celular
22	Cphy2008			Histidina quinasa de transducción de señal multisensor
	Cphy2009			Proteína de unión periplásmica/regulador transcripcional LacI
	Cphy2010			Relacionada con transportador ABC
	Cphy2011			ATPasa transportadora de monosacáridos
	Cphy2012			Proteína de unión periplásmica/regulador transcripcional LacI

Enzimas involucradas en la asimilación de xilosa

Algunas realizaciones descritas aquí se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican genes involucrados en la asimilación de xilosa. Otras realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican genes involucrados en la asimilación de xilosa.

Como se usa aquí, los genes involucrados en la asimilación de xilosa pueden incluir, por ejemplo, genes que codifican hidrolasas para la hidrólisis de polímeros a xilosa, transportadores ABC para el transporte de xilosa a la célula, reguladores de la

transcripción para la regulación de estos genes que codifican hidrolasas y/o transportadores ABC, y enzimas relacionadas con la fermentación de azúcares pentosas, tales como xilosa, a alcoholes. Los genes identificados como sobre-regulados cuando *C. phytofermentans* creció en xilosa incluyen Cphy3419, Cphy1219, y Cphy1585, Cphy1586, y Cphy1587 (ver Fig. 13).

Mientras que muchas especies de Clostridia pueden degradar hemicelulosa, la mayoría de las especies no pueden fermentar las pentosas que resultan de dicha hidrólisis. De forma notable, *C. phytofermentans* puede hidrolizar hemicelulosa a pentosas y fermentar pentosas a alcoholes. *C. phytofermentans* puede transportar pentosas a la célula como oligosacáridos o como monosacáridos. El genoma de *C. phytofermentans* contiene genes que codifican enzimas para la asimilación de xilosa incluyendo enzimas en la vía no oxidativa pentosa fosfato que esta relacionada con la conversión de pentosas a hexosas. Consistente con la habilidad para fermentar pentosas, los datos de expresión con células que crecen en xilano han mostrado que las enzimas clave en la vía pentosa fosfato, es decir, transaldolasa (EC 2.2.1.1 Cphy0013) y transcetolasa (EC 2.2.1.1, Cphy0014), están entre los transcriptos más abundantes. La gliceraldehído-3-fosfato deshidrogenasa (EC 1.2.1.12, Cphy2879), que conecta la vía pentosa fosfato y las reacciones iniciales de glicólisis a energía en los pasos iniciales de la glicólisis, está fuertemente inducida por xilano, celulosa, celobiosa, y glucosa. Otros genes sobre-regulados durante el crecimiento en xilano incluyen Cphy2105, Cphy2106, Cphy2108, Cphy1510, Cphy3158, Cphy3009, Cphy3010, Cphy3419, Cphy1219, Cphy2632, Cphy3206, Cphy3207, Cphy3208, Cphy3209, Cphy3210, Cphy3211, Cphy3212, Cphy1448, Cphy1449, Cphy1450, Cphy1451, Cphy1132, Cphy1133, Cphy1134, Cphy1528, Cphy1529, Cphy1530, Cphy1531, y Cphy1532.

La fermentación de hexosas y pentosas termina con la reducción de acetil-coA a etanol catalizada por enzimas que incluyen acetaldehído deshidrogenasa dependiente de NAD(P) (Aid) y alcohol deshidrogenasa dependiente de NAD (Adh). El genoma de *C. phytofermentans* contiene genes putativos que codifican al menos 7 Aid (Dominio PutA), y al menos 6 Adh, por ejemplo, la proteína putativa codificada en Cphy3925 que contiene los dominios Aid y Adh. 4 Aid y 3 Adh están codificados por genes en tres agrupamientos: Cphy1173-1183; Cphy1411-1430; y Cphy2634-2650.

Enzimas involucradas en la producción de propanol, el metabolismo de etanolamina y/o propanodiol

Algunas realizaciones descritas aquí se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican genes involucrados en la producción de propanol, el metabolismo de etanolamina y/o propanodiol. Algunas realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican genes involucrados en la producción de propanol, el metabolismo de etanolamina y/o propanodiol.

C. phytofermentans contiene microcompartimientos proteinaceos ("PMC", por sus siglas en inglés) que no se encuentran en otras bacterias de interés biotecnológico similar, tales como *C. cellulolyticum*, *C. thermocellum*, *C. acetobutylicum*, y *C. beijerinckii*. Estos microcompartimientos se han observado mediante microscopía electrónica. Las enzimas particulares involucradas en la conversión de carbohidratos a alcoholes están localizadas hacia estos microcompartimientos, sugiriendo la compartimentalización de vías particulares y mayor eficiencia metabólica (Conrado, R. J., Mansell, T. J., Varner, J. D. & DeLisa, M. P. Stochastic reaction-diffusion simulation of enzyme compartmentalization reveals improved catalytic efficiency for a synthetic metabolic pathway. *Metab. Eng.* 9, 355-363 (2007)).

Tres loci genéticos en *C. phytofermentans* codifican proteínas localizadas hacia compartimientos proteínicos. Estos compartimientos proteínicos son similares a los compartimientos proteínicos involucrados en la fijación de dióxido de carbono, y en la utilización de etanolamina y propanodiol encontrados en otros organismos. Cada locus incluye enzimas para la conversión de azúcares con cinco átomos de carbono y alcohol deshidrogenasas a alcoholes primarios.

De las 7 Aid y 6Adh identificadas en *C. phytofermentans*, 4 Aid y 3Adh, están localizadas hacia los microcompartimientos proteínicos. Las Adh localizadas hacia

los microcompartimientos proteínicos muestran identidad de secuencia para Fe-Adh o Zn-Adh, y están codificadas por genes en tres agrupamientos: Cphy1173-1183; Cphy1411-1430; y Cphy2634-2650.

Más enzimas localizadas en los microcompartimientos proteináceos pueden estar relacionadas con la vía de fucosa a propanol, así como con el metabolismo de etanolamina y propanodiol. Por ejemplo, el agrupamiento Cphy2634-2650 contiene ortólogos de genes implicados en el metabolismo de etanolamina en *Salmonella typhimurium*, y el agrupamiento Cphy1411-1430 contiene genes que codifican productos que pueden estar relacionados funcionalmente con el operón de utilización de propanodiol en *Salmonella typhimurium*.

Adicionalmente, el agrupamiento Cphy1173-1187 contiene genes homólogos a un microcompartimento encontrado en *Roseburia inulinivorans* (Scott, KP, Martin, J. C, Campbell, G. Mayer, C.D. & Flint, H.J. Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium *Roseburia inulinivorans*. J. Bacteriol. 188, 4340-4349 (2006)) y genes que codifican enzimas teóricas involucradas en la utilización de fucosa y ramnosa (ver Figs. 14 y 15). Genes adicionales identificados como sobrerregulados durante el crecimiento sobre fucosa o que de otra manera se ha predicho que están implicados en la utilización de fucosa incluyen Cphy3153, Cphy3154, Cphy3155, Cphy2010, Cphy2011 y Cphy2012 (Fig. 14). Genes adicionales identificados como sobrerregulados durante el crecimiento sobre ramnosa o que de otra manera se ha predicho que están implicados en la utilización de ramnosa incluyen Cphy0578, Cphy0579, Cphy0580, Cphy0581, Cphy0582, Cphy0583, Cphy0584, Cphy1146, Cphy1147, Cphy1148, Cphy1149 (Fig. 15).

Producción de Hidrógeno.

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican los genes involucrados en la producción de hidrógeno. Otras realizaciones se relacionan con métodos para producir combustible

utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y los microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* como genes codificantes involucrados en la producción de hidrógeno.

Puede producirse hidrógeno a partir de la fermentación de una variedad de azúcares. En algunas realizaciones, los polinucleótidos pueden comprender ácidos nucleicos que codifican ferredoxina hidrogenasas identificados en *C. phytofermentans*. Ejemplos de genes que codifican ferredoxina hidrogenasas identificados en *C. phytofermentans* incluyen Cphy0087, Cphy0090, Cphy0092, Cphy2056, Cphy3805, Cphy3798.

Polisacárido Liasa Multimodular

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican las enzimas/dominios de proteínas involucradas en la hidrólisis de la pectina. Algunas realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteínas involucradas en la hidrólisis de la pectina. Ejemplos de genes que codifican enzimas/dominios de proteínas involucradas en la hidrólisis de pectina pueden incluir genes en el locus Cphy1612. El locus Cphy1612 codifica dominios PL1 y PL9 predichos. PL1 incluye un dominio pectato liasa (CE 4.2.2.2); exo-pectato liasa (CE 4.2.2.9), y pectina liasa (CE 4.2.2.10). PL9 incluye un dominio pectato liasa (CE 4.2.2.2) y exopoligalacturonato liasa (CE 4.2.2.9).

Esterasa y Xilanasa Multimodular

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteína que incluyen

actividades xilanasa y esterasa. Otras realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* como codificadores de enzimas/dominios de proteínas que incluyen actividades xilanasa y esterasa. Ejemplos de genes que codifican enzimas/dominios de proteína que incluyen actividades xilanasa y esterasa, pueden incluir genes en el locus Cphy3862. El locus Cphy3862 incluye tres dominios predichos, a saber, dos dominios GH10 y un dominio CE15, que tienen las siguientes actividades: GH10 con actividad xilanasa (CE 3.2.1.8); GH10 con actividad endo-1,3--xilanasa (CE 3.2.1.32), y CE15, con actividad glucuronil esterasa (CE 3.1.1.-) y actividad 4-O-metil-glucuronil esterasa (CE 3.1.1.-).

Utilización de Laminarina

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteínas involucradas en la utilización de laminarina. Algunas realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótido, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteínas involucradas en la utilización de laminarina. La laminarina es un glucano (un polisacárido de glucosa) de almacenamiento que se encuentran en las algas pardas. Ejemplos de genes identificados como sobrerregulados durante el crecimiento sobre laminarina incluyen Cphy0857, Cphy0858, Cphy0859, Cphy0860, Cphy0861, Cphy0862, Cphy0863, Cphy0864, Cphy0865 y Cphy3388 (ver Fig. 16).

Utilización de Celobiosa

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteínas involucradas en la utilización de celobiosa. Otras realizaciones se relacionan con métodos para producir

combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/ dominios de proteínas involucradas en la utilización de celobiosa. La celobiosa es un disacárido derivado de la condensación de dos moléculas de glucosa unidas en un enlace $\beta(1\rightarrow4)$. Ejemplos de genes identificados como sobrerregulados durante el crecimiento sobre celobiosa incluyen Cphy0430, Cphy2464, Cphy2465, Cphy2466 y Cphy2467 (ver Fig. 17).

Utilización de Celulosa

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteína involucradas en la utilización de celulosa. Algunas realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteína involucrados en la utilización de celulosa. Ejemplos de genes identificados como sobrerregulados durante el crecimiento sobre celulosa o que de otra manera se ha predicho están involucrados en la utilización de celulosa incluyen Cphy3367, Cphy3368, Cphy1163, Cphy3202, Cphy3160, Cphy0430, Cphy3854, Cphy3855, Cphy3857, Cphy3858, Cphy3859, Cphy3860, Cphy3861, Cphy3862, Cphy2569, Cphy2570, Cphy2571, Cphy2464, Cphy2465, Cphy2466, Cphy2467, Cphy1528, Cphy1529, Cphy1530, Cphy1531 y Cphy1532.

Utilización de Pectina

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que comprenden ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteína involucrados en la utilización de pectina. Otras realizaciones se relacionan con métodos para producir combustible utilizando los polinucleótidos, casetes de polinucleótidos, casetes de expresión, vectores de expresión, y microorganismos que incluyen ácidos nucleicos identificados en *C. phytofermentans* que codifican enzimas/dominios de proteína involucradas en la utilización de pectina. Ejemplos de genes identificados como

sobrerregulados durante el crecimiento sobre pectina incluyen Cphy3585, Cphy3586, Cphy3587, Cphy3588, Cphy3589, Cphy3590, Cphy2262, Cphy2263, Cphy2264, Cphy2265, Cphy2266, Cphy2267, Cphy2268, Cphy2269, Cphy2272, Cphy2273, Cphy2274, Cphy2275, Cphy2276, Cphy2464, Cphy2465, Cphy2466, Cphy2467, Cphy1714, Cphy1715, Cphy1716, Cphy1717, Cphy1718, Cphy1719, Cphy1720, Cphy3153, Cphy3154, Cphy3155, Cphy2010, Cphy2011, Cphy1174, Cphy1175, Cphy1176, Cphy1177, Cphy1178, Cphy1179, Cphy1180, Cphy1181, Cphy1182, Cphy1183, Cphy1929, Cphy1612, Cphy0218, Cphy0219, Cphy0220, Cphy3160, y Cphy2919.

Los genes sobrerregulados durante el crecimiento sobre pectina y que se ha predicho están involucrados en la degradación y el transporte de la cadena lateral arabinogalactano del ramnogalacturonano-I, incluyen Cphy3585, Cphy3586, Cphy3587, Cphy3588, Cphy3589 y Cphy3590. Los genes sobrerregulados durante el crecimiento en pectina y que se ha predicho están involucrados en la degradación y el transporte de las cadenas laterales de ramnogalacturonano-I o ramnogalacturonano-II incluyen Cphy2262, Cphy2263, Cphy2264, Cphy2265, Cphy2266, Cphy2267, Cphy2268, Cphy2269, Cphy2272, Cphy2273, Cphy2274, Cphy2275, Cphy2276, Cphy1714, Cphy1715, Cphy1716, Cphy1717, Cphy1718, Cphy1719 y Cphy1720.

Los genes sobrerregulados durante el crecimiento sobre pectina y que se ha predicho están involucrados en el transporte de azúcar incluyen Cphy2464, Cphy2465, Cphy2466 y Cphy2467. Los genes que se ha predicho están involucrados en la degradación y el transporte de ácido poligalacturónico incluyen Cphy0288, Cphy0289, Cphy0290, Cphy0291, Cphy0292 y Cphy0293. Los genes que se ha predicho están involucrados en la lisis y el transporte de ramnogalacturonano incluyen Cphy0339, Cphy0340, Cphy0341, Cphy0342, Cphy0343. Los genes que se ha predicho están involucrados en el transporte y degradación de ramnosa incluyen Cphy0578, Cphy0579, Cphy0580, Cphy0581, Cphy0582, Cphy0583, Cphy0584, Cphy1146, Cphy1147, Cphy1148, y Cphy1149. Los genes sobrerregulados durante el crecimiento sobre pectina y/o que se ha predicho están involucrados en el transporte y degradación de fucosa incluyen Cphy3153, Cphy3154, Cphy3155, Cphy2010, Cphy2011 y Cphy2012. Los genes sobrerregulados durante el crecimiento sobre pectina y/o que se ha predicho participan en el metabolismo de fucosa y ramnosa

incluyen Cphy1174, Cphy1175, Cphy1176, Cphy1177, Cphy1178, Cphy1179, Cphy1180, Cphy1181, Cphy1182, Cphy1183, Cphy1184, Cphy1185, Cphy1186, y Cphy1187.

Los genes sobrerregulados durante el crecimiento sobre pectina y/o que se ha predicho están involucrados en la utilización de ácido poligalacturónico incluyen Cphy2919, Cphy0288, Cphy0289, Cphy0290, Cphy0291, Cphy0292, Cphy0293, Cphy3308, Cphy3309, Cphy3310, Cphy3311, Cphy3312, Cphy3313, Cphy3314, Cphy3315, Cphy3316, Cphy3317, Cphy1118, Cphy1119, Cphy1120, Cphy1121, Cphy1879, Cphy1880, Cphy1881, Cphy1882, Cphy1883, Cphy2736, Cphy2737, Cphy2738, Cphy2739, Cphy2740, Cphy2741, Cphy2742 y Cphy2743.

Identificación de Secuencias de Ácidos Nucléicos en *C. phytofermentans*

Algunas realizaciones aquí descritas se relacionan con métodos para identificar genes en *C. phytofermentans*. Tales métodos pueden identificar incluir secuencias de ácidos nucleicos que contienen secuencias codificantes, secuencias no codificantes, secuencias reguladoras, secuencias intergénicas, operones o agrupamientos de genes. En algunas realizaciones, los métodos para identificar genes en *C. phytofermentans* pueden incluir análisis genómico o de microarreglos.

En algunas realizaciones, puede identificarse un gen en *C. phytofermentans* mediante la similitud del gen con otra secuencia. Puede determinarse la similitud entre secuencias de polinucleótidos o secuencias de polipéptidos. En algunas realizaciones, otra secuencia puede ser una secuencia presente en otro organismo. Ejemplos de otros organismos pueden incluir un organismo de diferentes especies de Clostridia, tales como *C. beijerinckii* o *C. acetobutylicum*; o un organismo de un género diferente, tal como *Bacillus subtilis*.

En algunas realizaciones, la similitud puede medirse como un porcentaje de identidad. El porcentaje de identidad de secuencia puede ser una relación entre dos o más secuencias de polipéptidos o dos o más secuencias de polinucleótidos, determinado al comparar las secuencias. En realizaciones adicionales, la identidad

de las secuencias puede ser el grado de relación de secuencia entre secuencias de polipéptidos o polinucleótidos, según el caso, determinada por la coincidencia entre cadenas de tales secuencias. De manera típica, la identidad de secuencia y la similitud de secuencia pueden calcularse fácilmente mediante métodos conocidos, incluyendo pero sin limitarse a aquellos descritos en: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, Nueva York (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, Nueva York (1993); Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, Nueva Jersey (1994); Sequence Analysis in Molecular Biology (von Heinje, G., ed.) Academic Press (1987); y Sequence Analysis Primer (Gribskov, M. y Devereux, J., eds.) Stockton Press, NY (1991). Los métodos para determinar la identidad de secuencia pueden diseñarse para dar la mejor concordancia entre las secuencias evaluadas. Algunos métodos para determinar la identidad de secuencia y la similitud de secuencia están codificados en programas de computador a disposición del público. Los cálculos de porcentaje de identidad y alineaciones de secuencia pueden llevarse a cabo usando el programa Megalign del conjunto de procesamiento computacional en bioinformática LASERGENE (DNASTAR Inc., Madison, Wis.). La alineación múltiple de las secuencias puede llevarse a cabo usando el método Clustal de alineación (Higgins y Sharp (1989) CABIOS. 5:151-153) con los parámetros por defecto (PENALIZACIÓN POR GAP (GAP PENALTY)=10, PENALIZACIÓN POR LONGITUD DE GAP (GAP LENGTH PENALTY)=10). Los parámetros por defecto para alineaciones por pares usando el método Clustal fueron KTUPLE 1, PENALIZACIÓN POR GAP (GAP PENALTY)=3, VENTANA (WINDOW)=5 y DIAGONALES GUARDADAS (DIAGONALS SAVED)=5.

En otras realizaciones, puede identificarse un gen en *C. phytofermentans* prediciendo la presencia de un gen en una secuencia de ácido nucléico y/o secuencia de polipéptido teórico traducidos usando algoritmos bien conocidos en el arte. Por ejemplo, pueden usarse algoritmos de computador en programas tales como GeneMark™ (Besemer, J., y M. Borodovsky. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33:W451-4) y Glimmer (Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673-9).

En algunas realizaciones, las secuencias de nucleótidos o aminoácidos pueden analizarse usando un algoritmo de computador o un programa de software. En realizaciones relacionadas, el software de análisis de secuencias puede estar disponible comercialmente o ser desarrollado de manera independiente. Ejemplos de software de análisis de secuencias incluyen el conjunto de programas GCG (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, Wis.), BLASTP, BLASTN, BLASTX (Altschul et al, J. Mol. Biol. 215:403-410 (1990), y DNASTAR (DNASTAR, Inc. 1228 S. Park St. Madison, Wis. 53715 USA), y el programa FASTA que incorpora el algoritmo Smith-Waterman (W.R. Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(es): Suhai, Sandor. Editorial: Plenum, Nueva York, N.Y.). De manera típica, pueden usarse los valores por defecto de un programa, por ejemplo, una serie de valores o parámetros originalmente cargados con el software cuando es inicializado por primera vez.

En otras realizaciones pueden usarse bases de datos de dominios conservados de proteínas y familias de proteínas para identificar un gen en *C. phytofermentans*. Por ejemplo, la Base de Datos de Dominios Conservados (CDD, por las siglas en inglés de Conserved Domain Database) del Centro Nacional para Información en Biotecnología (NCBI, por las siglas en inglés de National Center for Biotechnology Information) comprende diversas bases de datos que incluyen los Dominios Conservados del NCBI, SMART (smart.embl-heidelberg.de/SMART), PFAM (disponible en la internet en sanger.ac.uk/Software/Pfam/PFAM), y COGS (clasificación filogenética de proteínas codificadas en genomas completos).

En algunas realizaciones pueden identificarse genes y predecirse rutas metabólicas de proteínas teóricas codificadas por los genes. En tales realizaciones, pueden usarse bases de datos de rutas metabólicas. Por ejemplo, la enciclopedia de Kyoto de Genes y Genomas (KEGG, por las siglas en inglés de Kyoto Encyclopedia of Genes and Genomes), donde el Servidor de Anotación Automático de la KEGG (disponible en internet en genome.jp/kegg/kaas/) puede proveer anotación funcional de genes identificados usando comparaciones BLAST contra la base de datos GENES de la KEGG.

Aislando Secuencias de Ácidos Nucléicos de *C. phytofermentans*

Las secuencias de ácidos nucleicos pueden clonarse a partir del genoma de *C. phytofermentans* usando técnicas ampliamente conocidas en el arte. Por ejemplo, las técnicas de clonación molecular y de ADN recombinante que pueden emplearse son descritas por Sambrook, J., Fritsch, E. F. y Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Segunda Edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); y por Silhavy, T. J., Bennis, M. L. Y Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Cold Press Spring Harbor, N.Y. (1984); y por Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, publicado por Greene Publishing Assoc. y Wiley-Interscience (1987). Adicionalmente, los métodos para aislar genes homólogos u ortólogos usando protocolos dependientes de secuencia son ampliamente conocidos en el arte. Ejemplos de protocolos dependientes de secuencia incluyen, pero no se limitan a, métodos de hibridación de ácidos nucleicos, y métodos de amplificación de ADN y ARN como se ejemplifica mediante diversos usos de las tecnologías de amplificación de ácidos nucleicos, tales como, reacción en cadena de la polimerasa (PCR, por las siglas en inglés de polymerase chain reaction; Mullis et al., Pat. U.S. 4,683,202), reacción en cadena de la ligasa (LCR, por las siglas en inglés de ligase chain reaction; Tabor, S. et al., Proc. Acad. Sci. USA 82, 1074, (1985)) o amplificación de desplazamiento de hebra (SDA, por las siglas en inglés de strand displacement amplification; Walker, et al., Proc. Natl. Acad. Sci. U.S.A., 89, 392, (1992)).

De manera típica, en las técnicas de amplificación de tipo PCR, los cebadores tienen diferentes secuencias y no son complementarios entre sí. Dependiendo de las condiciones de ensayo deseadas, las secuencias de los cebadores deben diseñarse para proveer una replicación tanto eficiente como fiel del los ácidos nucleicos objetivo. Los métodos de diseño de cebadores para PCR son comunes y ampliamente conocidos en el arte (Thein y Wallace, "The use of oligonucleotide as specific hybridization probes in the Diagnosis of Genetic Disorders", en *Human Genetic Diseases: A Practical Approach*, K. E. Davis Ed., (1986) pp. 33-50 IRL Press, Herndon, Va.; Rychlik, W. (1993) In White, B. A. (ed.), *Methods in Molecular Biology*, Vol. 15, páginas 31-39, PCR Protocols: Current Methods and Applications. Humana Press, Inc., Totowa, N.J.).

De manera general, dos segmentos cortos de una secuencia identificada pueden emplearse en protocolos de PCR para amplificar fragmentos más largos de ácidos nucleicos que codifican genes homólogos a partir de ADN o ARN. La PCR puede llevarse a cabo en una librería de fragmentos de ácidos nucleicos clonados en donde la secuencia de un cebador se deriva de la secuencia identificada de ácido nucleico, y la secuencia del otro cebador se deriva de los trectos 3' de ácido poliadenílico característicos del precursor mRNA que codifica los genes microbianos. De manera alternativa, la secuencia del segundo cebador puede basarse en las secuencias que se derivan de un vector de clonación. Por ejemplo, el protocolo RACE (Frohman et al., PNAS USA 85:8998 (1988)) provee un medio para generar cADNs usando PCR para amplificar copias de la región entre un único punto en el transcripto y el extremo 3' o 5'. Los cebadores orientados en las direcciones 3' y 5' pueden diseñarse a partir de la secuencia identificada. Usando los sistemas RACE 3' o RACE 5' disponibles comercialmente (BRL), pueden aislarse fragmentos de cADN 3' o 5' específicos (Ohara et al., PNAS USA 86:5673 (1989); Loh et al., Science 243:217 (1989)).

En algunas realizaciones, pueden aislarse secuencias de ácidos nucleicos identificadas mediante selección de una biblioteca de ADN de *C. phytofermentans* usando una porción del ácido nucleico identificado como una sonda de hibridación de ADN. Ejemplos de sondas pueden incluir sondas de ADN marcadas mediante métodos tales como, marcación aleatoria de cebador de ADN, desplazamiento de mella, o técnicas de marcación de extremos, y sondas de ARN producidas mediante métodos tales como, sistemas de transcripción *in vitro*. Adicionalmente, pueden diseñarse oligonucleótidos específicos y usarse para amplificar una parte de o toda la extensión de las presentes secuencias. Los productos de amplificación resultantes pueden marcarse de manera directa durante las reacciones de amplificación o marcarse después de las reacciones de amplificación, y usarse como sondas para aislar toda la extensión de los fragmentos de ADN bajo condiciones de adecuado rigor.

En algunas realizaciones, los ácidos nucleicos aislados son clonados en vectores. De manera típica, los vectores tienen la habilidad de replicarse en un microorganismo hospedero. Se conocen numerosos vectores, por ejemplo, bacteriófagos, plásmidos, virus, o híbridos de los mismos. Los vectores pueden ser

operables como vectores de clonación o vectores de expresión en la célula hospedero seleccionada. De manera típica, un vector comprende un ácido nucleico aislado, un marcador seleccionable, y secuencias que permitan la replicación autónoma o la integración cromosómica. Realizaciones adicionales pueden comprender una secuencia promotora que dirija la expresión de un ácido nucleico aislado, un potenciador, o una secuencia de terminación. En otras realizaciones, un vector puede comprender secuencias que permitan la escisión de las secuencias tras la integración al ADN cromosómico de las secuencias del vector. Ejemplos incluyen secuencias *loxP* o secuencias FRT, respondiendo estas secuencias a CRE recombinasa y FLP recombinasa, respectivamente.

Polinucleótidos, Casetes de Polinucleótidos, Casetes de Expresión, y Vectores de Expresión

Algunas realizaciones aquí descritas se relacionan con polinucleótidos, casetes de polinucleótidos, casetes de expresión, y vectores de expresión útiles para la producción de un combustible u otro producto en un microorganismo recombinante.

Los casetes de polinucleótidos pueden comprender al menos un polinucleótido de interés. En algunas realizaciones, un casete de polinucleótidos puede comprender más de un polinucleótido de interés. Por ejemplo, un casete de polinucleótidos puede comprender dos o más, tres o más, o cualquier número de genes y/o polinucleótidos de interés aquí descritos.

En algunas realizaciones, un polinucleótido de interés puede incluir uno o más ácidos nucleicos aquí descritos identificados en *C. phytofermentans*. En algunas realizaciones, el polinucleótido de interés puede tener por lo menos 50%, 55%, 60%, 65%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, y 100% de identidad con uno o más genes identificados en *C. phytofermentans*. En otras realizaciones, el polinucleótido de interés puede codificar una o más proteínas que comprenden sustituciones conservativas a la proteína de tipo nativo. En realizaciones adicionales, el polinucleótido de interés puede codificar una o más proteínas que comprenden sustituciones que alteran la eficiencia de la proteína para la producción de

combustible. Por ejemplo, las proteínas que codifican enzimas pueden hacerse más eficientes en catalizar reacciones.

Como se usa aquí, un casete de expresión puede ser un(unos) polinucleótido(s) de interés unido(s) de manera operable a una secuencia reguladora, tal como un promotor. Promotores adecuados para la presente invención incluyen cualquier promotor para la expresión del polinucleótido de interés. En algunas realizaciones el promotor puede ser la secuencia promotora identificada en *C. phytofermentans*. En algunas realizaciones, el promotor puede ser una secuencia promotora identificada en un organismo huésped. En algunas realizaciones, el promotor puede ser un promotor inducible, tal como, por ejemplo, un promotor inducible ligero o un promotor sensible a la temperatura. En otras realizaciones, el promotor puede ser un promotor constitutivo. En algunas realizaciones, un promotor puede seleccionarse con base en el nivel de expresión deseada para el(los) polinucleótido(s) de interés en el microorganismo hospedero. En algunas realizaciones, el promotor puede ubicarse alrededor de la misma distancia del sitio de inicio de transcripción heteróloga como lo está del sitio de inicio de transcripción en su configuración natural. Sin embargo, como es bien sabido en el arte, puede permitirse alguna variación en esta distancia sin perder la función promotora. En otras realizaciones, un casete de expresión puede comprender además secuencias reguladoras tales como potenciadores y/o secuencias de terminación.

Los elementos promotores pueden seleccionarse y movilizarse en un vector (e.g., pIMPCphy). Por ejemplo, una secuencia reguladora de transcripción se une de manera operable al(a los) gen(es) de interés (e.g., en un constructo de expresión). El promotor puede ser cualquier arreglo de secuencias de ADN que interactúe específicamente con factores de transcripción celular para regular la transcripción del gen corriente abajo. La selección de un promotor particular depende del tipo de célula que se vaya a emplear para expresar la proteína de interés. De manera general, una secuencia reguladora de transcripción útil es una del microorganismo hospedero. En diversas realizaciones, se seleccionan los promotores constitutivos o inducibles para el uso en una célula hospedero. Dependiendo de la célula hospedero, hay potencialmente cientos de promotores constitutivos e inducibles que son conocidos y que pueden ser modificados mediante ingeniería genética para funcionar en la célula hospedera.

Un promotor puede ser un arreglo de secuencias de ADN que interactúan específicamente con factores de transcripción celular para regular la transcripción del gen corriente abajo. La selección de un promotor particular depende del tipo de célula a emplear para expresar la proteína de interés. Las secuencias de transcripción reguladora pueden ser aquellas del microorganismo hospedero. En diversas realizaciones, se seleccionan promotores constitutivos o inducibles para el uso en una célula hospedera. Dependiendo de la célula hospedera, hay potencialmente cientos de promotores constitutivos e inducibles que son conocidos y que pueden ser modificados mediante ingeniería genética para funcionar en la célula hospedera.

En algunas ocasiones, pueden emplearse promotores ampliamente utilizados en tecnología recombinante, por ejemplo operones lac y trp de *Escherichia coli*, el promotor tac, el promotor pL de bacteriófago, los promotores T7 y SP6 de bacteriófago, promotor de beta-actina, promotor de insulina, polihedrina baculoviral y promotor p10.

En otras ocasiones, puede emplearse un promotor constitutivo. Ejemplos no limitantes de promotores constitutivos incluyen el promotor int del bacteriófago lambda, el promotor bla de la secuencia de gen de beta-lactamasa de pBR322, hydA o thIA en *Clostridium*, *Streptomyces coelicolor* hrdB, o whiE, el promotor CAT de la secuencia del gen de la cloramfenicol acetil transferasa de pPR325, promotor constitutivo de estafilococo blaZ y similares.

Un promotor útil para la presente invención puede ser también un promotor inducible que regula la expresión del gen corriente abajo de manera controlada, tal como bajo una condición específica del cultivo celular. Ejemplos de promotores procarióticos inducibles incluyen los promotores izquierdo y derecho principales de bacteriófago, los promotores trp, recA, lacZ, AraC, y gal de *E. coli*, los promotores alfa-amilasa (Ulmanen et al., J. Bacteriol. 162:176-182, 1985) y los promotores sigma-D-específicos de *Bacillus subtilis* (Gilman et al., Gene sequence 32:11-20 (1984)), los promotores de los bacteriófagos de *Bacillus* (Gryczan, In: The Molecular

Biology of the Bacilli, Academic Press, Inc., NY (1982)), promotores de *Streptomyces* (Ward et al., Mol. Gen. Genet. 203:468-478, 1986), y similares. Glick (J. Ind. Microbiol. 1:277-282, 1987); Cenatiempo (Biochimie 68:505-516, 1986); y Gottesman (Ann. Rev. Genet. 18:415-442, 1984) hacen una revisión de los promotores procarióticos ejemplares.

Un promotor que es constitutivamente activo bajo ciertas condiciones de cultivo, puede ser inactivo en otras condiciones. Por ejemplo, se sabe que la expresión del promotor del gen *hydA* de *Clostridium acetobutylicum*, está regulada por el pH ambiental. Adicionalmente, también se conocen, y pueden usarse, promotores regulados por la temperatura. Por consiguiente, en algunas realizaciones, dependiendo de la célula hospedera deseada, puede emplearse un promotor regulado por el pH o la temperatura con los constructos de expresión de la invención. Son conocidos otros promotores regulables por el pH, tales como P170 que funciona en las bacterias de ácido láctico, como se revela en la Solicitud de Patente U.S. No. 2002-0137140.

En general, para expresar la secuencia de gen/nucleótido deseada de manera eficiente, pueden emplearse diversos promotores; e.g., el promotor original del gen, promotores de genes de resistencia a antibióticos tales como por ejemplo el gen de resistencia a canamicina de Tn5, el gen de resistencia a ampicilina de pBR322, y promotores del lambda fago, y cualquier promotor que pueda ser funcional en la célula hospedera. Para la expresión, otros elementos regulatorios, tales como por ejemplo una secuencia Shine-Dalgarno (SD) que incluyen secuencias naturales y sintéticas operables en la célula hospedero) y un terminador transcripcional (estructura de repetición invertida que incluye cualquier secuencia natural y sintética) que sea operable en la célula hospedera (en donde la secuencia codificante será introducida para proveer una célula recombinante de esta invención) puede usarse con los promotores anteriormente descritos.

Ejemplos de promotores que pueden emplearse con los productos y procesos de la invención incluyen aquellos revelados en los siguientes documentos de patente: US 2004/0171824, US 6,410,317, WO 2005/024019. Diversos sistemas de promotor-operador, tales como *lac*, (D. V. Goeddel et al., "Expression in *Escherichia*

coli of Chemically Synthesized Genes for Human Insulin," Proc. Nat. Acad. Sci. U.S.A., 76:106-110 (1979)); trp (J. D. Windass et al. "The Construction of a Synthetic *Escherichia coli* Trp Promoter and Its Use In the Expression of a Synthetic Interferon Gene", Nucl. Acids. Res., 10:6639-57 (1982)) y operones λ PL (R. Crowl et al., "Versatile Expression Vectors for High-Level Synthesis of Cloned Gene Products in *Escherichia coli*", Gene, 38:31-38 (1985)) existen en *E. coli* y han sido usados para la regulación de la expresión de genes en células recombinantes. Los reguladores correspondientes son el represor lac, los represores trpR, y cl, respectivamente.

Los represores son moléculas de proteína que se unen de manera específica a operadores particulares. Por ejemplo, la molécula represora lac se une al operador del sistema promotor-operador lac, mientras que el represor cro se une al operador del promotor λP_R . En el arte se conocen otras combinaciones de represor y operador. Ver, e.g., J. D. Watson et al., Molecular Biology Of The Gene, p. 373 (4th ed. 1987). La estructura formada por el represor y el operador bloquea la interacción productiva del promotor asociado con la ARN polimerasa, previniendo de esta manera la transcripción. Otras moléculas, denominadas inductores se unen a los represores evitando de esta manera que el represor se una a su operador. Por lo tanto, puede reversarse la supresión de la expresión de proteína por las moléculas represoras reduciendo la concentración del represor o neutralizando el represor con un inductor.

Los inductores y sistemas de promotor-operador análogos con conocidos en otros microorganismos. En las levaduras, los promotores GAL10 y GAL1 son reprimidos por la glucosa extracelular, y activados mediante la adición de galactosa, un inductor. La proteína GAL80 es un represor para el sistema, y GAL4 es un activador transcripcional. La unión de GAL80 a galactosa evita la unión de GAL80 a GAL4. Por lo tanto, GAL4 puede unirse a una secuencia de activación corriente arriba (UAS, por las siglas en inglés de upstream activation sequence) que activa la transcripción. Ver Y. Oshima, "Regulatory Circuits for Gene Expression: The Metabolisms Of Galactose And Phosphate" in The Molecular Biology Of The Yeast *Sacharomyces*, Metabolism And Gene Expression, J. N. Strathern et al. eds. (1982).

La transcripción bajo el control del promotor PHO5 es reprimida por el fosfato inorgánico extracelular, e inducida a un alto nivel cuando el fosfato es reducido. R. A. Kramer y N. Andersen, "Isolation of Yeast Genes with mRNA Levels Controlled By Phosphate Concentration," *Proc. Nat. Acad. Sci. U.S.A.*, 77:6451-6545 (1980). Se han identificado varios genes reguladores para la expresión de PHO5, incluyendo algunos involucrados en la regulación de fosfato.

Mat α 2 es un sistema promotor en levadura regulado por la temperatura. En este sistema se han identificado una proteína represora, sitios promotores y operador. A. Z. Sledziewski et al., "Construction Of Temperature-Regulated Yeast Promoters Using The Mat α 2 Repression System," *Bio/Technology*, 6:411-16 (1988).

Otro ejemplo de un sistema represor en levadura es el promotor CUP1, el cual puede ser inducido por iones Cu²⁺. El promotor CUP1 es regulado por una proteína metalotionina. J. A. Gorman et al., "Regulation of The Yeast Metallothioneine Gene," *Gene*, 48:13-22 (1986).

De manera similar, para obtener la expresión deseada de una o más celulasas, puede utilizarse un mayor número de copias de plásmido en un producto o proceso de la invención. Pueden prepararse constructos para la integración cromosómica de los genes deseados. La integración cromosómica de los genes foráneos puede ofrecer muchas ventajas sobre las construcciones basadas en plásmidos, teniendo estas últimas ciertas limitaciones para procesos comerciales. Se han integrado cromosómicamente genes etanológicos en *E. coli* B; ver Ohta et al. (1991) *Appl. Environ. Microbiol.* 57:893-9. En general, esto se logra mediante la purificación de un fragmento de ADN que contiene (1) los genes deseados corriente arriba de un gen de resistencia a antibióticos y (2) un fragmento de ADN homólogo del microorganismo objetivo. Este ADN puede ligarse para formar círculos sin replicones y usarse para transformación. Por consiguiente, el gen de interés puede introducirse en un hospedero heterólogo tal como *E. coli*, y pueden aislarse cortos fragmentos aleatorios y unirse operablemente a los genes objetivo (e.g., genes que codifican enzimas celulasas) para promover la recombinación homóloga.

Vectores de Expresión

Los vectores de expresión pueden comprender cualquier casete de expresión aquí descrito, e incluir de manera típica todos los elementos requeridos para la expresión de uno o más polinucleótidos de interés en una célula hospedera. En algunas realizaciones, se introduce un polinucleótido de interés en un vector para crear un vector de expresión recombinante adecuado para la transformación de una célula hospedera para la producción de un combustible en un microorganismo recombinante. En otras realizaciones, puede introducirse un casete de expresión en un vector para crear un vector de expresión recombinante adecuado para la transformación de una célula hospedera. En algunas realizaciones, se proveen vectores de expresión que comprenden uno o más casetes de expresión.

Los vectores de expresión pueden replicarse de manera autónoma, o pueden replicarse al ser insertados en el genoma de la célula hospedera. En algunas realizaciones, puede integrarse de manera homóloga un casete de expresión en el genoma de la célula hospedera. En otras realizaciones, los genes pueden integrarse de manera no homóloga en el genoma de una célula hospedera. En algunas realizaciones, puede integrarse el casete de expresión en un locus deseado mediante recombinación homóloga doble.

En algunas realizaciones, puede ser deseable que un vector pueda ser usado en más de una célula hospedera. Por ejemplo, puede usarse un vector para clonación en *E. coli* y para expresión en una especie de *Clostridium*. Tal vector incluirá de manera típica un origen de replicación de *E. coli* y un origen compatible con *Clostridium* u otra bacteria Gram positiva. Se conocen diversos orígenes de replicación de plásmidos Gram positivos y de *E. coli*. Elementos adicionales del vector pueden incluir, por ejemplo, marcadores seleccionables, e.g., resistencia a canamicina o resistencia a ampicilina, lo que permite la detección y/o la selección de aquellas células transformadas con las secuencias de polinucleótido deseadas. Vectores transportadores de *Clostridium* ejemplares se describen en Mauchline et al. (1999) en: *Clostridia: Manual of Industrial Microbiology and Biotechnology*, AL Demain y JE Davies, ed. (ASM Press), pp. 475-492; y Heap et al., *J. Microbiol. Methods*, 78:79-85 (2009).

En algunas realizaciones, el vector de expresión puede incluir uno o más genes cuya presencia y/o expresión permita la tolerancia de una célula hospedera a concentraciones de etanol económicamente relevantes. Por ejemplo, pueden incluirse genes tales como *omrA*, *lmrA*, e *lmrCD* en el vector de expresión. Se ha mostrado que *OmrA* de la bacteria de ácido láctico en vino *Oenococcus oeni* y su homólogo *LmrA* de *Lactococcus lactis* incrementan la resistencia relativa de *tolC(-)* *E. coli* de 100 a 10.000 veces (Bourdineaud et al., A bacterial gene homologous to ABC transporters protect *Oenococcus oeni* from ethanol and other stress factors in wine. *Int. J. Food Microbiol.* 2004 Apr 1;92(1):1-14). Por lo tanto, puede resultar benéfico incorporar *omrA*, *lmrA*, y otros homólogos para incrementar la tolerancia a etanol de una célula hospedera.

En algunas realizaciones, los vectores que se proveen aquí pueden incluir uno o más segmentos de ácido nucléico genómico para facilitar la integración dirigida en el genoma del organismo hospedero. Un segmento de ácido nucléico genómico para integración dirigida puede ser de aproximadamente diez nucleótidos a aproximadamente 20.000 nucleótidos de largo. En algunas realizaciones, un segmento de ácido nucléico genómico para integración dirigida puede ser de aproximadamente 1.000 a aproximadamente 10.000 nucleótidos de largo. En otras realizaciones, un segmento de ácido nucléico genómico para integración dirigida es de entre aproximadamente 1 kb a aproximadamente 2 kb de largo. En algunas realizaciones, una pieza "contigua" de ácido nucléico nuclear genómico puede dividirse en dos piezas adyacentes cuando los genes de interés se clonan en la región no codificante del ADN contiguo. Esto permite la integración de la región de ácido nucléico intermedia en el cromosoma bacteriano mediante una recombinación cruzada doble. En otras realizaciones, las piezas adyacentes pueden comprender segmentos de secuencias de ácido nucléico nuclear que no son contiguos entre sí. En algunas realizaciones, un primer segmento adyacente de ácido nucléico genómico se localiza entre aproximadamente 0 a aproximadamente 10.000 pares de bases lejos de un segundo segmento adyacente de ácido nucléico genómico en el genoma nuclear.

En algunas realizaciones, pueden introducirse segmentos de ácidos nucléicos genómicos en un vector para generar un vector de expresión principal para la integración dirigida de cualquier casete de expresión aquí revelado en el genoma

nuclear del organismo hospedero. Puede usarse cualquiera de una variedad de métodos conocidos en el arte para introducir secuencias de ácidos nucleicos. Por ejemplo, pueden amplificarse segmentos de ácido nucleico a partir de ácido nucleico nuclear genómico aislados usando cebadores apropiados y PCR. Los productos amplificados pueden luego ser introducidos en cualquiera de una variedad de vectores de clonación adecuados mediante, por ejemplo, ligación. Algunos vectores útiles incluyen, por ejemplo sin limitación, pGEM13z, pGEMT y pGEMTEasy (Promega, Madison, WI); pSTBlueI (EMD Chemicals Inc. San Diego, CA); y pcDNA3.1, pCR4-TOPO, pCR-TOPO-II, pCRBlunt-II-TOPO (Invitrogen, Carlsbad, CA). En algunas realizaciones, al menos un segmento de ácido nucleico de un núcleo se introduce en un vector. En otras realizaciones, dos o más segmentos de ácido nucleico de un núcleo se introducen en un vector. En algunas realizaciones, los dos segmentos de ácido nucleico pueden ser adyacentes el uno al otro en el vector. En algunas realizaciones, los dos segmentos de ácido nucleico introducidos en un vector pueden estar separados por, por ejemplo, entre aproximadamente uno a treinta pares de bases. En algunas realizaciones, las secuencias que separan los dos segmentos de ácido nucleico pueden contener por lo menos un sitio de reconocimiento de restricción por endonucleasa.

En diversas realizaciones, pueden incluirse secuencias regulatorias en los vectores de la presente invención. En algunas realizaciones, las secuencias regulatorias comprenden secuencias de ácido nucleico para regular la expresión de genes (e.g., un gen de interés) introducidos en el genoma nuclear. En diversas realizaciones, pueden introducirse las secuencias regulatorias en un vector de expresión principal. Por ejemplo, pueden identificarse diversas secuencias regulatorias a partir del genoma del microorganismo hospedero. Las secuencias regulatorias pueden comprender, por ejemplo, un promotor, un potenciador, un intron, un exon, un UTR 5', un UTR 3', o cualquier porción de los mismos de cualquiera de los anteriores, de un gen nuclear. Usando técnicas estándar de biología molecular, pueden introducirse las secuencias regulatorias en el vector deseado. En algunas realizaciones, los vectores comprenden un vector de clonación o un vector que comprende segmentos de ácido nucleico para la integración dirigida.

En algunas realizaciones, las secuencias de ácido nucleico para regular la expresión de genes introducidos en el genoma nuclear pueden introducirse en un

vector mediante amplificación por PCR de un UTR 5', UTR 3', un promotor y/o un potenciador, o porción del mismo, uno o más genes nucleares. Usando condiciones de ciclación de PCR adecuadas, se emplean cebadores adyacentes a las secuencias a ser amplificadas para amplificar las secuencias reguladoras. En algunas realizaciones, los cebadores pueden incluir secuencias de reconocimiento para cualquiera de una variedad de enzimas de restricción, introduciendo de esta manera aquellas secuencias de reconocimiento en los productos de amplificación por PCR. Los productos de PCR pueden someterse a digestión con las enzimas de restricción apropiadas e introducirse en los sitios correspondientes de un vector.

En otras realizaciones, uno o más genes a ser expresados pueden ser integrados en el genoma del microorganismo usando sistemas disponibles comercialmente o métodos similares. Se ha demostrado la aplicabilidad de estos métodos e *Clostridia*, incluyendo la integración y expresión de un gen foráneo en una célula de *Clostridium* (ver, e.g., Heap et al. (2007). J. Microbiol. Methods. 70:452-464; Chen et al. (2007). Plasmid. 58:182-189).

Microorganismos Hospederos

Algunas realizaciones se relacionan con microorganismos que contienen cualquiera de los polinucleótidos, casetes de polinucleótido, casetes de expresión, o vectores expresión aquí descritos. Las células hospederas pueden incluir, pero no se limitan a, células eucariotas, tales como células animales, células de insectos, células fúngicas, y levaduras, y células procariotas, tales como bacterias. En algunas realizaciones, el hospedero es *C. phytofermentans*. En algunas realizaciones, un organismo hospedero potencial puede comprender un organismo recombinante.

En algunas realizaciones, el microorganismo recombinante puede ser un microorganismo celulolítico o sacarolítico. En ciertas realizaciones, el microorganismo puede ser *Clostridium cellulovorans*, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium josui*, *Clostridium papyrosolvens*, *Clostridium cellobioparum*, *Clostridium hungatei*, *Clostridium cellulosi*, *Clostridium stercorarium*, *Clostridium termitidis*, *Clostridium thermocopriae*, *Clostridium celerecrescens*, *Clostridium polysaccharolyticum*, *Clostridium populeti*, *Clostridium lentocellum*,

Clostridium chartatabidum, *Clostridium aldrichii*, *Clostridium herbivorans*, *Acetivibrio cellulolyticus*, *Bacteroides cellulosolvens*, *Caldicellulosiruptor saccharolyticum*, *Ruminococcus albus*, *Ruminococcus flavefaciens*, *Fibrobacter succinogenes*, *Eubacterium cellulosolvens*, *Butyrivibrio fibrisolvens*, *Anaerocellum thermophilum*, *Halocella cellulolytica*, *Thermoanaerobacterium thermosaccharolyticum* o *Thermoanaerobacterium saccharolyticum*.

En algunas realizaciones, puede seleccionarse un microorganismo hospedero, por ejemplo, a partir de las amplias categorías de bacterias Gram-negativas, tales como especies de *Xanthomonas*, y bacterias Gram-positivas, incluyendo miembros del género *Bacillus*, tal como *B. pumilus*, *B. subtilis* y *B. coagulans*; *Clostridium*, por ejemplo, *C. acetobutylicum*, *C. aerotolerans*, *C. thermocellum*, *C. thermohydrosulfuricum* y *C. thermosaccharolyticum*; especies de *Cellulomonas* como *Cellulomonas uda*; y *Butyrivibrio fibrisolvens*. Adicional a *E. coli* pueden usarse, por ejemplo, otras bacterias entéricas del género *Erwinia*, como *E. chrysanthemi*, y *Klebsiella*, como *K. planticola* y *K. oxytoca*. En algunas realizaciones, el microorganismo hospedero puede ser *Zymomonas mobilis*. Organismos hospederos similarmente aceptables son diversas levaduras, ejemplificadas por especies de *Cryptococcus* como *Cr. albidus*, especies de *Monilia*, *Pichia stipitis* y *Pullularia pullulans*, y *Saccharomyces cerevisiae*; y otras bacterias metabolizantes de oligosacáridos, incluyendo pero sin limitarse a *Bacteroides succinogenes*, especies de *Thermoanaerobacter* como *T. ethanolicus*, especies de *Thermoanaerobium* tales como *T. brockii*, especies de *Thermobacteroides* como *T. acetoethylicus*, y especies del género *Ruminococcus* (por ejemplo, *R. flavefaciens*), *Thermonospora* (tal como *T. fusca*) y *Acetivibrio* (por ejemplo, *A. cellulolyticus*). En algunas realizaciones, puede seleccionarse un organismo hospedero, por ejemplo, de un alga tal como, por ejemplo, *Amphora*, *Anabaena*, *Anikstrodesmis*, *Botryococcus*, *Chaetoceros*, *Chlorella*, *Chlorococcum*, *Cyclotella*, *Cylindrotheca*, *Dunaliella*, *Euglena*, *Hematococcus*, *Isochrysis*, *Monoraphidium*, *Nannochloris*, *Nannochloropsis*, *Navicula*, *Nephrochloris*, *Nephroselmis*, *Nitzschia*, *Nodularia*, *Nostoc*, *Oochromonas*, *Oocystis*, *Oscillatoria*, *Pavlova*, *Phaeodactylum*, *Playtmonas*, *Pleurochrysis*, *Porphyra*, *Pseudoanabaena*, *Pyramimonas*, *Stichococcus*, *Synechococcus*, *Tetraselmis*, *Thalassiosira*, *Trichodesmium*. La literatura que se relaciona con microorganismos que cumplen con el criterio objeto se refleja, por ejemplo, en Biely, *Trends in Biotech.* 3: 286-90 (1985), en Robsen et al, *Enzyme Microb. Technol.* 11: 626-44 (1989), y en Beguin *Ann. Rev. Microbiol.* 44: 219-48 (1990), cada uno de los cuales se incorpora

aquí por referencia en su totalidad. La metodología de transformación apropiada se encuentra disponible para cada uno de estos diferentes tipos de hospederos y se describe en detalle más abajo. Ver también, e.g., Brat et al. *Appl. Env. Microbiol.* 29: 75:2304-2311, que revela la expresión de xilosa isomerasa en *Saccharomyces cerevisiae*.

En algunas realizaciones, puede seleccionarse un microorganismo hospedero mediante, por ejemplo, su habilidad para producir las proteínas necesarias para transportar un oligosacárido dentro de la célula y sus niveles intracelulares de enzimas que metabolizan aquellos oligosacáridos. Ejemplos de tales microorganismos incluyen bacterias entéricas como *E. chrysanthemi* y otras especies de *Klebsiella* y *Erwinia* tales como *K. oxytoca*, que produce naturalmente una β -xilosidasa, y *K. planticola*. Ciertas *E. coli* son hospederos atractivos porque transportan y metabolizan celobiosa, maltosa y/o maltotriosa. Ver, por ejemplo, Hall et al., *J. Bacteriol.* 169: 2713-17 (1987).

En algunas realizaciones, puede seleccionarse un microorganismo hospedero mediante selección para determinar si el microorganismo evaluado transporta y metaboliza oligosacáridos. Tal selección puede lograrse de diversas maneras. Por ejemplo, puede hacerse una selección de microorganismos para determinar cuáles crecen en sustratos de oligosacáridos adecuados, siendo la selección diseñada para escoger aquellos microorganismos que no transportan solamente monómeros al interior de la célula. Ver, por ejemplo, Hall et al. (1987), *supra*. De manera alternativa, puede evaluarse en los microorganismos la actividad enzimática intracelular apropiada, e.g., actividad β -xilosidasa. Puede hacerse una selección adicional en el crecimiento de microorganismos hospederos portenciales con base en la tolerancia a etanol, tolerancia a sal, y tolerancia a la temperatura. Ver Alterhum et al., *Appl. Environ. Microbiol.* 55: 1943-48 (1989); Beall et al, *Biotechnol. & Bioeng.* 38:296-303 (1991).

En algunas realizaciones, un microorganismo hospedero puede exhibir una o más de las siguientes características: la habilidad para crecer en concentraciones de etanol por encima del 1%, la habilidad para tolerar niveles de sal de, por ejemplo, 0.3 molar, la habilidad para tolerar niveles de acetato de, por ejemplo, 0.2 molar, y la

habilidad para tolerar temperaturas de, por ejemplo, 40°C, y la habilidad para producir mayores niveles de enzimas útiles para la despolimerización de celulosa, hemicelulosa y pectina con actividad de proteasa mínima. En algunas realizaciones un microorganismo hospedero puede contener también xilanasas o celulasas nativas. En algunas realizaciones, después de la introducción de vectores de expresión para producción de combustible, un hospedero puede producir etanol a partir de diversos sacáridos evaluados con rendimientos teóricos mayores a, por ejemplo, 90% reteniendo a la vez uno o más de los rasgos útiles arriba mencionados.

Transformación de Células Hospederas

Algunas realizaciones se relacionan con métodos para introducir cualquiera de los polinucleótidos, casetes de polinucleótidos, casetes de expresión, y vectores de expresión aquí descritos en una célula de un microorganismo hospedero. Tales realizaciones producen por lo tanto un microorganismo recombinante que es capaz de producir un combustible al ser cultivado bajo una variedad de condiciones de fermentación. Los métodos para transformar células son bien conocidos en el arte, y pueden incluir, por ejemplo, electroporación, lipofección, transfección, conjugación, transformación química, inyección, bombardeo con pistola de flujo entrante de partículas, y magnetoforesis. La magnetoforesis emplea magnetoforesis y fabricación por nanotecnología de magnetos lineales de tamaño micro para introducir ácidos nucleicos en células (Kuehnle *et al.*, Patente U.S. No. 6,706,394; 2004; Kuehnle *et al.*, Patente U.S. No. 5,516,670; 1996). En algunas realizaciones, puede llevarse a cabo la electrotransformación de plásmidos metilados en *C. phytofermentans* de acuerdo con un protocolo desarrollado por Mermelstein (Mermelstein, *et al.* Bio/Technology 10:190-195 (1992)). Más métodos pueden incluir transformación por conjugación. En otras realizaciones, pueden aislarse transformantes positivos en CGM solidificado con agar suplementado con el antibiótico apropiado.

En diversas realizaciones, los métodos de transformación pueden acoplarse con uno o más métodos para la visualización o cuantificación de la introducción de ácido nucleico a uno o más microorganismos. Además, se enseña que esto puede acoplarse con la identificación de cualquier línea que muestre una diferencia estadística en, por ejemplo, crecimiento, fluorescencia, metabolismo de carbono, flujo de isoprenoide, o contenido de ácidos grasos con respecto al fenotipo inalterado. Los métodos de transformación pueden también acoplarse con la

visualización o cuantificación de un producto resultante de la expresión de los ácidos nucleicos introducidos.

De manera típica, previo a la transformación en *C. phytofermentans*, pueden metilarse vectores que comprenden el ADN de plásmido para prevenir la restricción por endonucleasas de *Clostridium*. (Mermelstein y Papoutsakis. Appl. Environ. Microbiol. 59: 1077-1081 (1993)). En algunas realizaciones, puede lograrse la metilación mediante la phi3TI metiltransferasa. En realizaciones adicionales, puede transformarse el ADN de plásmido en el vector de *E. coli* pDHKM que alberga DH10β (Zhao, *et al.* Appl. Environ. Microbiol. 69: 2831-41 (2003)) portando una copia activa del gen de metiltransferasa phi3TI.

De manera general, puede permitirse que crezcan cepas de *C. phytofermentans* de manera anaeróbica en Medio de Crecimiento de *Clostridium* (CGM) a 37°C suplementado con un antibiótico apropiado, tal como 40 µg/ml de eritromicina/cloranfenicol o 25 µg/ml de tiamfenicol (Hartmanis y Gatenbeck. Appl. Environ. Microbiol. 47: 1277-83 (1984)). Adicionalmente, las cepas de *C. phytofermentans* pueden ser cultivadas en fermentaciones en contenedor cerrado de lotes de 100 mL de CGM suplementado con el antibiótico apropiado a 37°C en una cámara anaeróbica FORMA SCIENTIFIC™ (THERMO FORMA™, Marietta, Ohio).

En otras realizaciones, puede cultivarse, *C. phytofermentans* de acuerdo con las técnicas de Hungate (Hungate, R. E. (1969). A roll tube method for cultivation of strict anaerobes. Methods Microbiol 3B, 117-132). Puede usarse el medio GS-2C para el enriquecimiento, aislamiento y el cultivo de rutina de cepas de *C. phytofermentans*, y puede derivarse de GS-2 de Johnson *et al* (Johnson, E. A., Madia, A. & Demain, A. L. (1981). Chemically defined minimal medium for growth of the anaerobic cellulolytic thermophile *Clostridium thermocellum*. Appl Environ Microbiol 41, 1060-1062). GS-2C puede contener lo siguiente: 6.0 g/L de celulosa molida en esferas (Leschine, S. B. & Canale-Parola, E. (1983). Mesophilic cellulolytic Clostridia from freshwater environments. Appl Environ Microbiol 46, 728-737.); 6.0 g/L de extracto de levadura; 2.1 g/L de urea; 2.9 g/L de K₂HPO₄; 1.5 g/L de KH₂PO₄; 10.0 g/L de MOPS; 3.0 g/L de citrato de trisodio dihidratado; 2.0 g/L de clorhidrato de cisteína; 0.001 de g/L de resazurin; con el pH ajustado a 7.0. Pueden incubarse

caldos de cultivo en una atmósfera de N₂ libre de O₂ a 30°C. Pueden incubarse los cultivos en placas de medio agar a temperatura ambiente en una atmósfera de N₂/CO₂/H₂ (83:10:7) en una cámara anaeróbica (Coy Laboratory Products).

Crecimiento, Expresión, y Producción de Combustible

Algunas realizaciones se relacionan con la producción de combustible empleando cualquier microorganismo recombinante aquí descrito. En algunas realizaciones, pueden usarse uno o más microorganismos recombinantes diferentes en combinación para producir combustible. Tales combinaciones pueden incluir más de un tipo diferente de microorganismo recombinante en una sola reacción de fermentación. Otras combinaciones pueden incluir uno o más tipos diferentes de microorganismos recombinantes usados en pasos secuenciales de un proceso para producir combustible a partir de biomasa. En algunas realizaciones, puede usarse un único microorganismo recombinante para producir combustible a partir de biomasa. En algunas realizaciones, puede usarse un microorganismo recombinante para catalizar la producción de productos tales como sacáridos y polisacáridos a partir de lignocelulosa y otros sustratos.

En algunas realizaciones, puede cultivarse un microorganismo recombinante bajo condiciones adecuadas para la expresión de genes a partir de casetes de expresión allí contenidos para la producción de combustible. En ciertas realizaciones, pueden variar las condiciones de incubación dependiendo del microorganismo hospedero empleado. En algunas realizaciones, pueden variar las condiciones de incubación de acuerdo con el tipo de elemento regulador que pueda asociarse con casetes de expresión. Por ejemplo, un organismo recombinante que contiene un casete de expresión que comprende un promotor inducible unido a un ácido nucleico puede requerir la adición de un agente particular al medio de cultivo para la expresión del ácido nucleico.

En otras realizaciones, el microorganismo recombinante puede ser una cepa de *C. phytofermentans* empleada para fermentar un amplio espectro de materiales en combustibles con alta eficiencia como se describe en la solicitud copendiente U.S. No. 2007/0178569 y en la Solicitud de Patente Provisional U.S. No. 61/032,048,

presentadas en Febrero 28, 2008; ambas referencias incorporadas aquí de manera expresa en su totalidad. En algunas realizaciones, la cepa de *C. phytofermentans* puede ser la 700394^T de American Type Culture Collection.

En algunas realizaciones, el proceso empleado para fermentar un sustrato (e.g., materia prima lignocelulósica) pueden incluir: (1) proveer un material derivado de biomasa pretratado que comprende un polisacárido de planta (en donde el pretratamiento puede ser cortar, picar, moler o similares); (2) inocular el material derivado de biomasa pretratado con un primer cultivo que comprende un microorganismo anaeróbico celulolítico (e.g., un microorganismo aquí revelado) en presencia de oxígeno para generar un caldo aerobico, en donde el microorganismo anaeróbico es capaz de hidrolizar por lo menos parcialmente el polisacárido de planta; y (3) fermentar el caldo anaeróbico inoculado hasta que una porción del polisacárido de planta se haya convertido en etanol. En otras realizaciones, el proceso empleado para fermentar un sustrato puede incluir: (1) proveer un material derivado de biomasa pretratado que comprende un polisacárido de planta (en donde el pretratamiento puede ser cortar, picar, moler o similares); (2) inocular el material derivado de biomasa pretratado con un primer cultivo que comprende un microorganismo aeróbico celulolítico (e.g., un microorganismo aquí revelado) en presencia de oxígeno para generar un caldo aeróbico, en donde el microorganismo aeróbico es capaz de hidrolizar al menos parcialmente el polisacárido de planta; (3) incubar el caldo aeróbico hasta que el microorganismo aeróbico celulolítico consuma al menos una porción del oxígeno e hidrolice al menos una porción del polisacárido de planta, convirtiendo de esta manera el caldo aeróbico en un caldo anaeróbico que comprende un hidrolizado que comprende azúcares fermentables; (4) inocular el caldo anaeróbico con un segundo cultivo que comprende un microorganismo anaeróbico (e.g., un microorganismo aquí revelado) capaz de convertir los azúcares fermentables en etanol; y (5) fermentar el caldo anaeróbico inoculado hasta que una porción de los azúcares fermentables se haya convertido en etanol.

La eficiencia de una fermentación puede medirse de diferentes maneras, por ejemplo pueden medirse cambios en la eficiencia en comparación con un organismo de tipo nativo. Además, pueden medirse los cambios en la eficiencia como la relación de producción de un combustible a partir de un sustrato, tal como celulosa, por unidad de tiempo entre un organismo recombinante y un organismo de tipo nativo.

En algunas realizaciones, los cambios en la eficiencia entre un organismo recombinante y un organismo de tipo nativo pueden ser más de 1%, más de 5%, más de 10%, más de 15%, más de 20%, más de 25%, más de 30%, más de 35%, más de 40%, más de 45%, más de 50%, más de 55%, más de 60%, más de 65%, más de 70%, más de 75%, más de 80%, más de 85%, más de 90%, más de 95%, más de 100%, y más de 200%.

En el arte se conocen diversos medios para el crecimiento de una variedad de microorganismos. El medio de crecimiento puede ser mínimo y/o definido, o completo y/o complejo. Fuentes de carbono fermentables pueden incluir materia prima pretratada o no pretratada que contiene material celulósico, hemicelulósico, y/o lignocelulósico tal como, aserrín, harina de madera, pulpa de madera, pulpa de papel, corriente de desechos de pulpa de papel, pastos, tales como, césped de pradera, plantas y cultivos de biomasa, tales como, col marítima, algas, cascarilla de arroz, bagazo, yute, hojas, hierba cortada, forraje de maíz, olote de maíz, granos de maíz, maíz molido, granos de destilerías, y pectina.

Los nutrientes adicionales pueden estar presentes en una reacción de fermentación, incluyendo compuestos que contienen nitrógeno tales como aminoácidos, proteínas, hidrolizados de proteínas, amoníaco, úrea, nitrato, nitrito, soya, derivados de soya, caseína, derivados de caseína, polvo de leche, derivados de leche, suero, extracto de levadura, levadura hidrolizada, levadura autolizada, licor de maceración del maíz, sólidos de maceración del maíz, glutamato monosódico, y/o otras fuentes de fermentación de nitrógeno, vitaminas, y/o suplementos minerales. En algunas realizaciones, pueden adicionarse o estar presentes una o más fuentes de carbono de bajo peso molecular adicionales tales como glucosa, sacarosa, maltosa, jarabe de maíz, ácido láctico, etc. En algunas realizaciones, una forma posible de medio de crecimiento puede ser caldo Luria-Bertani (LB) modificado (con 10 g de Difco triptona, 5 g de extracto de levadura Difco, y 5 g de cloruro de sodio por litro) como describe Miller J. H. (1992).

Puede observarse producción mejorada de combustible luego de transformar las células hospedero competentes para producir combustible con los vectores de expresión aquí descritos y de permitir que crezcan microorganismos recombinantes bajo condiciones adecuadas. Puede observarse la producción mejorada de

combustible mediante métodos estándar conocidos por aquellos versados en la materia.

En algunas realizaciones, puede llevarse a cabo el crecimiento y la producción de los microorganismos recombinantes aquí revelados en lotes de fermentación normales, fermentaciones de alimentación de lote o fermentaciones continuas. En ciertas realizaciones, es deseable llevar a cabo fermentaciones bajo condiciones de oxígeno reducido o anaeróbicas para ciertos hospederos. En otras realizaciones, la producción de combustible puede llevarse a cabo con niveles de oxígeno suficientes para permitir el crecimiento de organismos aeróbicos; y, opcionalmente con el uso de fermentadores desplazados por aire o equivalentes. En algunas realizaciones, los microorganismos recombinantes crecen usando lotes de cultivos. En algunas realizaciones, se permite que los microorganismos recombinantes crezcan empleando fermentación en bioreactores. En algunas realizaciones, se cambia el medio de crecimiento en el crecen los microorganismos recombinantes, permitiendo de esta manera niveles incrementados de producción de combustible. El número de cambios de medio puede variar.

El pH de la fermentación puede ser suficientemente alto para permitir el crecimiento y la producción de combustible por el hospedero. El ajuste del pH del caldo de fermentación puede llevarse a cabo usando agentes neutralizantes tales como carbonato de calcio o hidróxidos. La selección e incorporación de cualquiera de los anteriores métodos de fermentación es altamente dependiente de la cepa del hospedero y el proceso utilizado corriente abajo.

En algunas realizaciones, pueden purificarse solventes orgánicos a partir de biomasa fermentada con *C. phytofermentans* mediante diversos medios. En ciertas realizaciones, se purifican solventes orgánicos mediante destilación. En realizaciones ejemplares, puede destilarse etanol al 96% aproximadamente a partir de la mezcla fermentada. En realizaciones adicionales, puede obtenerse etanol grado combustible, a saber etanol al 99-100% aproximadamente, mediante destilación azeotrópica de etanol al 96% aproximadamente. La destilación azeotrópica puede lograrse mediante la adición de benceno a etanol al 96% aproximadamente y destilando nuevamente la mezcla. De manera alternativa, puede pasarse etanol al 96% aproximadamente a través de un tamiz molecular para remover el agua.

En algunas realizaciones, los métodos para producir el combustible pueden incluir cultivar cualquier microorganismo aquí descrito y suministrar una proteína expresada por un polinucleótido, casete de polinucleótidos, casete de expresión, vector de expresión que comprende cualquier ácido nucleico codificado por un gen predicho identificado en *C. phytofermentans* aquí descrito al medio de cultivo. En realizaciones particulares, el ácido nucleico puede codificar una hidrolasa. En ciertas realizaciones, pueden suministrarse proteínas aisladas a un medio de cultivo.

Los siguientes ejemplos se presentan a manera de ilustración y no a manera de limitación.

EJEMPLOS

Ejemplo 1. Identificación de Secuencias de ADN en *C. phytofermentans*

Construcción, aislamiento y secuenciación de bibliotecas de insertos. Se secuenció el ADN genómico utilizando una estrategia aleatoria convencional del genoma. En resumen, se aislaron al azar fragmentos de ADN de 2-3 kb después del cizallamiento mecánico. Estos fragmentos extraídos con gel se concentraron, se repararon sus extremos y se clonaron en pUC18. Se llevaron a cabo reacciones de secuenciamiento del plásmido de doble extremo utilizando PE BigDye™ Terminator chemistry (Perkin Elmer) y se resolvieron escaleras de secuenciación Secuenciadores de ADN Automatizados PE 3700. Se realizó una ronda (x lecturas) de la secuenciación de una biblioteca de pequeños insertos, generando una redundancia de x veces.

Montaje de secuencia y cierre de brecha. Se procesaron trazas de secuencias con Phred43, 44 para lectura automática de nucleótidos y evaluación de calidad de los datos antes del montaje con Phrap (P. Green, University of Washington, Seattle, Washington, EE.UU.) y visualización con Consed45.

Anotación y análisis de secuencias. El modelamiento genético se realizó

utilizando los paquetes de modelamiento Critica47, Glimmer48 y Generation (compbio.ornl.gov/generation/index.shtml), se combinaron los resultados y se condujo una búsqueda con herramienta básica de búsqueda de alineación local para proteínas (BLASTP) de las traducciones frente a la base de datos no redundante (NR) de GenBank. Se empleó la alineación del N terminal de cada modelo de genes versus la mejor coincidencia NR para elegir un modelo de genes. Si no se encuentra ninguna coincidencia BLAST, se mantiene el modelo Critica. Se señalaron los modelos de genes que se superponían en más del 10% de su longitud, dando preferencia a los genes con una coincidencia BLAST. Se buscó el conjunto gen/proteína revisado contra KEGG GENES, InterPro (que incorpora Pfam, TIGRFams, SmartHMM, PROSITE, PRINTS y ProDom) y bases de datos de agrupamientos de grupos ortólogos de proteínas (COG, por las siglas en inglés de Clusters of Orthologous Groups), además de BLASTP versus NR. A partir de estos resultados, se desarrollaron categorizaciones utilizando las jerarquías KEGG y COG. Los criterios iniciales para la asignación funcional automatizada requirieron un mínimo de identidad de residuos del 50% sobre el 80% de la extensión de la coincidencia para alineaciones BLASTP, además de evidencia concurrente de los métodos de patrón o perfil. Se hicieron asignaciones teóricas para las identidades de hasta 30%, sobre 80% de la extensión.

Usando BLASTP, se realizó una búsqueda para cada gen de *C. phytofermentans* contra todos los genes de los genomas secuenciados y se extrajo el primer blast de cada proteína predicha. El análisis de la localización subcelular teórica y los sitios señal para ruptura de péptidos se llevó a cabo utilizando PSORT (psort.hgc.jp/form.html). Los dominios CAZy se anotaron mediante CAzy ((enzimas de carbohidrato-activas, www.cazy.org)). Los transportadores se anotaron usando TransportDB (www.membranetransport.org). La secuencia completa de *C. phytofermentans* se puso a disposición del público en agosto del 2007 (número de acceso NC_010001).

Ejemplo 2. Análisis de Expresión de secuencias de ADN en *C. phytofermentans*

Diseño de microarreglos. El diseño de microarreglos Affymetrix a la medida de *C. phytofermentans* (Figura 3) permite la medición del nivel de expresión de todos los marcos de lectura abierta (ORFs, por las siglas en inglés de Open Reading Frames) identificados, la estimación de las regiones 5' y 3' de mRNA no traducidas, la determinación de operones, el descubrimiento del tARN, y la discriminación entre

modelos alternativos de genes (que difieren fundamentalmente en la selección del codón de inicio).

Se identificaron secuencias de codificación de proteínas teóricas usando los programas de predicción GeneMark™ (Besemer, J., y M. Borodovsky. 2005 GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acid Res 33:W451-4) y Glimmer (Delcher, A. L., K. A. Bratke, C.E. Powers, y S.L. Salzberg. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673-9). La unión de estas dos predicciones se utilizó como el conjunto de expresión. Si dos proteínas diferían en su región N-terminal, la más pequeña de las dos proteínas se usaba para el análisis de transcriptos, pero la región extendida era representada por sondas con el fin de definir la N-terminal real. Este arreglo resultó en la inclusión de todas las proteínas representadas en el registro GenBank e incluyó ORFs adicionales que no se encontraban en el registro GenBank. Se siguieron protocolos estándar de diseño de arreglos Affymetrix para asegurar que cada sonda fuera única con el fin de minimizar la hibridación cruzada. El diseño de arreglos se implementó en un GeneChip™ Affymetrix de formato 49-5241arreglado con características11 u.

Crecimiento del cultivo celular y el aislamiento del ARN. Se cultivó *C.phytofermentans* en tubos o matraces erlenmeyer de 500 mL a 30°C bajo 100% de N₂ en medio GS2 suplementado con 0,3% (p/v) con una de catorce fuentes de carbono específicas (glucosa; xilano; celobiosa; celulosa; D-arabinosa; L-arabinosa; fucosa; galactosa; laminarina; manosa; pectina; ramnosa; xilosa; o extracto de levadura). El crecimiento se determinó por espectrofotometría monitoreando cambios en la densidad óptica a 660 nm.

Se purificó ARN a partir de cultivos en fase exponencial media (OD₆₆₀ = 0.5). Las muestras de 1 ml se congelaron rápidamente mediante inmersión en nitrógeno líquido. Las células se recolectaron mediante centrifugación durante 5 minutos a 8.000 rpm a 4°C, y el ARN total se aisló utilizando Qiagen RNeasy™ Mini Kit y tratamiento con DNasa I libre de RNAsa. La concentración de ARN se determinó mediante la absorbancia a 260/280 nm usando un espectrofotómetro Nanodrop™.

Procesamiento de microarreglos. La síntesis de cADN, la hibridación y la obtención de imágenes del arreglo se realizaron en las instalaciones centrales de genómica del Centro Médico de la Universidad de Massachusetts (University of Massachusetts Medical Center). Se usó 10 µg de ARN total de cada muestra como plantilla para sintetizar cADN marcado usando el Kit de Reactivo de Marcación de ADN GeneChip™ de Affymetrix. Las muestras de cADN marcadas se hibridaron con los Arreglos GeneChip™ de Affymetrix de acuerdo con la guía de Affymetrix. Los arreglos de hibridados fueron escaneados con un escáner GeneChip™ 3000. Los archivos de datos crudos de imágenes resultantes se procesaron, se hizo un reporte de calidad y se normalizó la intensidad de la sonda usando Microarray Suite version 5.0 (MAS 5.0). Los valores de expresión se calcularon usando un paquete de software implementando el método GCRMA.

La calidad de los datos del microarreglo se analizó usando procedimientos de modelamiento a nivel de sonda proporcionados por el paquete affyPLM (Bolstad, B. M., F. Collin, J. Brettschneider, K. Simpson, L. Cope, R. Irizarry, y T. P. Speed. 2005. Quality Assessment of Affymetrix GeneChip Data, p. 33-47. en R. Gentleman, V. Carey, W. Huber, y S. Dudoit (ed.), *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, Heidelberg.) en BioConductor (Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5:R80). No se observaron artefactos en la imagen debido al procesamiento o a la fabricación del arreglo. Los valores de señal de fondo del microarreglo de 34 (glucosa), 32 (celobiosa), 30 (xilosa) y 34 (celulosa) se encontraban dentro de los valores de señal de fondo de 20 a 100 típicos promedio para arreglos Affymetrix. Las verificaciones de control de calidad para los procedimientos adaptados para el uso en *C. phytofermentans*, a saber, la purificación de ARN, la síntesis de cADN, la marcación y la hibridación, indicaron una alta calidad de los datos.

Estimación de los límites del transcripto de ARNm. Para identificar secuencias teóricas del promotor se estimó la longitud de los transcriptos de mRNA

con un error de ± 24 bases. Se compararon los niveles de expresión de las regiones intergénicas adyacentes a la ORF y las regiones dentro de la ORF usando sondas específicas (Figura 4). Las lecturas por encima de 1000 UA indicaron que una sonda específica representaba parte de una región expresada. Por el contrario, las lecturas por debajo de 250 UA indicaron que no había hibridación específica entre la sonda y una región expresada. Las sondas que indican una región teórica expresada tienen una lectura mayor que:

Promedio (gen 1) – desv.est (gen 1) o promedio (gen 2) – desv. est (gen 2).

Para evitar los errores de una única sonda, se usaron lecturas de al menos dos sondas consecutivas para indicar una región expresada. Sin embargo, debido a que algunas sondas pueden tener propiedades que hacen que no presenten respuesta, se incluyeron sondas con lecturas por debajo del umbral en el mapeo de regiones expresadas donde sondas consecutivas corriente arriba y corriente abajo de la sonda que no presentaba respuesta cumplieron con los criterios. Esto permitió una mejor cuantificación de los límites del transcripto de genes de bajo nivel de expresión y regiones intergénicas expresadas adyacentes consolidadas.

Se empleó BLAST para identificar fuentes potenciales de hibridación cruzada, corriendo BLAST para cada sonda detectada contra el genoma de *C. phytofermentans*. Para cualquier coincidencia con valores E inferiores a 0.01, se midieron las intensidades de las sondas en el arreglo correspondiente a la coincidencia de BLAST. Si alguna de las coincidencias exhibía un valor de expresión mayor que la sonda en cuestión, se señalaba la sonda como una posible fuente de hibridación cruzada. Para cada región expresada teórica, se reportó el número de sondas positivas y el número de estas sondas positivas que se consideraba con posibilidades de hibridación cruzada. Se reportaron los límites del transcripto para cada proteína predicha relacionada con glucósido hidrolasa y alcohol deshidrogenasa teórica.

Los genes correspondientes a transcriptos que se observó eran expresados diferencialmente más de 4 veces durante el crecimiento sobre D-arabinosa, en

comparación con el crecimiento sobre la glucosa se presentan en la tabla 10.

Tabla 10. Expresión sobre D-arabinosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
5.1	Cphy1174	piruvato formiato-liasa
5.0	Cphy1175	Familia de proteínas activadoras de la enzima de radical glicilo
4.9	Cphy1176	Proteínas de microcompartimientos
6.1	Cphy1177	Familia de proteínas adducina/aldolasa clase II
5.9	Cphy1178	Aldehido Deshidrogenasa
5.3	Cphy1179	Proteína alcohol deshidrogenasa con dominio que se une a zinc
5.6	Cphy1180	Proteína de microcompartimientos
5.7	Cphy1181	Proteína de microcompartimientos
5.0	Cphy1182	Proteína de microcompartimientos
4.5	Cphy1183	Proteína de utilización de propanodiol
3.9	Cphy1184	Proteína de utilización de Etanolamina EutN/ proteína estructural de carboxisoma Ccml
3.4	Cphy1185	Subunidad de 51 kDa del dominio de NADH deshidrogenasa de la cadena respiratoria
3.9	Cphy3153	Transporte FucU o RbsD
4.4	Cphy3154	carbohidrato quinasa FGGY
4.4	Cphy3155	L-fucosa isomerasa

2.6	Cphy3367	Celulosa 1,4-beta-celobiosidasa
2.8	Cphy3368	Cellulosa 1,4-beta-celobiosidasa

Los genes correspondientes a los transcritos que se observó eran expresados diferencialmente más de 4 veces durante el crecimiento sobre L-arabinosa, en comparación con el crecimiento sobre glucosa se presentan en la Tabla 11.

Tabla 11. Expresión sobre L-arabinosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
2.1	Cphy0580	relacionados con el transportador ABC
2.1	Cphy0581	ATPasa que transporta monosacárido
2.7	Cphy0582	ATPasa que transporta monosacárido
2.3	Cphy0583	transportador ABC de azúcar teórico, proteína de unión a sustrato
4.3	Cphy1071	familia 26 glicósido hidrolasa
2.0	Cphy1169	Alfa-N-arabinofuranosidasa
3.9	Cphy1219	xilosa isomerasa
6.9	Cphy1799	Glucósido hidrolasa, familia 18: precursor XII/V familia de unión a carbohidrato
6.3	Cphy1800	Precursor de quitinasa
3.2	Cphy2105	Endo-1,4-beta-xilanasa
2.5	Cphy2128	Manano endo-1,4-beta-manosidasa, celulosa 1,4-beta-celobiosidasa

2.3	Cphy2569	Familia 1 proteína que se une a soluto extracelular
3.7	Cphy2570	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.6	Cphy2571	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.2	Cphy2919	proteína de función desconocida DUF1565
2.0	Cphy3202	Celulasa
4.9	Cphy3367	Celulosa 1,4-beta-celobiosidasa
5.7	Cphy3368	Celulosa 1,4-beta-celobiosidasa
2.7	Cphy3419	Xiluloquinasa
3.0	Cphy3854	glicosiltransferasa 36
3.5	Cphy3855	Fosfomanomutasa
2.5	Cphy3858	Familia 1 de proteína que se une a soluto extracelular
3.9	Cphy3859	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.9	Cphy3860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.1	Cphy3861	regulador transcripcional de dos componentes, familia AraC
2.3	Cphy3862	Endo-1,4-beta-xilanasa

Los Genes correspondientes a transcriptos que se observó eran expresados diferencialmente más de 4 veces durante el crecimiento sobre celobiosa en comparación con el crecimiento sobre glucosa se presentan en la Tabla 12.

Tabla 12. Expresión sobre celobiosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
3.5	Cphy0430	glicosiltransferasa 36
2.1	Cphy1586	relacionados con el transportador ABC
2.3	Cphy1587	ATPasa que transporta monosacárido
2.0	Cphy2264	Relacionado con glicosidasa PH1107
1.9	Cphy2265	Familia 1 de proteína que se une a soluto extracelular
2.0	Cphy2266	Proteína hipotética
2.2	Cphy2267	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.3	Cphy2268	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.3	Cphy2269	Proteína hipotética
2.3	Cphy2270	-
2.0	Cphy2271	-
2.1	Cphy2272	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.2	Cphy2273	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
4.3	Cphy2464	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.8	Cphy2465	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína

2.7	Cphy2466	Familia 1 de proteína que se une a soluto extracelular
3.4	Cphy2467	regulador transcripcional, familia LacI

Los genes correspondientes a transcritos que se observa son expresados diferencialmente más de 4 veces durante el crecimiento sobre celulosa, en comparación con el crecimiento sobre glucosa se presentan en la Tabla 13.

Tabla 13. Expresión sobre Celulosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
3.4	Cphy0430	glicosiltransferasa 36
3.1	Cphy1071	familia 26 glicósido hidrolasa
3.0	Cphy1163	Celulasa
4.1	Cphy1529	Familia 1 de proteína que se une a soluto extracelular
3.2	Cphy1530	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.6	Cphy1531	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.1	Cphy1586	relacionados con el transportador ABC
2.3	Cphy1587	ATPasa que transporta monosacárido
6.2	Cphy1799	familia 18 de glicósido hidrolasa
6.2	Cphy1800	familia 18 de glicósido hidrolasa
2.0	Cphy1929	glicosiltransferasa 36

3.2	Cphy2105	Endo-1,4-beta-xilanasas
2.6	Cphy2263	Proteína hipotética
2.0	Cphy2264	Relacionado con glicosidasa PH1107
2.9	Cphy2265	familia 1 de proteína de unión a soluto extracelular
2.0	Cphy2266	Proteína hipotética
2.2	Cphy2267	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.3	Cphy2268	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.3	Cphy2269	Proteína hipotética
3.0	Cphy2270	-
2.7	Cphy2271	-
2.1	Cphy2272	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.2	Cphy2273	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.3	Cphy2274	familia 1 de proteína de unión a soluto extracelular
4.0	Cphy2464	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.5	Cphy2465	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.5	Cphy2466	Familia 1 de proteína de unión a soluto extracelular
3.0	Cphy2467	regulador transcripcional, familia LacI
2.9	Cphy2569	Familia 1 de proteína de unión a soluto extracelular
3.9	Cphy2570	Componente de membrana interna de sistemas de

		transporte dependientes de unión a proteína
3.9	Cphy2571	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.0	Cphy3209	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.0	Cphy3210	Proteína de unión a sustrato múltiple teórico del sistema de transporte de azúcar
4.8	Cphy3367	Celulosa 1,4-beta-celobiosidasa
5.5	Cphy3368	Celulosa 1,4-beta-celobiosidasa
2.5	Cphy3854	glicosiltransferasa 36
3.1	Cphy3855	Fosfomanomutasa
2.5	Cphy3858	familia 1 de proteína de unión a soluto extracelular
3.8	Cphy3859	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.7	Cphy3860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.9	Cphy3861	regulador transcripcional de dos componentes, familia AraC
2.3	Cphy3862	Endo-1,4-beta-xilanasa

Los genes correspondientes a transcritos que se observa son expresados diferencialmente 4 o más veces durante el crecimiento sobre fucosa, en comparación con el crecimiento sobre glucosa se presentan en la Tabla 14.

Tabla 14. Expresión sobre Fucosa

Expresión Diferencial	JGI No.	Descripción COG
-----------------------	---------	-----------------

(log ₂)		
2.7	Cphy0580	relacionados con el transportador ABC
2.6	Cphy0581	ATPasa que transporta monosacárido
2.8	Cphy0582	ATPasa que transporta monosacárido
2.2	Cphy0583	transportador de azúcar teórico ABC, proteína de unión a sustrato
2.2	Cphy0584	L-arabinosa isomerasa
2.8	Cphy1071	familia 26 de glicósido hidrolasa
2.8	Cphy1163	Celulasa
5.6	Cphy1174	piruvato formiato liasa
6.2	Cphy1175	familia de proteínas activadoras de la enzima radical glicilo
5.8	Cphy1176	proteína de microcompartimientos
6.5	Cphy1177	proteína de la familia aducina/aldolasa clase II
6.4	Cphy1178	Aldehído deshidrogenasa
6.3	Cphy1179	proteína de dominio que se une a zinc Alcohol deshidrogenasa
6.4	Cphy1180	proteína de microcompartimientos
6.4	Cphy1181	proteína de microcompartimientos
6.0	Cphy1182	proteína de microcompartimientos
5.9	Cphy1183	proteína de utilización de propanodiol
5.4	Cphy1184	proteína de utilización de etanolamina EutN/proteína estructural carboxisoma Ccml
5.2	Cphy1185	subunidad de 51 kDa de dominio de NADH

		deshidrogenasa de cadena respiratoria
4.7	Cphy1186	proteína de microcompartimientos
4.9	Cphy1799	familia 18 de glicósido hidrolasa
5.2	Cphy1800	familia 18 de glicósido hidrolasa
6.1	Cphy2010	relacionado con el transportador ABC
6.6	Cphy2011	ATPasa que transporta monosacárido
5.9	Cphy2012	proteína de unión periplásmica/regulador transcripcional LacI
2.0	Cphy2105	Endo-1,4-beta-xilanasa
2.4	Cphy2569	familia 1 de proteína de unión a soluto extracelular
3.3	Cphy2570	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.0	Cphy2571	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.5	Cphy2919	proteína de función desconocida DUF1565
4.9	Cphy3153	Transporte de FucU o RbsD
5.3	Cphy3154	carbohidrato quinasa FGGY
5.3	Cphy3155	L-fucosa isomerasa
2.3	Cphy3308	Proteína hipotética
4.2	Cphy3367	Celulosa 1,4-beta-celobiosidasa
4.7	Cphy3368	Celulosa 1,4-beta-celobiosidasa
2.1	Cphy3854	glicosiltransferasa 36
2.3	Cphy3855	Fosfomanomutasa
2.3	Cphy3858	familia 1 de proteína de unión a soluto extracelular

3.3	Cphy3859	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.2	Cphy3860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.3	Cphy3861	regulador transcripcional de dos componentes, familia AraC

Los genes correspondientes a los transcritos que se observó eran expresados diferencialmente 4 o más veces durante el crecimiento sobre galactosa, en comparación con el crecimiento sobre glucosa se presentan en la Tabla 15.

Tabla 15. Expresión sobre Galactosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
2.1	Cphy3367	Celulosa 1,4-beta-celobiosidasa
2.0	Cphy3368	Celulosa 1,4-beta-celobiosidasa

Los genes correspondientes a los transcritos que se observa son expresados diferencialmente más de 4 veces durante el crecimiento sobre Laminarina, en comparación con el crecimiento sobre glucosa se presentan en la tabla 16.

Tabla 16. Expresión sobre Laminaria

Expresión Diferencial (log ₂)	JGI No.	Descripción COG

5.4	Cphy0857	proteína de tipo celobiosa fosforilasa
5.1	Cphy0858	familia 30 de glicósido hidrolasa
4.9	Cphy0859	Proteína hipotética
5.7	Cphy0860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
5.8	Cphy0861	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.9	Cphy0862	familia 1 de proteína de unión a soluto extracelular
3.8	Cphy0863	región interna histidina quinasa
3.8	Cphy0864	regulador transcripcional de dos componentes, familia AraC
4.0	Cphy0865	Proteína hipotética
2.2	Cphy1448	transportador de fosfonato ABC, proteína de unión a fosfonato periplásmico
1.8	Cphy1449	transportador de fosfonato ABC, subunidad ATPasa
1.9	Cphy1450	transportador de fosfonato ABC, subunidad de membrana interna
2.0	Cphy1451	transportador de fosfonato ABC, subunidad de membrana interna
2.0	Cphy1929	glicosiltransferasa 36
4.9	Cphy3388	glucano endo-1,3-beta-D-glucosidasa

Los genes correspondientes a transcritos que se observó eran expresados diferencialmente más de 4 veces durante el crecimiento sobre manosa, en comparación con el crecimiento sobre glucosa se presentan en la Tabla 17.

Tabla 17. Expresión sobre Manosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
2.6	Cphy1071	familia 26 de glicósido hidrolasa
2.5	Cphy1585	Componente teórico de unión a soluto del transportador ABC
2.5	Cphy1586	relacionados con el transportador ABC
2.9	Cphy1587	ATPasa que transporta monosacáridos
3.9	Cphy1799	glicósido hidrolasa de la familia 18
4.2	Cphy1800	glicósido hidrolasa de la familia 18
2.5	Cphy2105	Endo-1,4-beta-xilanasas
3.1	Cphy2569	familia 1 de proteína de unión a soluto extracelular
3.6	Cphy2570	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.4	Cphy2571	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.8	Cphy3367	Celulosa 1,4-beta-celobiosidasas
4.1	Cphy3368	Celulosa 1,4-beta-celobiosidasas
2.1	Cphy3855	Fosfomanomutasa
2.4	Cphy3858	familia 1 de proteína de unión a soluto extracelular
3.3	Cphy3859	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.1	Cphy3860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.4	Cphy3861	regulador transcripcional de dos componentes familia

		AraC
--	--	------

Los genes correspondientes a los transcritos que se observa son expresados diferencialmente más de 4 veces durante el crecimiento sobre pectina, en comparación con el crecimiento sobre glucosa se presentan en la tabla 18.

Tabla 18. Expresión sobre Pectina

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
2.7	Cphy0218	familia 31 de glicósido hidrolasa
2.2	Cphy0219	Proteína hipotética
2.5	Cphy0220	proteína de dominio de la Familia 3 de glicósido hidrolasa
3.5	Cphy0430	glicosiltransferasa 36
2.2	Cphy1071	familia 26 de glicósido hidrolasa
2.3	Cphy1174	formiato de piruvato-liasa
2.4	Cphy1175	familia de proteínas activadoras de la enzima de radical glicilo
2.1	Cphy1176	proteínas de microcompartimientos
3.3	Cphy1177	proteína de familia adducina/aldolasa clase II
2.9	Cphy1178	Aldehído deshidrogenasa
2.3	Cphy1179	Proteína de dominio de unión a zinc alcohol deshidrogenasa
2.9	Cphy1180	Proteína de microcompartimientos

2.8	Cphy1181	Proteína de microcompartimientos
2.2	Cphy1182	Proteína de microcompartimientos
2.1	Cphy1183	Proteína de utilización de propanodiol
2.0	Cphy1219	xilosa isomerasa
3.2	Cphy1612	Alergeno Amb/pectato liasa
2.0	Cphy1714	familia 85 de glicósido hidrolasa
3.6	Cphy1715	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.4	Cphy1716	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.3	Cphy1717	familia 1 de proteína de unión a soluto extracelular
2.7	Cphy1718	Relacionado con glicosidasa PH1107
2.5	Cphy1719	Proteína hipotética
2.4	Cphy1720	Familia 38 de glicósido hidrolasa
2.4	Cphy1888	Proteína hipotética
3.4	Cphy1929	glicosiltransferasa 36
2.1	Cphy2010	relacionado con el transportador ABC
2.1	Cphy2011	ATPasa que transporta monosacárido
2.1	Cphy2262	N-acilglucosamina 2-epimerasa
3.1	Cphy2263	Proteína hipotética
3.3	Cphy2264	Relacionado con glicosidasa PH1107
3.1	Cphy2265	familia 1 de la proteína de unión a soluto extracelular
3.1	Cphy2266	Proteína hipotética
4.0	Cphy2267	Componente de membrana interna de sistemas de

		transporte dependientes de unión a proteína
3.8	Cphy2268	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.9	Cphy2269	Proteína hipotética
3.6	Cphy2270	-
3.6	Cphy2271	-
4.0	Cphy2272	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
4.0	Cphy2273	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.2	Cphy2274	familia 1 de proteína de unión a soluto extracelular
2.4	Cphy2275	Proteína hipotética
2.1	Cphy2276	Manano endo-1,4-beta-manosidasa
3.9	Cphy2464	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.5	Cphy2465	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.4	Cphy2466	familia 1 de proteína de unión a soluto extracelular
2.9	Cphy2467	regulador transcripcional, familia LacI
2.9	Cphy2919	proteína de función desconocida DUF1565
2.4	Cphy3153	Transporte de FucU o RbsD
2.5	Cphy3154	carbohidrato quinasa FGGY
2.6	Cphy3155	L-fucosa isomerasa
2.2	Cphy3160	familia 2 de glicósido hidrolasa que se une a azúcar
3.4	Cphy3367	Celulosa 1,4-beta-celobiosidasa

3.5	Cphy3368	Celulosa 1,4-beta-celobiosidasa
4.2	Cphy3585	regulador transcripcional, familia LacI
6.5	Cphy3586	Arabinogalactan endo-1,4-beta-galactosidasa
6.8	Cphy3587	Proteína hipotética
6.5	Cphy3588	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
6.1	Cphy3589	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
6.7	Cphy3590	familia 1 de proteína de unión a soluto extracelular
2.2	Cphy3859	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.0	Cphy3860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína

Los genes correspondientes a los transcritos que se observó son expresados diferencialmente más de 4 veces durante el crecimiento sobre ramnosa, en comparación con el crecimiento sobre glucosa se presentan en la Tabla 19.

Tabla 19. Expresión sobre Ramnosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
3.7	Cphy1071	familia 26 de glicósido hidrolasa
3.4	Cphy1163	Celulasa
3.2	Cphy1219	xilosa isomerasa
2.8	Cphy1585	componente de unión a soluto teórico del

		transportador ABC
3.4	Cphy1586	relacionados con el transportador ABC
3.7	Cphy1587	ATPasa que transporta monosacárido
6.4	Cphy1799	familia 18 de glicósido hidrolasa
6.1	Cphy1800	familia 18 de glicósido hidrolasa
5.0	Cphy2105	Endo-1,4-beta-xilanasasa
2.0	Cphy2106	proteína de función desconocida DUF323
2.1	Cphy2128	Manan Endo-1,4-beta-manosidasa, celulosa 1,4-beta-celobiosidasa
2.9	Cphy2265	Familia 1 de proteína de unión a soluto extracelular
2.8	Cphy2569	Familia 1 de proteína de unión a soluto extracelular
4.0	Cphy2570	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.9	Cphy2571	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.1	Cphy2919	proteína de función desconocida DUF1565
5.1	Cphy3367	Celulosa 1,4-beta-celobiosidasa
5.9	Cphy3368	Celulosa 1,4-beta-celobiosidasa
2.5	Cphy3419	Xiluloquinasa
2.4	Cphy3854	glicosiltransferasa 36
2.7	Cphy3855	Fosfomanomutasa
2.3	Cphy3858	Familia 1 de proteína de unión a soluto extracelular
3.8	Cphy3859	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína

3.8	Cphy3860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.9	Cphy3861	regulador transcripcional de dos componentes, familia AraC

Los genes que corresponden a los transcriptos que se observa son expresados diferencialmente más de 4 veces durante el crecimiento sobre xilano, en comparación con el crecimiento sobre glucosa se presentan en la tabla 20.

Tabla 20. Expresión sobre Xilano

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
2.4	Cphy1132	componente de unión a soluto teórico del transportador ABC
2.7	Cphy1133	ATPasa que transporta monosacárido
2.6	Cphy1134	Relacionado con el transportador ABC
2.2	Cphy1177	proteína de familia aducina/aldolasa clase II
2.1	Cphy1178	Aldehído deshidrogenasa
2.0	Cphy1181	proteína de microcompartimientos
3.6	Cphy1219	xilosa isomerasa
2.7	Cphy1448	transportador ABC de fosfonato, proteína de unión a fosfonato periplásmico
2.3	Cphy1449	transportador ABC de fosfonato, subunidad ATPasa
2.5	Cphy1450	transportador ABC de fosfonato, subunidad membrana interna

2.6	Cphy1451	transportador ABC de fosfonato, subunidad membrana interna
2.8	Cphy1528	regulador transcripcional, familia AraC
6.7	Cphy1529	familia 1 de proteína de unión a soluto extracelular
6.8	Cphy1530	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
6.7	Cphy1531	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
5.3	Cphy1532	Dominio de función desconocida DUF1801
4.5	Cphy2105	Endo-1,4-beta-xilanasasa
4.6	Cphy2106	proteína de función desconocida DUF323
4.8	Cphy2108	Endo-1,4-beta-xilanasasa
2.4	Cphy2632	familia 43 de glicósido hidrolasa
2.4	Cphy2654	proteína de unión a sustrato del transportador ABC ed azúcar
4.2	Cphy2655	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
4.3	Cphy2656	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
4.3	Cphy3009	Proteína de dominio de la familia 3 de glicósido hidrolasa
4.1	Cphy3010	Endo-1,4-beta-xilanasasa
4.6	Cphy3158	Alfa-glucuronidasa
4.3	Cphy3206	Transductor sensorial de quimiotaxis que acepta metilo
4.5	Cphy3207	Familia 8 de glicósido hidrolasa

4.4	Cphy3208	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
4.3	Cphy3209	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
4.6	Cphy3210	proteínas de unión a sustrato del sistema de transporte de azúcar múltiple teórico
3.3	Cphy3211	regulador transcripcional de dos componentes, familia AraC
3.0	Cphy3212	región interna de histidina quinasa
3.4	Cphy3419	Xiluloquinasa

Los genes correspondientes a transcritos para los que se observa son expresados diferencialmente más de 4 veces durante el crecimiento sobre xilosa, en comparación con el crecimiento sobre glucosa se presentan en la tabla 21.

Tabla 21. Expresión sobre Xilosa

Expresión Diferencial (log ₂)	JGI No.	Descripción COG
3.4	Cphy1071	familia 26 de glicósido hidrolasa
3.2	Cphy1219	xilosa isomerasa
2.8	Cphy1585	Componente de unión a soluto teórico del transportador ABC
3.4	Cphy1586	relacionado con el transportador ABC
3.7	Cphy1587	ATPasa que transporta monosacáridos
4.6	Cphy1799	familia 18 de glicósido hidrolasa

4.8	Cphy1800	familia 18 de glicósido hidrolasa
5.0	Cphy2105	Endo-1,4-beta-xilanasas
2.0	Cphy2106	proteína de función desconocida DUF323
2.2	Cphy2128	Manano Endo-1,4-beta-manosidasa, celulosa 1,4-beta-celobiosidasas
2.8	Cphy2569	Familia 1 de proteína de unión a soluto extracelular
3.7	Cphy2570	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.5	Cphy2571	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.1	Cphy2919	proteína de función desconocida DUF1565
4.4	Cphy3367	Celulosa 1,4-beta-celobiosidasa
4.9	Cphy3368	Celulosa 1,4-beta-celobiosidasa
2.5	Cphy3419	Xiluloquinasa
2.1	Cphy3854	glicosiltransferasa 36
2.4	Cphy3855	Fosfomanomutasa
2.4	Cphy3858	Familia 1 de proteína de unión a soluto extracelular
3.4	Cphy3859	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.4	Cphy3860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
2.6	Cphy3861	Regulador transcripcional de dos componentes, familia AraC

Los genes correspondientes a transcritos para los que se observa son expresados diferencialmente más de 4 veces durante el crecimiento sobre extracto de levadura, en comparación con el crecimiento sobre glucosa se presentan en la Tabla 22.

Tabla 22. Expresión sobre Extracto de Levadura

Expresión Diferencial (log2)	JGI No.	Descripción COG
2.4	Cphy0857	Proteína de tipo celobiosa fosforilasa
2.2	Cphy0858	familia 30glicósido hidrolasa
3.2	Cphy0860	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.6	Cphy0861	Componente de membrana interna de sistemas de transporte dependientes de unión a proteína
3.0	Cphy1448	transportador ABC de fosfonato
2.0	Cphy1449	transportador ABC de fosfonato
2.1	Cphy1450	transportador ABC de fosfonato
2.0	Cphy1451	transportador ABC de fosfonato

Ejemplo 3. Disección Genómica del Genoma de *C. phytofermentans*

Organización del genoma. ISDg ATCC 700394 de *C.phytofermentans* tiene un solo cromosoma circular de 4.847.594 pares de bases y no alberga ningún plásmido. El origen de replicación del cromosoma se definió utilizando la posición del punto de transición de la posición oblicua GC y la presencia de la replicación característica de la proteína dnaA (Figura 5). El contenido de G + C es 35,3%. El

trazado del contenido de G + C, de ventanas 1kb en función de la posición en el genoma (Figura 6) revela varias islas genómicas aisladas, con un mayor contenido G + C. La ubicación de 6 islas específicas se definió como regiones 1 kb con un contenido promedio de G + C> 50%), que se muestra en la Figura 6. Los genes fueron identificados ya sea dentro o alrededor de cada una de estas islas genómicas (Tabla 23).

Tabla 23: Características generales del genoma de *C. phytofermentans*

Parámetro	Valor
Tamaño (bp)	4,847,594
contenido G+C (%)	35.3
Genes que codifican proteínas	
No. (%) similar a las proteínas conocidas	2,870 (73.1)
No. (%) similar a las proteínas de función desconocida ^a	170 (4.3)
No. (%) de hipotéticos conservados ^b	265 (6.7)
No. (%) de hipotéticos ^c	621 (15.8)
Total	3,926
tamaño ORF promedio (bp)	1,009
Codificación (%)	81
No. de agrupamientos de rARN	8
No. de genes tARN	61
^a Función desconocida, importante similitud de secuencia a una proteína nombrada a la que ninguna función específica es actualmente atribuida.	
^b Conservadas hipotética Proteína con similitud de secuencia una traducción de un marco de lectura abierta (ORF) en otro organismo, sin embargo no hay evidencia	

experimental de que expresión de la proteína existe.
° Proteínas hipotéticas con ninguna similitud significativa a cualquier otro gen secuenciado.

En general, estas islas altas de G + C parecen tener baja densidad de genes. De las dieciséis regiones de 1 Kb, con un contenido de G + C> 50% que componen las 6 islas genómicas, 12 de las regiones no contienen genes. Los genes que únicos se encuentran dentro de las islas de alto G + C es un sistema de dos componentes (histidina quinasa y regulador de la respuesta) y una proteína con una repetición teórica de triple hélice de colágeno. La mayoría de los genes que rodean a estas regiones de alto G + C son de función desconocida. Uno de los genes adyacentes a la RegiónV, codifica un proteína del fago (Figura 6). El genoma codifica 3,926 secuencias codificates predichas (CDS) (Tabla 23).

Los genomas de Clostridim exhiben de manera típica un fuerte sesgo de codificación, sin embargo, en *C.phytofermentans* los CDS están codificados por igual en la hebra principal (52%) y rezagados (48%) (Seedorf, H. *et al.* The genome of Clostridium kluyveri, a strict anaerobe with unique metabolic features. *Proc. Natl. Acad. Sci. EE.UU.* 105, 2128-2133 (2008)). A setenta y tres por ciento de los CDS se les asignó funciones teóricas, mientras que el 11% tenía similitud con genes de función desconocida, y el 16% eran exclusivas de *C. phytofermentans*.

Se predicen sesenta y un genes de tARN en el genoma que cubren 20 aminoácidos (Tabla 23, Tabla 24).

Tabla 24: Genes de tARN de *C.phytofermentans*

tARN	Locus				
Cys	CphyR0019				1
His	CphyR0041				1

Phe	CphyR0010				1
Trp	CphyR0094				1
Asn	CphyR0018	CphyR0030			2
Asp	CphyR0005	CphyR0034			2
Gln	CphyR0042	CphyR0098			2
Ile	CphyR0015	CphyR0069			2
Tyr	CphyR0008	CphyR0021			2
Xaa	CphyR0055	CphyR0074			2
Ala	CphyR0016	CphyR0028	CphyR0068		3
Pro	CphyR0038	CphyR0045	CphyR0083		3
Val	CphyR0006	CphyR0035	CphyR0049		3
Arg	CphyR0037	CphyR0040	CphyR0073	CphyR0109	4
Glu	CphyR0031	CphyR0059	CphyR0060	CphyR0063	4
Lys	CphyR0011	CphyR0043	CphyR0107	CphyR0108	4
Met	CphyR0009	CphyR0020	CphyR0033	CphyR0050	4
Ser	CphyR0001	CphyR0024	CphyR0025	CphyR0056	4
Thr	CphyR0007	CphyR0032	CphyR0080	CphyR0082	4
Gly	CphyR0023	CphyR0039	CphyR0075	CphyR0078	6
tARN	Locus				
Leu	CphyR0022	CphyR0036	CphyR0044	CphyR0062	6
Total					61
tARN	Locus				
Cys	CphyR0019				1

His	CphyR0041				1
Phe	CphyR0010				1
Trp	CphyR0094				1
Asn	CphyR0018	CphyR0030			2
Asp	CphyR0005	CphyR0034			2
Gln	CphyR0042	CphyR0098			2
Ile	CphyR0015	CphyR0069			2
Tyr	CphyR0008	CphyR0021			2
Xaa	CphyR0055	CphyR0074			2
Ala	CphyR0016	CphyR0028	CphyR0068		3
Pro	CphyR0038	CphyR0045	CphyR0083		3
Val	CphyR0006	CphyR0035	CphyR0049		3
Arg	CphyR0037	CphyR0040	CphyR0073	CphyR0109	4
Glu	CphyR0031	CphyR0059	CphyR0060	CphyR0063	4
Lys	CphyR000H	CphyR0043	CphyR0107	CphyR0108	4
Met	CphyR0009	CphyR0020	CphyR0033	CphyR0050	4
Ser	CphyR0001	CphyR0024	CphyR0025	CphyR0056	4
Thr	CphyR0007	CphyR0032	CphyR0080	CphyR0082	4
Gly	CphyR0023	CphyR0039	CphyR0075	CphyR0078	6
tARN	Locus				
Leu	CphyR0022	CphyR0036	CphyR0044	CphyR0062	6
Total					61

Los ocho operones ribosomales, de los cuales tres están orientados en la hebra principal y cinco en la hebra retrasada, se agrupan en proximidad general al origen de replicación (Tabla 23). La abundancia de operones de rARN en *C. phytofermentans* puede ser una adaptación evolutiva y una ventaja para los organismos que experimentan fluctuación en las condiciones de crecimiento como lo sugiere el aumento de la capacidad de respuesta rápida a las condiciones favorables de crecimiento para las bacterias con mayor número de operones (Schmidt, T.M. en *Bacterial genomes: physical structure and analysis*. 221 (Chapman and Hall Co., New York, N.Y., 1997); Klappenbach, J. A., Dunbar, J. M. & Schmidt, T. M. rRNA operon copy number reflects ecological strategies of bacteria *Appl. Environ. Microbiol.* 66, 1328-1333 (2000); Condon, C., Liveris, D., Squires, C., Schwartz, I. & Squires, C. L. rRNA operon multiplicity in *Escherichia coli* and the physiological implications of *rrn* inactivation. *J. Bacteriol.* 177, 4152-4156 (1995)).

Aparentemente hay un profago teórico en el genoma revelado por la presencia agrupada de genes relacionados con fago. El fago-agrupamiento abarca aproximadamente 39 kb, e incluye 40 genes (Cphy2953-2993). Quince genes, responsables de los componentes estructurales de la cabeza y la cola y su montaje, son homólogos a genes en fagos de *Clostridium difficile* ΦC2 (Goh, S., Ong, P.F., Song, K.P., Riley, T.V. & Chang, B.J. The complete genome sequence of *Clostridium difficile* ohage phiC2 and comparisons to phiCD119 and inducible prophages of CD630. *Microbiology* 153, 676-685 (2007)). No está claro si los genes equivalentes funcionales necesarios para el fago para completar su ciclo de vida, i.e., empaquetamiento del ADN, montaje de la cola, lisis celular, control de lisogenia y replicación, recombinación y modificación del ADN están presentes en el genoma (*Id.*). Veintisiete genes relacionados con la inserción de secuencias (transposasas) están presentes, lo que es menor que en los genomas estrechamente relacionados (Tabla 25).

Tabla 25: Comparación de genomas de *Clostridium*

	C. phytofermentans	C. bolteae	R. gnavus	R. obeum	C. beijerinckii	C. botulinum	C. perfringens	C. cellulolyticum	C. thermocellum	Caldicellulosiruptor saccharolyticus	Thermoanaerobacter ethanolicus
Agrupamiento de clostridium	XIVa	XIVa	XIVa	XIVa	I	I	I	III	III	X	V
Características generales del cromosoma											
Tamaño del cromosoma, pb	4,847,594	6,556,988	####	3,624,708	6,000,632	3,995,387	2,897,393	3,958,683	3,843,301	2,970,275	2,362,816
contenido de GC, %	35	49	42	41	29	28	28	37	38	35	34
Codificación, %	81	91	92	90	79	81	80	86	83	86	86
Genes de codificación de proteína	3,926	7,284	3,913	4,175	5,020	3,572	3,635	3,283	3,189	2,679	2,243
Transposasas (COG0675, pfam01548, pfam02371)	27	77	61	50	42	1	93	100	139	106	56
Glucósido hidrolasas ³	109	na	na	na	75	23	38	na	70	61	15
Familias de glucósido hidrolasa ^a	39	na	na	na	25	10	21	na	23	31	26
Proteínas de unión a soluto (pfam01547)	21	30	9	8	11	4	7	11	5	18	8
Transportadores de polisacáridos ABC (Lplb COG4209)	20	0	2	0	1	0	1	6	0	7	0
Transportadores de Xilosa ABC (xylF PRK10355)	9	32	8	13	35	3	3	6	2	3	3

PurR (COG1609)	23	42	18	15	20	6	15	11	8	11	12
AraC (pfam00165)	70	66	33	28	48	21	10	35	9	19	1
	C. phytofermentans	C. bolteae	R. gnnavus	R. obeum	C. beijerinckii	C. botulinum	C. perfringens	C. cellulolyticum	C. thermocellum	Caldicellulosiruptor saccharolyticus	Therm oanaerobacter ethanolicus
AraC+CheY (COG4753)	18	34	11	11	8	1	2	12	2	13	0
^a www.CAZy.org											

C. phytofermentans está relacionado evolutivamente con los microbios del suelo asociados a los desechos de plantas. Para elucidar la relación filogenética entre *C.phytofermentans* y otros miembros de la clase Clostridia incluyendo genomas no secuenciados, se usaron secuencias de genes 16S rARN (1.611 pb) del aislado y de los miembros más estrechamente relacionados para el análisis de integración de vecinos. La filogenia confirmó que la cepa ISDg es un miembro del agrupamiento XIVa integrado por una mayoría de los microbios del intestino de humano/rata/pollo, y sólo lejanamente relacionada con el agrupamiento I, que contiene muchos agentes patógenos y el *Clostridium acetobutylicum* solventegénico, y el agrupaminto III, que contiene bacterias celulolíticas tales como *Clostridium cellulolyticum* y *Clostridium thermocellum* (Figura 7) (Warnick, T. A., Methe, B. A. & Leschine, S. B. *Clostridium phytofermentans* sp. nov., a cellulolytic mesophile from forest soil. *Int. J. Syst. Evol. Microbiol.* 52, 1155-1160 (2002); Collins, M. D. *et al.* The phylogeny of the genus Clostridium: proposal of five new genera and eleven new species combinations. *Int. J. Syst. Bacteriol.* 44, 812-826 (1994)). Dentro del agrupamiento XIVa, *C. phytofermentans* es parte de un subtipo que contiene las bacterias no cultivadas derivadas de los análisis de metagenómica del suelo anóxico de arrozal, biorreactor metanogénico de lixiviados de vertederos (93,7-93,8% de similitud) (Burrell, P. C., O'Sullivan, C., Song, H., Clarke, W. P. & Blackall, L. L. Identification, detection, and spatial resolution of Clostridium populations responsible for cellulose degradation in a methanogenic landfill leachate bioreactor. *Appl. Environ. Microbiol.* 70, 2414-2419 (2004); Hengstmann, U., Chin, K. J., Janssen, P. H. & Liesack, W. Comparative phylogenetic assignment of environmental sequences of genes encoding 16S rRNA and numerically abundant culturable bacteria from an anoxic rice paddy soil. *Appl. Environ. Microbiol.* 65, 5050-5058 (1999)), las especies de *Clostridium aminovalericum* (92,4% de similitud) y *Clostridium jejuense* (92,2% de

de *Clostridium aminovalericum* (92,4% de similitud) y *Clostridium jejuense* (92,2% de similitud), divergentes del subtipo de microbios intestinales (89,7% de similitud) (Figura 7).

El grupo de *C.phytofermentans* dentro de la clase Clostridia basado en el análisis de rARN es consistente con la distribución global de genes CDS de *C.phytofermentans* de acuerdo a su similitud con otros genes en genomas completamente secuenciados utilizando BLASTP. Treinta y ocho por ciento de los CDS eran más similares al agrupamiento XIVa, seguido de 10% en el agrupamiento I y 7% en el agrupamiento III (Figura 8). Sin embargo una proporción significativa de los CDS (14%), no tenía homología evidente en la clase Clostridia y exhibió el más alto nivel de similitud con CDS en cepas filogenéticamente distantes. Esto sugiere que el genoma de *C. phytofermentans* puede contener muchos genes adquiridos por una transferencia horizontal de genes. Estos orígenes dispersos en los genes subrayan la heterogeneidad del género *Clostridium* y la singularidad de *C. phytofermentans* entre los genomas secuenciados.

Montaje de un conjunto único de GH de diversos orígenes. La presencia simultánea de glicosido hidrolasas (GHs) con una amplia gama de funciones en un solo genoma tal como *C. phytofermentans* es notable. A pesar de la diversidad organización de los sistemas microbianos para la utilización de polisacáridos, los bloques de construcción básicos muestran una uniformidad considerable. Los dominios catalíticos que se encuentran en las enzimas que degradan polisacáridos se pueden organizar en las familias de acuerdo con sus secuencias principales y sus topologías de plegado. Pueden encontrarse representantes de estas familias entre muchas diversas bacterias y microorganismos eucariotas. Al cuantificar las similitudes o diferencias en los dominios catalíticos de GH de otras bacterias así como su organización al nivel de gen a genoma, tratamos de comprender mejor su función o medir la singularidad de *C. phytofermentans*.

En comparación con enzimas en otras bacterias secuenciadas utilizando BLASTP, las GH de *C. phytofermentans* son similares a una amplia diversidad de bacterias que representan seis filos y 46 especies. Hay más genes de GH similares a bacterias lejanamente relacionados de lo esperado de la distribución de todos los

genes en *C. phytofermentans*, (prueba de chi cuadrado, $P = 0.0004998$) (Figura 9). Alrededor del 50% de las GH en comparación con el 14%, esperado son más similares a especies fuera de la clase Clostridia (Figura 9). Alrededor del 18% de las GH es más similar a los bacilos, seguido de un 17% más similar al agrupamiento III de bacterias celulolíticas (Figura 9). Esto sugiere que la transferencia horizontal de genes juega un papel clave en la evolución de las capacidades para degradar plantas en *C. phytofermentans* y el montaje de un único conjunto de GH de orígenes muy diferentes.

Con la excepción de los genes de degradación de almidón agrupados en el genoma (Cphy2304-2352), en general, hay muy poca co-localización de los genes hemicelulasa o celulasa. Esto apoya la hipótesis de que estos últimos fueron adquiridos a través de la transferencia horizontal de genes independientes. Sin embargo, un ejemplo tándem de genes con función relacionada es un agrupamiento degradador de xilano con una beta-glucosidasa Cphy3009 (GH3), una endo-xilanasa Cphy3010 (GH10) y un arabinofuranosidasa Cphy3011 (GH43). Este tándem que parece ser único para *C. phytofermentans*. Además, los genes que codifican las dos celulasas principales, Cphy3367 (GH9) y Cphy3368 (GH48) son contiguos en el genoma. Esto es consistente con el alto sinergismo observado entre las celulasas bacterianas GH48 y GH9, presentes en todos los sistemas de enzimas celulasas conocido hasta ahora en las bacterias (Riedel, K., Ritter, J. & Bronnenmeier, K. Synergistic interaction of the *Clostridium stercorarium* cellulases Avicelase I (CelZ) and Avicelase II (CelY) in the degradation of microcrystalline cellulose. *FEMS Microbiol. Lett.* 147, 239 (1997)).

En la filogenia molecular, los dominios catalíticos GH9 y GH48 de *C. phytofermentans* (ver Figuras 10 y 11) son los más similares al precursor de endoglucanasa Z (Avicelasa I) (Jauris, S. *et al.* Sequence analysis of the *Clostridium stercorarium* celZ gene encoding a thermoactive cellulase (Avicelase I): identification of catalytic and cellulose-binding domains. *Mol. Gen. Genet.* 223, 258-267 (1990)) y celodextrinohidrolasa (Avicelasa II), respectivamente (Bronnenmeier, K., Rucknagel, K. P. & Staudenbauer, W. L. Purification and properties of a novel type of exo-1,4-beta-glucanase (avicelase II) from the cellulolytic thermophile *Clostridium stercorarium*. *Eur. J. Biochem.* 200, 379-385 (1991)) de los termófilos celulolíticos y xilanolíticos *Clostridium Stercorarium*. En *C. stercorarium*, GH9 y GH48 también son

adyacentes (Schwarz, W. H., Zverlov, V. V. & Bahl, H. Extracellular glycosyl hydrolases from Clostridia. *Adv. Appl. Microbiol.* 56, 215-261 (2004)). Esto es aún más extremo en el termófilo *C. saccharolyticus*, donde GH9 y GH48 son altamente similares a aquellos de *C. phytofermentans* y *C. stercorarium*, y se funden en una sola proteína. Estas observaciones sugieren un origen común y el funcionamiento sinérgico de estas enzimas clave en estas tres bacterias.

Dado que una determinada familia de GH contiene enzimas con un rango muy amplio de actividades conocidas, uno podría preguntarse si la redundancia en un familia refleja la redundancia en la función o una falta de especificidad del dominio catalítico. Estudios moleculares más detallados de las familias de GH de interés para la degradación de la pared celular de los vegetales muestran que a menudo hay variación en las secuencias de estos genes relacionados pero no idénticos. Cabe mencionar que en comparación con el especialista celulósico *C. thermocellum* con un número sobresaliente de celulasas GH9 (16), GH48 (2), GH5 (11) lo que refleja su especialización en degradar celulosa, *C. phytofermentans* tiene de manera general menos redundancia por familia, pero un rango más amplio de familias de GH lo que reflejan su ecología más generalista (Figura 9). No obstante, algunas familias de GH en *C. phytofermentans* todavía contienen un número significativo de genes. Este es el caso de las glucosidasas GH3, celulasas GH5, y enzimas que degradan xilano GH10, GH26, GH43. La filogenia molecular de las celulasas GH5 (pfam00150) de *C. phytofermentans* reveló que son diversas, separadas en 2 subagrupamientos (Figura 11). El agrupamiento B contiene celulasas fúngicas. Este ejemplo refuerza cómo la transferencia lateral de genes ha afectado la evolución de la GH. Más particularmente, se hace énfasis en la importancia de la transferencia de genes entre microorganismos que pertenecen a diferentes reinos lo que conjetura un papel aún más importante de la transferencia de genes entre reinos. La filogenia de los 6 dominios de GH10 de *C. phytofermentans* sugiere un amplio rango de variabilidad y probablemente función en *C. phytofermentans*. Cphy2108 (GH10) es muy similar a la de xilanasa multimodular de *C. stercorarium* Xyn10C, una proteína termoestable unida a célula y de unión a celulosa y xilano, uniendo por lo tanto la célula al sustrato (Ali, M. K., Kimura, T., Sakka, K. & Ohmiya, K. The multidomain xylanase Xyn10B as a cellulose-binding protein in *Clostridium stercorarium*. *FEMS Microbiol. Lett.* 198, 79-83 (2001)). Dos dominios de GH10 únicos pero estrechamente relacionados se encuentran en una sola enzima fusionados a un dominio CE lo que sugiere un evento de duplicación seguido de fusión y divergencia. Esta disposición específica de

la función catalítica en una sola proteína es única para *C. phytofermentans*.

Las duplicaciones, seguidas de fusiones y rearrreglos, y la divergencia de secuencia generaron un enorme arreglo de enzimas multimodulares en *C. phytofermentans* que varían en sus especificidades de sustrato y características cinéticas. Pero en general, la característica más llamativa de *C. phytofermentans* es la importancia de la transferencia horizontal de genes que permitió la adquisición de un complejo arreglo de genes, y agrupamientos de genes, de otros miembros de la comunidad del lugar.

Degradación de la pared celular vegetal sin celulosoma.

C. phytofermentans comparte una ecología similar con las bacterias que producen celulosoma. Sin embargo, no hay evidencia bioquímica ni genética (no hay dominios de docquerina, cohesina, o ancorina) para la producción de celulosomas en estas bacterias. Se cree que los complejos de celulosoma están involucrados en la degradación de la pared celular de los vegetales, ya que proporcionan un mecanismo de superficie celular bacteriano para la retención de una alta concentración de proteínas que representan el arreglo de especificidades de sustrato que son necesarias para romper los diversos enlaces existentes en los polisacáridos de la pared celular vegetal; potencialmente maximizan la estequiometría y la sinergia entre las diferentes especificidades catalíticas de enzimas y de unión; y pueden ayudar a limitar la difusión de los productos de degradación fuera de la célula, proporcionando un ambiente especial entre la membrana celular y el sustrato (Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R. & White, B. A. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat. Rev. Microbiol.* 6, 121- 131 (2008)). La estrategia que emplea *C. phytofermentans* para una degradación eficiente de la pared celular vegetal y la captación de producto sin un celulosoma no es claro.

En primer lugar, es sorprendente que para todas las celulasas GH5, GH9 y GH48, no hay ninguna agrupación en los árboles filogenéticos de acuerdo con la afiliación de la proteína a un celulosoma. Esto sugiere que, o bien el dominio de docquerina que ancla la proteína en el complejo se perdió o que fue adquirido muchas veces de manera independiente (Figuras 10 y 11). La característica multimodular

más destacada de las proteínas celulosómicas donde múltiples dominios de diversas familias de GH, CE, PL y el módulo de unión a carbohidratos (CBM, por las siglas en inglés de carbohydrate-binding module) se fusionan en una sola proteína se conserva en *C. phytofermentans*. *C. phytofermentans* cuenta con 19 proteínas GH modulares, lo que representa aproximadamente el 17% de todas las GH. En las bacterias no celulolíticas, la GH correspondiente se encuentra principalmente en los polipéptidos de dominio único (Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R. & White, B. A. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat. Rev. Microbiol.* 6, 121-131 (2008)). Esto podría reflejar la diferente ubicación celular de los productos génicos, y el hecho de que actúan sobre sustratos carbohidratos más pequeños, solubles (*Id.*). Según esta hipótesis, la organización de multimodular que parece ser característica de las enzimas de las especies celulolíticas puede ser conservada principalmente para el procesamiento extracelular de sustratos insolubles heterogéneos, tales como paredes celulares vegetales(*Id.*).

En ausencia de un celulosoma, el CBM podría fijar las enzimas firmemente a la pared celular vegetal y así mantenerlas en las inmediaciones de su sustrato. Se identificaron treinta y cinco CBM teóricos que representan 15 familias CAZy (Tabla 5). Se ha demostrado que CBM2, CBM3, CBM4, CBM6 y CBM46 se unen a celulosa (Tabla 5). Se ha demostrado que CBM2, CBM4, CBM6, CBM13, CBM22, CBM35 y CBM36 se unen a xilano (Tabla 5). La presencia de diversas combinaciones de dominios de CBM con una especificidad que no coincide con la especificidad del dominio catalítico puede suponer una ventaja para una acción en diferentes topologías de la pared celular vegetal donde múltiples tipos de polisacáridos están entrecruzados. Las xilanasas con CBM que se une a celulosa podrían ayudar a *C. phytofermentans* a adherirse a las fibras de celulosa mientras degrada el xilano entrecruzado. La presencia de los CBM independientes de los dominios catalíticos podría también explicarse por su acción termoestabilizante que se ha demostrado en algunos casos. Otro tipo de dominio, X2, puede encontrarse entre los dominios catalíticos y de CBM o entre los dominios de CBM en una mananasa y tres celulasas en *C. phytofermentans* (Tabla 6). Hay muy poca información disponible sobre la función de X2 de las enzimas extracelulares de las bacterias. Puede postularse que sirven como separadores o engarces que permiten la interacción óptima entre los módulos catalítico de unión a sustrato, para la interacción proteína-proteína o como un dominio potencial de unión a carbohidratos.

Además de unirse al polisacárido, algunas enzimas parecen estar unidas a la célula permitiendo que la célula permanezca cerca de su sustrato. Entre las 31 enzimas GH que se ha predicho se segregan (se ha predicho que tienen un péptido señal y/o extracelular), la predicción recurrente de las hélices transmembranales, caja TonB (COG0810 y PS00430), el dominio SORT y / o de unión a la pared celular sugiere que estas enzimas pueden asociarse con la membrana o la pared celular (Tabla 6). Entre estas proteínas, se ha predicho que las enzimas 8 GH tengan tanto CBM como habilidad de adherirse a la célula para conseguir potencialmente la proximidad física de la célula a su sustrato degradado. Por último, el peculiar gen Cphy1775 (SLH-GH*-CBM32-CBM32) coincidió con un dominio SLH predicho (pfam00395) para el anclaje a la pared celular y también dos veces de tipo inmunoglobulina (CBM32) y puede comportarse como un dominio CBM, que unen la célula a su sustrato. Otras enzimas GH podrían estar ancladas a la superficie celular por otros mecanismos desconocidos. Las células podrían adherirse a través de diferentes dominios tales como pfam07705 (dominio de adhesión celular en las bacterias, CARDB, por las siglas en inglés de cell adhesion domain in bacteria) y pfam01391 (colágeno, repetición de triple hélice del colágeno). La formación de biopelículas puede también desempeñar un papel en la orquestación de la degradación de los polisacáridos de la pared celular vegetal.

Degradación acoplada con captación y fosforilación de un amplio rango de transportadores ABC inducibles. Aunque se encontró un sistema fosfotransferasa dependiente de fosfoenolpiruvato (PEP) en el genoma, los datos del perfil preliminar de expresión sugieren que éste puede no expresarse cuando las células se cultivan en cualquiera de los componentes principales de las paredes celulares de plantas tal como xilano, celulosa, celobiosa, o glucosa. Más bien, el notable número (137) de proteínas con dominio ABC_tran (pfam00005) en el genoma es consistente con la captación de carbohidratos a través de los transportadores del casete de unión a ATP (ABC, por las siglas en inglés de ATP-binding cassette). Por ello, en comparación con sus familiares dentro de la clase Clostridia (Tabla 25), *C. phytofermentans* tiene un número inusualmente alto (21) de dominios de unión a soluto (SBP_bac_1, pfam01547), típicamente asociados con la captación de transportadores ABC y que permite la unión específica de diferentes solutos. Esto sugiere una necesidad de afinidad a los distintos tipos de solutos, lo cual es consistente con la hipótesis de que *C. phytofermentans* puede captar diversos

oligosacáridos. Por último, el dominio Lplb (COG4209) de los transportadores ABC de polisacáridos, un subcomponente tipo permeasa de algunos transportadores ABC está sobre-representado (20) en *C. phytofermentans* en comparación con otras bacterias en la clase Clostridia (Tabla 25).

El número y la variedad de este dominio puede permitir la captación de un amplio rango de oligosacáridos (Tabla 25). Que *C. phytofermentans* tiene 0.5% de sus genes dedicados a esta función, mientras que *C. saccharolyticus*, que también es un generalista, sólo tiene un 0,2% sugiere que *C. saccharolyticus* capta tipos más uniforme de sacáridos. Otra característica notable es la presencia de GH (53 de 109) adyacente a 41 transportadores de tipo ABC, junto con los reguladores lo que sugiere captación y degradación acoplada, así como regulación específica (Tabla 7). El gran número y la variedad de GH94 (celobiosa fosforilasa/celodextrin fosforilasa) y GH65 (maltosa fosforilasa) (Tabla 25) es consistente con la hipótesis de que un amplio rango de tipos de oligosacáridos entran en la célula. La presencia de 4 de cada 5 proteínas celobiosa / celodextrina fosforilasa GH94 unidas a membrana cerca a un transportador ABC son consistentes con el transporte de celobiosa y celodextrina a través de una proteína ABC, que es también el caso de *C. cellulolyticum* (Desvaux, M., Guedon, E. & Petitdemange, H. Cellulose catabolism by *Clostridium cellulolyticum* growing in batch culture on defined medium. *Appl. Environ. Microbiol.* 66, 2461-2470 (2000)).

También hay un gran número de beta-glucosidasas (8 GH3) que pueden tener actividad contra celobiosa o xilobiosa. *C. phytofermentans* podría alimentar a los oligosacáridos en su catabolismo mediante la fosforilación energéticamente favorable a través de la celobiosa/celodextrina fosforilasa o por la acción hidrolítica beta-glucosidasa que desperdicia energía. Es probable que la concentración de celodextrinas y la disponibilidad de otros sustratos de crecimiento (e.g., celulosa o celobiosa) estén involucrados en la determinación del destino de celodextrins, así como la importancia relativa del rompimiento hidrolítico y fosforolítico. Dada la presencia generalizada de las rutas fosforolítica e hidrolítica para el metabolismo de celodextrina en microorganismos celulolíticos, es posible que esta aparente redundancia sea de valor selectivo. La regulación del flujo relativo mediante estas dos vías pueden permitir que el microbio ajuste la tasa de ATP en respuesta a factores ambientales (por ejemplo, la disponibilidad de sustrato). *C. phytofermentans*

también es capaz de captar monosacáridos tales como xilosa, atestiguado por la presencia de 9 XylF, que se ha predicho capta xilosa (Tabla 25).

Regulación afinada del metabolismo de carbohidratos. En comparación con sus familiares en Clostridia, *C. phytofermentans* tiene una abundancia de los reguladores transcripcionales AraC (70) y PurR (23) (Tabla 25). Los reguladores transcripcionales procariotas se clasifican en familias con base en la similitud de secuencias y criterios estructurales y funcionales. Los reguladores AraC normalmente activan la transcripción de genes involucrados en el metabolismo del carbono, la respuesta al estrés y la patogénesis (Ramos, J. L. *et al.*, "The TetR family of transcriptional repressors," *Microbiol. Mol. Biol. Rev.* 69, 326-356 (2005)). PurR pertenece a la familia del represor lactosa (LAC) y el producto del gen normalmente actúa como un represor, donde las concentraciones fisiológicas de ligando causan la disociación del complejo PurR-ADN (Id.). La abundancia de estos reguladores es consistente con una amplia variedad de utilización de sustratos y una compleja red de regulación.

De los 23 genes que tienen similitud significativa a PurR, 8 son adyacentes a los transportadores ABC, agrupados con enzimas GH. Entre los 70 genes que tienen una importante similitud con AraC, 20 se encuentran cerca de los transportadores ABC agrupados a enzimas GH. Entre los 41 transportadores ABC que se encuentran agrupados con enzimas de GH, 20 son adyacentes a una AraC y 8 están cerca de PurR. Con base en estas observaciones, planteamos la hipótesis de que éstos regulan la expresión de las respectivas regiones genómicas (Tabla 7).

Ejemplo 4. Evaluando la Actividad Hydrolasa

Puede emplearse una variedad de métodos para evaluar la actividad biológica de un hidrolasa predicha. En un ejemplo, un gen predicho identificado en *C. phytofermentans* que codifica una hidrolasa es aislado y clonado en un vector de expresión. El vector de expresión se transforma en un microorganismo, por ejemplo, *E. coli*. La actividad del gen expresado se mide mediante el suministro de los microorganismos transformados con el sustrato de la hidrolasa predicho y midiendo el agotamiento del sustrato y el aumento de los productos de hidrólisis, y comparando

el nivel de esta actividad con la actividad en un microorganismo de control no transformado. En algunos experimentos, el vector de expresión es diseñado para la expresión extracelular de la hidrolasa predicha. Un incremento en la hidrólisis del sustrato puede indicar que la hidrolasa predicha es de hecho una hidrolasa.

Ejemplo 5. Evaluación de la Actividad de transportadores ABC

Puede emplearse una variedad de métodos para evaluar la actividad biológica de un transportador ABC. En un ejemplo, se aísla un gen o genes predichos identificados en *C. phytofermentans* que codifican un transportador ABC y se clona en un vector de expresión. El vector de expresión se transforma en un microorganismo, por ejemplo, *E. coli*. La actividad del gen expresado se suministra al microorganismo transformado con el sustrato del transportador ABC predicho y midiendo el transporte del sustrato a la célula, y comparando el nivel de esta captación, con la captación de un microorganismo de control no transformado. Un aumento en la captación puede indicar que el transportador ABC predicho es un transportador ABC.

Ejemplo 6. Evaluación de la Actividad Reguladora transcripcional

Puede emplearse una variedad de métodos para evaluar la actividad biológica de un regulador transcripcional predicho. En un ejemplo, se aísla un gen predicho identificado en *C. phytofermentans* que codifica un regulador transcripcional y se clona en un vector de expresión. El vector de expresión se transforma en un microorganismo, por ejemplo, *E. coli*. La actividad del gen expresado se mide mediante la co-transfección del organismo transformado con un plásmido que contiene una secuencia de nucleótidos objetivo para el regulador transcripcional y un gen reportero. La actividad del gen reportero se mide y se compara con el nivel de actividad del mismo gen reportero en un microorganismo de control. Un aumento en la actividad del gen reportero indica que el regulador transcripcional predicho puede ser un regulador transcripcional.

Ejemplo 7. Utilización de la Ingeniería Genética de Celobiosa en *E. coli*

La mayoría de las cepas de laboratorio y aislados naturales de *E. coli* no

expresan genes funcionales para la utilización de celobiosa, aunque sí suelen contener genes críptico de utilización celobiosa en sus cromosomas (Hall et al, J Bacteriol, 1987 June, 169: 2713-2717). Se ha modificado genéticamente *E. coli* para utilizar celobiosa por la expresión de Cphy2464-2466, que codifica un transportador ABC y Cphy0430, que codifica un celobiosa fosforilasa que convierte celobiosa en glucosa y glucosa-1-fosfato. Los genes Cphy2464-2466 y Cphy0430 se expresan a partir de un promotor constitutivo en un plásmido. La secuencia de la señal de Cphy2466 se reemplaza con la secuencia señal de una proteína endógena de unión al transportador ABC de *E. coli* para la expresión directa de la proteína en el periplasma. *E. coli* modificada mediante ingeniería genética es capaz de crecer usando celobiosa como única fuente de carbono.

Ejemplo 8. Ingeniería genética de la descomposición mejorada de Pectina en *S. cerevisiae*

Se clonan Cphy1714, Cphy1720 y Cphy3586, en un vector transportador de *E. coli* - *S. cerevisiae* y se expresan de manera heteróloga a partir del plásmido en *S. cerevisiae*. Para permitir la secreción de los productos de genes, las secuencias señal se sustituyen por secuencias señal de proteínas de *S. cerevisiae*. La levadura modificada con ingeniería genética muestra pectinólisis mejorada.

Ejemplo 9. Modificación de Microorganismos

pIMPCphy

El vector pIMPCphy se construyó como un vector transportador para *C. phytofermentans*. Tiene un casete de resistencia a Ampicilina y un origen de Replicación (ORI) para la selección y la replicación en *E. coli*. Contiene un origen de replicación Gram-positivo que permite la replicación del plásmido en *C. phytofermentans*. Para seleccionar la presencia del plásmido, el vector pIMPCphy lleva un gen de resistencia a eritromicina bajo el control del promotor del gen Cphy1029 de *C. phytofermentans*. Este plásmido se transfiere a *C. phytofermentans* mediante electroporación o mediante transconjugación con una cepa de *E. coli* que tiene un plásmido de movilización, por ejemplo pRK2030. En la FIG. 18 se muestra un mapa del plásmido de pIMPCphy.

Promotor constitutivo

En un primer paso, muchos promotores de *C. phytofermentans* fueron escogidos entre los que muestran alta expresión de sus correspondientes genes en todos los estadios de crecimiento así como en diferentes substratos. Un elemento promotor puede ser escogido seleccionando genes clave que podrían necesariamente estar implicados en vías constitutivas (e.g. genes ribosómicos o para la producción de etanol, genes de alcohol deshidrogenasa). Ejemplos de promotores que incluyen dichos genes, pero que no están limitados a éstos, son:

Cphy_1029: alcohol deshidrogenasa que contiene hierro

Cphy_3510: proteína que contiene el dominio IgG

Cphy_3925: acetaldehído-CoA/alcohol deshidrogenasa bifuncional

Clonación de Promotores

Diversos promotores en regiones corriente arriba de los genes fueron amplificadas por PCR. Los cebadores de esta reacción de PCR fueron escogidos de tal forma que incluyeran la región del promotor, pero no los sitios de enlace ribosomal del gen corriente abajo. Los cebadores fueron diseñados para introducir sitios de restricción al final de los fragmentos del promotor que están presentes en un sitio de múltiple clonación de pIMPCphy, pero que por otro lado no están presentes en la región del promotor por si misma, por ejemplo Sail, BamHI, XmaI, SmaI, EcoRI.

La reacción PCR fue llevada a cabo con un Kit PCR que se encuentra en el mercado, GoTaq™ Green Master Mix (Promega), de acuerdo a las condiciones del fabricante. La reacción se realizó en un termociclador, Gene Amp System 24 (Perkin Elmer). Los productos de la PCR fueron purificados con GenElute™ PCR Clean-Up Kit (Sigma). Los dos productos purificados, lo mismo que el plásmido pIMPCphy

fueron llevados a digestión con las enzimas correspondientes en cantidades apropiadas según las condiciones del fabricante (Enzimas de restricción de New England Bio Labs and Promega). Los productos de la PCR fueron subsecuentemente ligados al plásmido con el Quick Ligation Kit (New England Biolabs) y las células competentes de E. coli (DH5α) fueron transformadas con las mezclas de ligación y colocadas sobre placas LB con 1 µg/ml de ampicilina. Las placas son incubadas toda la noche a 37°C.

Las colonias de E.Coli resistentes a la ampicilina fueron removidas de las placas y colocadas en nuevas placas selectivas. Después del crecimiento a 37°C, el medio LB líquido con 1 µg/ml de ampicilina fue inoculado a una única colonia y se dejó crecer a 37°C toda la noche. Los plásmidos fueron aislados del medio con el Gene Elute Plasmid Isolation Kit.

Miniprep Kit

Los plasmidos fueron revisados del inserto derecho por PCR y por la digestión de restricción con los cebadores apropiados y con enzimas de restricción respectivamente. Para asegurar la integridad de la secuencia , el inserto es secuenciado en este paso.

Clonación de Genes Celulasa.

Uno o más genes celulasa pueden incluir sus propios sitios de enlace ribosómico característicos del gen, éstos son amplificados via PCR, y subsecuentemente digeridos con las enzimas apropiadas como se ha descrito anteriormente en Clonación del Promotor. Los plásmidos resultantes son también tratados con las enzimas de restricción correspondientes y los genes amplificados son movilizados dentro de los plásmidos a través de ligación estándar. El plásmido pCphyP3510-3367 (FIG. 19; SEQ ID NO: 1) es creado por ligación de Cphy_3367 de corriente baja y del promotor Cphy_3510 . E. Coli es transformado con los plásmidos, y los insertos correctos son verificados de los transformantes seleccionados sobre las placas de selección.

Transconjugación

El *E. coli* DH5 α junto con el plásmido ayudante pRK2030, son transformados con los diferentes plásmidos discutidos previamente. Las colonias de *E. Coli* con los dos plásmidos precedentes son seleccionados de las placas LB con 1 μ g/ml de ampicilina y 50 μ g/ml de kanamycina después de ser cultivadas por toda una noche a 37°C. El medio de cultivo de *E. Coli* (e.g. LB o LB enriquecido con 1% de Glucosa y 1% de Celobiosa) es inoculado a una única colonia y luego se permite crecer aeróbicamente a 37°C o anaeróbicamente a 35°C durante toda una noche. El medio de cultivo fresco es inoculado 1:1 con los cultivos de toda la noche y se cultiva hasta la fase mid log. Los A.C. fitofermentantes son también cultivados in el mismo medio hasta mid log.

Los dos cultivos diferentes, *C. Fitofermentantes* y *E. Coli* con pRK2030 y uno de los plásmidos, son luego mezclados in diferentes relaciones, e.g 1:10, 1:1, 1:10, 1:1, 10:1, 1:1, 10:1. El acoplamiento se realiza en cada medio líquido, sobre placas o sobre membranas Nucleopore™ Track-Etch (Whatman) a 35 °C. El tiempo se varía entre 2 y 24 horas, y el acoplamiento de los medios es el mismo medio de cultivo en el que el cultivo ha crecido antes del acoplamiento. Después del procedimiento de acoplamiento, la mezcla de bacterias es separada directamente sobre las placas o el primer cultivo sobre el medio líquido por 6 horas a 18 horas y luego sembradas. Las placas contienen 10 μ g/ml de eritromicina como agente selectivo de fitofermentantes and 10 μ g/ml de Trimetoprim, 150 μ g/ml de Ciclosporina, and 1 μ g/ml de ácido Nalidíxico como un medio, contador selectivo para *E. Coli*.

Después de 3 a 5 días de incubación a 35°C, las colonias resistentes a la eritromicina son removidas de las placas y sembradas nuevamente en placas selectivas frescas. Las colonias individuales son removidas y se confirma la presencia del plásmido por PCR.

Expresión de Gen de Celulasa

Posteriormente se evalúa la expresión de los genes celulasa en los diferentes plásmidos bajo condiciones donde hay una expresión mínima o nula de los genes

correspondientes del locus cromosómico. Los candidatos positivos muestran expresión constitutiva de las celulasas clonadas.

ID de SEC No: 1

Plásmido - pCphyP3510-3367

1	ccgggaattc	actggccgtc	gtttttacaac	gtcgtgactg	ggaaaaccct	ggcgttaccc
61	aacttaatcg	ccttgcagca	catccccctt	tcgccagctg	gcgtaatagc	gaagaggccc
121	gcaccgatcg	cccttcccaa	cagttgcgca	gcctgaatgg	cgaatggcgc	ctgatgcggt
181	atthttctcct	tacgcatctg	tgcggtatth	cacaccgcat	atggtgcact	ctcagtacaa
241	tctgctctga	tgccgcatag	ttaagccagc	cccgacaccc	gccaacaccc	gctgacgcgc
301	cctgacgggc	ttgtctgctc	ccggcatccg	cttacagaca	agctgtgacc	gtctccggga
361	gctgcatgtg	tcagaggtht	tcaccgtcat	caccgaaacg	cgcgagacga	aagggcctcg
421	tgatacgctt	atthtttatag	gttaatgtca	tgataataat	ggthttcttag	acgtcagggt
481	gcactthttcg	gggaaatgtg	cgcggaaccc	ctatthtgtht	atthtttctaa	atacattcaa
541	atatgtatcc	gctcatgaga	caataaccct	gataaatgct	tcaataatat	tgaanaagga
601	agagtatgag	tattcaacat	ttccgtgtcg	cccttatthcc	ctthttttgcg	gcattthtgcc
661	thctctgthtt	tgctcaccca	gaaacgctgg	tgaaagtaaa	agatgctgaa	gatcagthtg
721	gtgcacgagt	gggttacatc	gaactggatc	tcaacagcgg	taagatcctt	gagagthttc
781	gccccgaaga	acgtthttcca	atgatgagca	ctthttaaagt	tctgctatgt	ggcgcggtat
841	tatcccgat	tgacgccggg	caagagcaac	tcggtcgccg	catacactat	tctcagaatg
901	acttggttga	gtactcacca	gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag
961	aattatgcag	tgctgccata	accatgagtg	ataaactgc	ggccaactta	cttctgacaa
1021	cgatcggagg	accgaaggag	ctaaccgctt	thtttgacaa	catgggggat	catgtaactc
1081	gccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca
1141	cgatgcctgt	agcaatggca	acaacgttgc	gcaaactatt	aactggcgaa	ctacttactc

1201	tagcttcccg	gcaacaatta	atagactgga	tggaggcgga	taaagttgca	ggaccacttc
1261	tgcgctcggc	ccttcgggt	ggctggttta	ttgctgataa	atctggagcc	ggtagcggtg
1321	ggtctcgcgg	tatcattgca	gcactggggc	cagatggtaa	gccctcccgt	atcgtagtta
1381	tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	tagacagatc	gctgagatag
1441	gtgcctcact	gattaagcat	tggtaactgt	cagaccaagt	ttactcatat	atactttaga
1501	ttgatttaaa	acttcatttt	taattttaaaa	ggatctaggt	gaagatcctt	tttgataatc
1561	tcatgaccaa	aatcccttaa	cgtgagtttt	cgttccactg	agcgtcagac	cccgtagaaa
1621	agatcaaagg	atcttcttga	gacccctttt	ttctgcgcgt	aatctgctgc	ttgcaaacaa
1681	aaaaaccacc	gctaccagcg	gtggtttggt	tgccggatca	agagctacca	actctttttc
1741	cgaaggtaac	tggcttcagc	agagcgcaga	taccaaatat	tgtccttcta	gtgtagccgt
1801	agttaggcca	ccacttcaag	aactctgtag	caccgcctac	atacctcgct	ctgctaattc
1861	tgttaccagt	ggctgctgcc	agtggcgata	agtcgtgtct	taccgggttg	gactcaagac
1921	gatagttacc	ggataaggcg	cagcggtcgg	gctgaacggg	gggttcgtgc	acacagccca
1981	gcttggagcg	aacgacctac	accgaactga	gataacctaca	gcgtgagcta	tgagaaagcg
2041	ccacgcttcc	cgaagggaga	aaggcggaca	ggtatccggt	aagcggcagg	gtcggaacag
2101	gagagcgcac	gagggagctt	ccagggggaa	acgcctggta	tctttatagt	cctgtcgggt
2161	ttcgccacct	ctgacttgag	cgctgatttt	tgtgatgctc	gtcagggggg	cggagcctat
2221	ggaaaaacgc	cagcaacgcg	gcctttttac	ggttcctggc	cttttgctgg	ccttttgctc
2281	acatgttctt	tcctgcgtta	tcccctgatt	ctgtggataa	ccgtattacc	gcctttgagt
2341	gagctgatac	cgctcgccgc	agccgaacga	ccgagcgagc	cgagtcagtg	agcgaggaag
2401	cggaagagcg	ccaataacgc	aaaccgcctc	tccccgcgcg	ttggccgatt	cattaatgca
2461	gctggcacga	caggtttccc	gactggaaag	cgggcagtga	gcgcaacgca	attaatgtga
2521	gttagctcac	tcattaggca	ccccaggctt	tacactttat	gcttccggct	cgtatgttgt

2581	gtggaattgt	gagcggataa	caatttcaca	caggaaacag	ctatgaccat	gattacgcca
2641	aagctttggc	taacacacac	gccattccaa	ccaatagttt	tctcggcata	aagccatgct
2701	ctgacgctta	aatgcactaa	tgccttaaaa	aaacattaaa	gtctaacaca	ctagacttat
2761	ttacttcgta	attaagtcgt	taaaccgtgt	gctctacgac	caaaagtata	aaacctttaa
2821	gaactttctt	ttttcttgta	aaaaaagaaa	ctagataaat	ctctcatatc	ttttattcaa
2881	taatcgcata	agattgcagt	ataaatTTaa	cgatcactca	tcatgttcat	atttatcaga
2941	gctccttata	ttttatttcg	atttatttgt	tatttattta	acatttttct	attgacctca
3001	tcttttctat	gtgttattct	tttgTTaatt	gtttacaaat	aatctacgat	acatagaagg
3061	aggaaaaact	agtatactag	tatgaacgag	aaaaatataa	aacacagtca	aaactttatt
3121	acttcaaaac	ataatataga	taaaataatg	acaaatataa	gattaaatga	acatgataat
3181	atctttgaaa	tcggctcagg	aaaagggcat	tttacccttg	aattagtaca	gaggtgtaat
3241	ttcgtaactg	ccattgaaat	agaccataaa	ttatgcaaaa	ctacagaaaa	taaacttggt
3301	gatcacgata	atttccaagt	tttaaacaag	gatatatTgc	agtttaaatt	tcctaaaaac
3361	caatcctata	aaatatTTtg	taatatacct	tataacataa	gtacggatat	aatacgcaaa
3421	attgTTTTtg	atagtatagc	tgatgagatt	tatttfaatcg	tggaatacgg	gtttgctaaa
3481	agattattaa	atacaaaacg	ctcattggca	ttatttttaa	tggcagaagt	tgatatttct
3541	atattaagta	tggttccaag	agaatatTTt	catcctaaac	ctaaagtgaa	tagctcactt
3601	atcagattaa	atagaaaaaa	atcaagaata	tcacacaaaag	ataaacagaa	gtataattat
3661	ttcgttatga	aatgggttaa	caaagaatac	aagaaaatat	ttacaaaaaa	tcaatttaac
3721	aattccttaa	aacatgcagg	aattgacgat	ttaaacaata	ttagctttga	acaattctta
3781	tctcttttca	atagctataa	attatttfaat	aagtaagtta	agggatgcat	aaactgcatc
3841	ccttaacttg	tttttcgtgt	acctattTTt	tgtgaatcga	tccggccagc	ctcgcagagc
3901	aggattcccg	ttgagcaccg	ccaggtgcga	ataagggaca	gtgaagaagg	aacacccgct

3961	cgcggtggg	cctacttcac	ctatcctgcc	cggatcgatt	atgtcttttg	cgcatcact
4021	tcttttctat	ataaatatga	gcgaagcgaa	taagcgtcgg	aaaagcagca	aaaagtttcc
4081	tttttgctgt	tggagcatgg	gggttcaggg	ggtgcagtat	ctgacgtcaa	tgccgagcga
4141	aagcgagccg	aagggtagca	tttacgttag	ataaccccct	gatatgctcc	gacgctttat
4201	atagaaaaga	agattcaact	aggtaaaatc	ttaatatagg	ttgagatgat	aaggtttata
4261	aggaatttgt	ttgttcta	ttttcactca	ttttgttcta	atttctttta	acaaatgttc
4321	tttttttttt	agaacagtta	tgatatagtt	agaatagttt	aaaataagga	gtgagaaaaa
4381	gatgaaagaa	agatatggaa	cagtctataa	aggctctcag	aggctcatag	acgaagaaag
4441	tggagaagtc	atagaggtag	acaagttata	ccgtaaacaa	acgtctggta	acttcgtaaa
4501	ggcatatata	gtgcaattaa	taagtatggt	agatatgatt	ggcggaaaaa	aacttaaaat
4561	cgttaactat	atcctagata	atgtccactt	aagtaacaat	acaatgatag	ctacaacaag
4621	agaaatagca	aaagctacag	gaacaagtct	acaaacagta	ataacaacac	ttaaaatctt
4681	agaagaagga	aatattataa	aaagaaaaac	tggagtatta	atgttaaacc	ctgaactact
4741	aatgagaggc	gacgaccaa	aacaaaaata	cctcttactc	gaatttgagg	actttgagca
4801	agaggcaa	gaaatagatt	gacctccaa	taacaccacg	tagttattgg	gaggtcaatc
4861	tatgaaatgc	gattaagctt	agcttggctg	caggtcgaca	gacagcataa	gtcacatcca
4921	gacaaatgtc	ctataggatg	ttagtagggg	tttgagaaat	tgcccgtaa	gcaggttatt
4981	tggctagata	taatcaatcc	agttacagga	tagtaggatt	gcaaccag	cgttttgacc
5041	agtttgtaca	agaattttta	ttgtcgaaa	tattgtggca	aatcaa	agttctttga
5101	tgaaatgttt	agaaacatga	cttagaatgg	ggtacaaaa	gtgaattt	aagcaaaaag
5161	acttgacctt	tcctacgata	gttggtataa	tcatcttg	attggaacga	ttatatttac
5221	ttatgcacat	tttagagttt	ttcgaattgt	taatacatca	ttaacaattt	aattatactc
5281	gttatgtgac	gtaagtcaat	ataatacaaa	accatatatt	ttaagccgcg	ggcagaaagg

5341	atgagagata	tgaaaaagat	aataagtctt	ttattagtga	taacacttct	gatatccatg
5401	gcaccatcga	aagctgacgc	agcggaaacc	aattataatt	acggagaagc	tcttcaaaaa
5461	tcaatcatgt	tttatgagtt	tcaacgttct	ggtaaactgc	caagtaccat	tcggaataat
5521	tggagaggtg	actctggttt	aaccgatgga	gcagatgttg	gtttggatct	aactggtggc
5581	tggtatgatg	ctggtgatca	tgtaaaatth	aatcttcctt	tggcttatac	tgtaacaatg
5641	ttagcatggg	cagtatatga	agaagaggct	actctttcaa	aggcaggcca	attaagttat
5701	ttattagatg	aaattaagtg	gtctagtgat	tacctaatta	aatgtcatcc	acaagcaaat
5761	gtattttatt	atcaggttgg	taatggaaat	acagatcact	cttgggtggg	acctgctgaa
5821	gttatgcaga	tggctagacc	gtcctataag	gttgatttaa	ataaccagg	ttctactgta
5881	gtaggagaag	cagcagcagc	tcttgacgca	acagcactta	tatataagac	aaaagaccct
5941	acttattcag	caacttgcct	tcgtcatgca	aaagagcttt	ttaattttgc	agatacaaca
6001	aaaagcgatg	ctggatatac	agcagcaagt	gggttctata	cttcctatag	tggattttat
6061	gatgaattat	cctgggcagc	tacatggatt	taccttgcaa	gtggagaagc	gacctatttg
6121	gataaggcag	aatcttatgt	agccaaatgg	ggaacagaac	ctcaatcttc	cacattaagt
6181	tataagtggg	cacaaaactg	ggatgatgtt	cactatgggtg	cagctttatt	attagcaaga
6241	attacaaata	aagcaattta	taagaacaat	attgaaatgc	atcttgacta	ttggactaca
6301	gggtataatg	gtagtcgtat	tacttataca	ccaaaaggac	ttgcttggtt	agattcctgg
6361	ggtgcattaa	gatatgacg	gacaacagca	tttctagcaa	gtgtttatgc	tgattggagc
6421	ggatgtagtg	ctggaaaagt	tagtacttac	aatgcatttg	cgaaacagca	ggtagattat
6481	gcattaggaa	gtaccggaag	aagttttgtg	gttgatgatg	gtgtaaattc	tccaacaaga
6541	cctcatcata	gaactgctca	tagttcatgg	gcagacagtc	agacggagcc	aaattaccat
6601	agacacacca	tttatggtgc	tttagtaggt	ggacctggta	ataatgatag	ttatgaggat
6661	aacattaata	attatgtaaa	caatgaaatc	gcttgtgact	ataatgcagg	ttttgttggc

6721	gcattggcta	aagtttataa	aacatatggc	ggaacaccaa	ttgcaaactt	taaggcaatc
6781	gaaacagtaa	caaacgatga	gttattttatt	caagctggta	ttaatgcctc	tggtccatct
6841	tttatcgaag	taaaggcatt	ggttttcaat	gagacaggtt	ggccagctcg	tgttaccgat
6901	aaattatcct	ttaagtattt	tattgatatc	tcggaatatg	tagcaaaggg	atatacaaag
6961	aatgatttta	cggtatcgac	aaattataac	aatggagcaa	ccacatcggc	attgcttcct
7021	tgggatgctg	cgaataatat	ctattatgtg	aatgtagact	tctctggaac	taagatttat
7081	cctggtggac	agtctgcata	taagaaagaa	gtacaattta	gaattgctgg	tccacaaaac
7141	gttaatatat	gggacaattc	caatgactac	tcctttacac	aaattgctaa	tgttagttca
7201	ggaaataccg	taaagaccac	atatatacca	ttgtatgata	atggtaaatt	agtatttggt
7261	aatgagccaa	agacgggtgt	tccttctgca	agtcttgata	agactacagc	aaactttgac
7321	aaaaacccag	ctgtatccgc	agatatacca	gtaaccatta	actataatgg	taatacatta
7381	acagcggtta	agaatggaac	aacggtttta	acgaaaggta	ctgattatac	tgtatctggt
7441	aatgtagtaa	cgttatctaa	gaattatttc	ttagcacaga	gcgctagtag	ggttacttta
7501	acatttgtat	ttagtggcgg	taacgatgca	acattaaaag	tgactttagt	agatacttct
7561	ccaagtgcac	ccattaatcc	aaattctgct	gtctttgata	aggctagcgg	aaaacaggaa
7621	aatatagtta	ttacgcttac	accaaattggc	aataccttag	ctggacttaa	gaatgggtct
7681	aagagcctgg	taactggaac	tgattatacc	gtttccggaa	caacagtgac	gattctatct
7741	tcttatttta	gtcaatttgc	agtaggaagt	caatctattg	tatttgaaat	gaataaaggg
7801	acaaatccag	tcttagcagt	taccattaag	gattcttctg	ttgttactcc	aacaggaaat
7861	attaaacttc	aaatgtttta	tggaaattct	tctgcaacaa	cgaatggcat	tgaccaaga
7921	attaaattaa	ttacaccgg	aactactgca	atcaacttat	ccgatgttaa	gattcgctat
7981	tattatacaa	tcaatggcga	aaaggatcag	gcattctggt	gtgattattc	gacgattggt
8041	agttccaatg	taaatggtac	tttcgtaaag	atgagtacac	caaaaacaaa	tcgagattac

8101 tatctagaat ttccatttaa gtccgctgcc ggaactttaa acgcagggca aagtattgaa
8161 gttcaaggaa gattttctaa ggtagactgg acaaactata cacaacaga tgattattcg
8221 tttggtgata gtaactcaag ttatgctgat tggaataaga caacagtata tatctctgat
8281 gttttgggtt ggggagtcga accataatag gagaaaaaat gtaataattt ttagaggggt
8341 cataacttag tatacatgtc tgtatatgag gtccgacacg tgccacacgg catgtgtcgg
8401 gcctcatttt tatacagcgt gtatgtgacc ttattcatga caagggatcg tccgcc

OTRAS REALIZACIONES

La anterior descripción revela diversos métodos y materiales de la presente invención. Esta invención es susceptible de modificaciones en los métodos y materiales, así como alteraciones en los métodos de fabricación y equipos. Tales modificaciones resultarán evidentes para aquellos versados en la materia al considerar esta divulgación o llevar a la práctica la invención aquí revelada. En consecuencia, no se pretende que esta invención esté limitada a las realizaciones específicas aquí reveladas, sino que cubra todas las modificaciones y alternativas que se incluyen en el verdadero alcance y espíritu de la invención según lo señalado en las reivindicaciones adjuntas.

LO QUE SE REIVINDICA ES:

1. Un producto para la producción de un biocombustible que comprende: una biomasa lignocelulósica, un microorganismo que es capaz de hidrolizar directamente y fermentar dicha biomasa, en donde dicho microorganismo está modificado para proveer actividad mejorada de una o más celulasas.
2. El producto de 1, en donde dicho microorganismo es capaz de fermentar directamente azúcares de cinco de carbonos y de seis carbonos.
3. El producto de la reivindicación 2, en donde dicho microorganismo es una bacteria.
4. El producto de la reivindicación 1, en donde dicho microorganismo es una especie de *Clostridium*.
5. El producto de la reivindicación 1, en donde dicho microorganismo es *Clostridium phytofermentans*.
6. El producto de la reivindicación 1, en donde el microorganismo es no recombinante o recombinante.
7. El producto de la reivindicación 2, en donde el microorganismo comprende uno o más polinucleótidos heterólogos que mejoran dicha actividad de una o más celulasas.
8. Un producto para la producción de un biocombustible que comprende: una biomasa carbonácea, un microorganismo que es capaz de hidrolizar directamente y fermentar dicha biomasa, en donde dicho microorganismo está modificado para proveer actividad mejorada de una o más celulasas.

9. El producto de la reivindicación 8, en donde el microorganismo es capaz de producir productos finales de fermentación, una parte importante de los cuales es etanol.
10. El producto de la reivindicación 8, en donde el microorganismo es capaz de producir productos finales de fermentación que comprenden ácido láctico, ácido acético y ácido fórmico.
11. El producto de la reivindicación 8, en donde dicho microorganismo es capaz de captar uno o más carbohidratos complejos.
12. El producto de la reivindicación 8, en donde dicha biomasa comprende una mayor concentración de carbohidratos oligoméricos con respecto a carbohidratos monoméricos.
13. El producto de la reivindicación 8, en donde dichas una o más celulasas se seleccionan del grupo que consiste de Cphy3367, Cphy3368, Cphy0218, Cphy3207, Cphy2058 y Cphy1163.
14. El producto de la reivindicación 8, en donde dichas una o más celulasas es Cphy3367.
15. Un proceso para producir un biocombustible, comprendiendo el proceso:
(a) poner en contacto una biomasa carbonácea con: un microorganismo que es capaz de hidrolizar directamente y fermentar dicha biomasa, en donde el microorganismo se ha modificado para mejorar la actividad de una o más enzimas celulasas, y (b) permitir que transcurra el tiempo suficiente para que dicha hidrólisis y fermentación produzcan un biocombustible.

16. El proceso de la reivindicación 15, en donde dichas una o más enzimas celulasas se selecciona del grupo que consiste de Cphy3367, Cphy3368, Cphy0218, Cphy3207, Cphy2058 y Cphy1163.

17. El proceso de la reivindicación 15, en donde dicho microorganismo es capaz de captar uno o más carbohidratos complejos.

18. El proceso de la reivindicación 15, en donde dicha biomasa comprende una mayor concentración de carbohidratos oligoméricos con respecto a los carbohidratos monoméricos.

19. El proceso de la reivindicación 18, en donde dicha hidrólisis resulta en una mayor concentración de celobiosa y/o oligómeros más grandes, con relación a los carbohidratos monoméricos.

20. El producto de la reivindicación 15, en donde dichas una o más celulasas es Cphy3367.

21. El nucleótido aislados de acuerdo con el ID de SEC No: 1.

POSIBLE COMBINACIÓN DE GENES

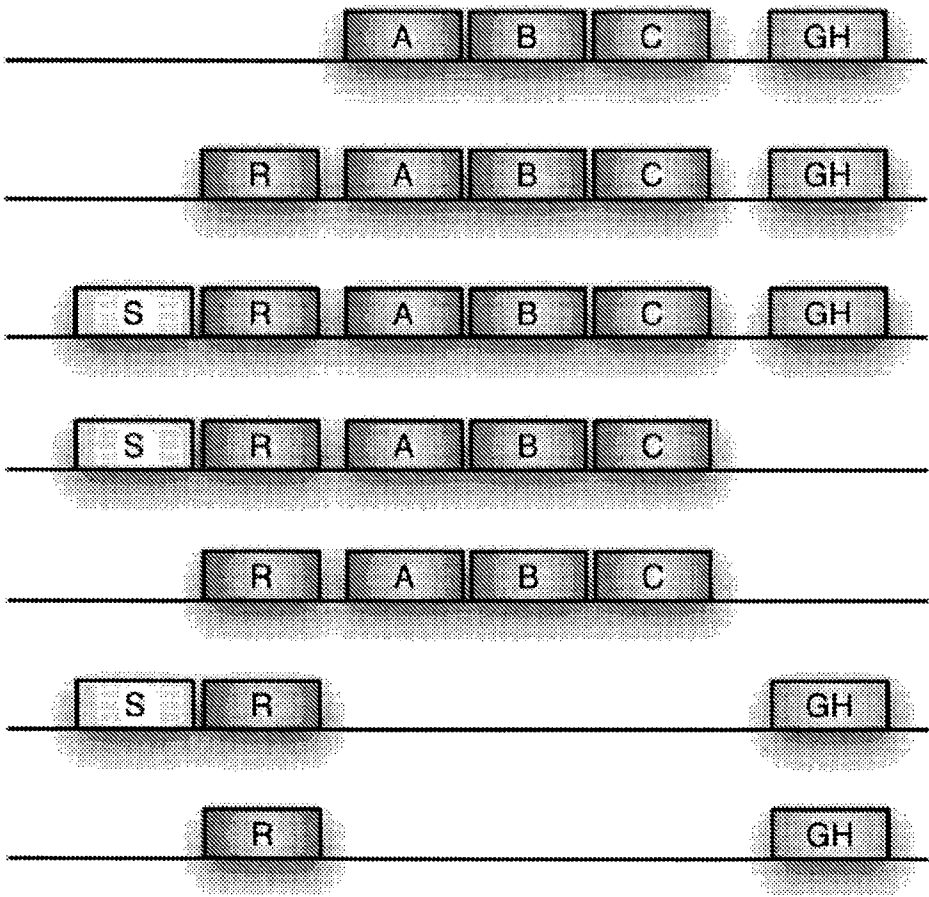


FIG. 1

EJEMPLOS ESPECÍFICOS DE Cphy RELACIONADOS
CON CELULOSA Y XILANO

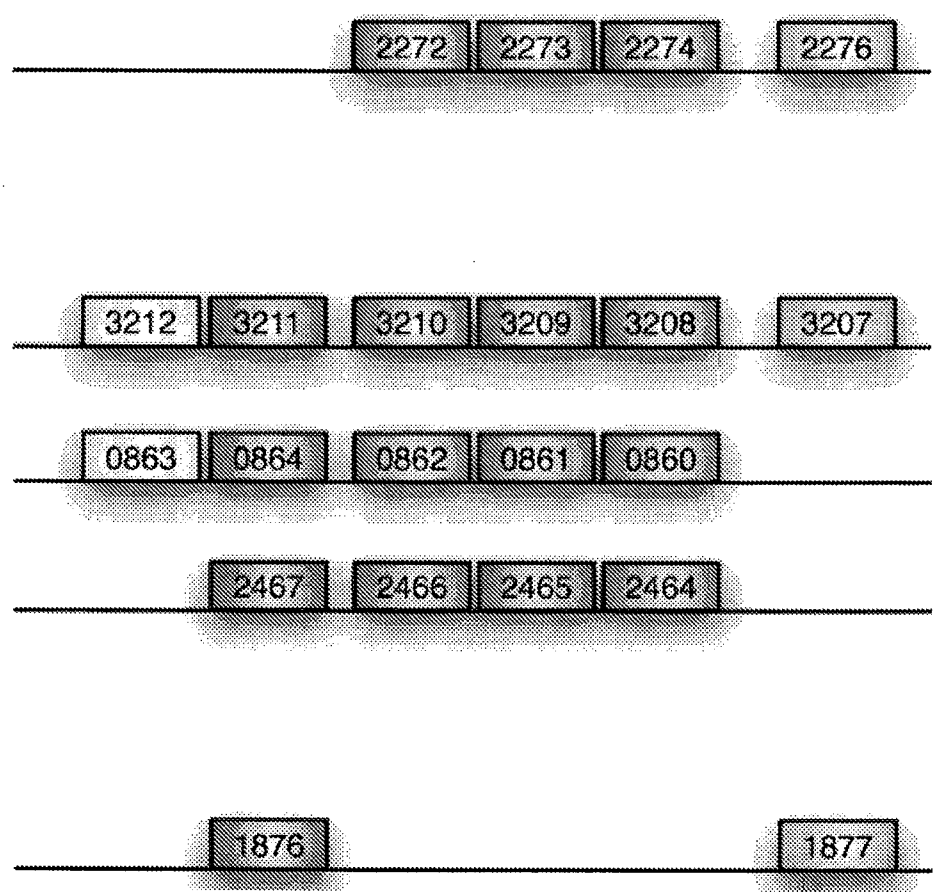


FIG. 2

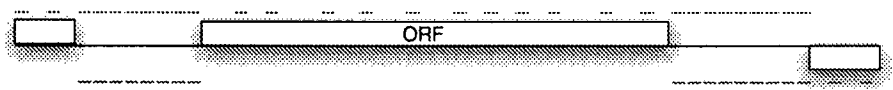


FIG. 3

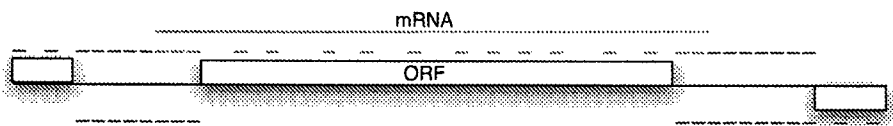


FIG. 4

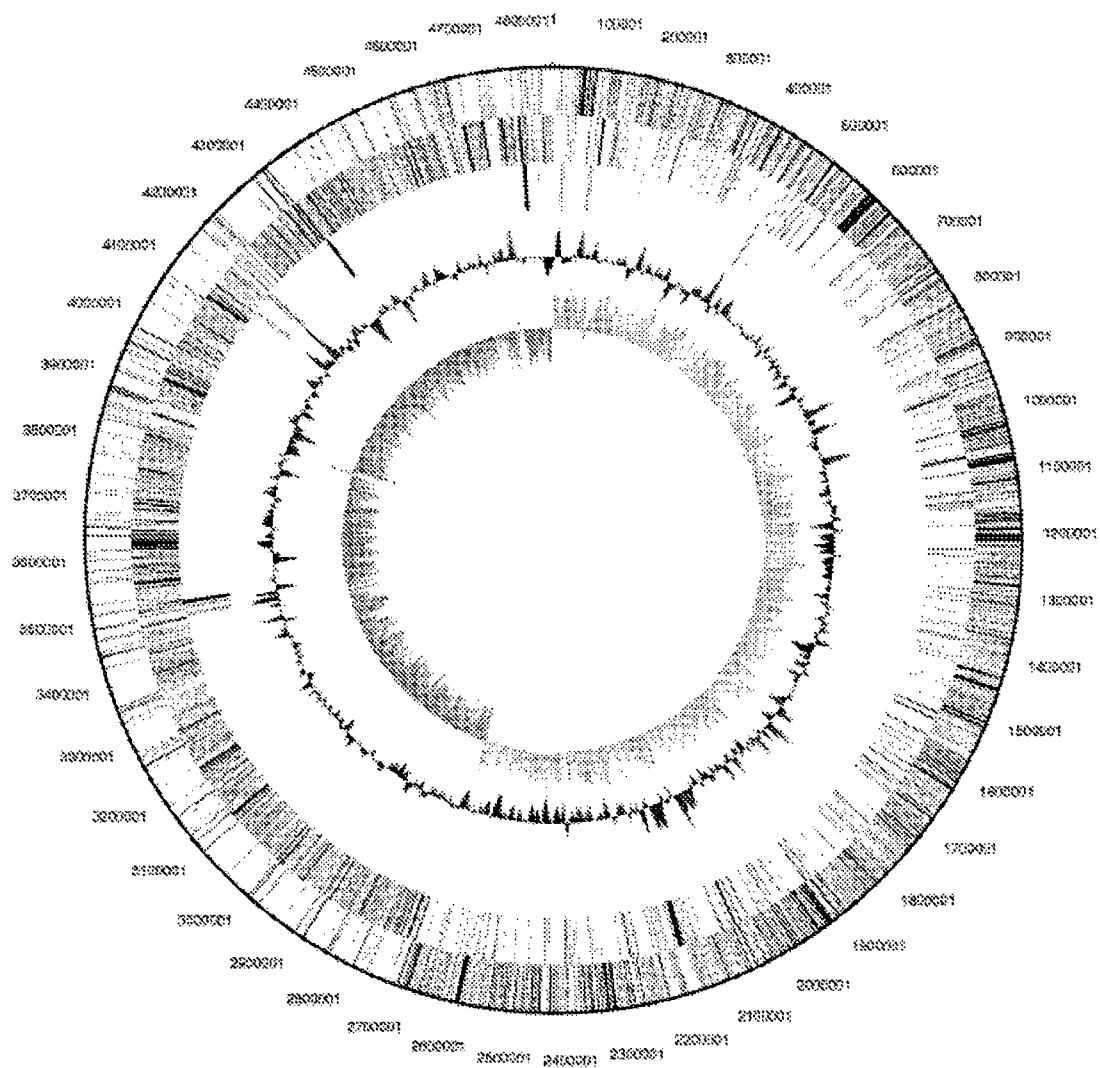


FIG. 5

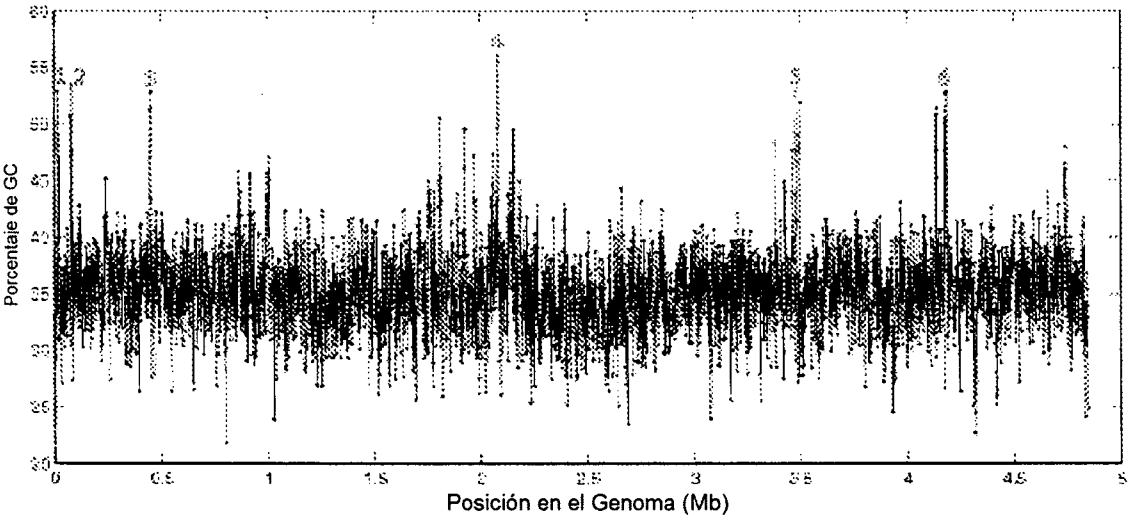


FIG. 6

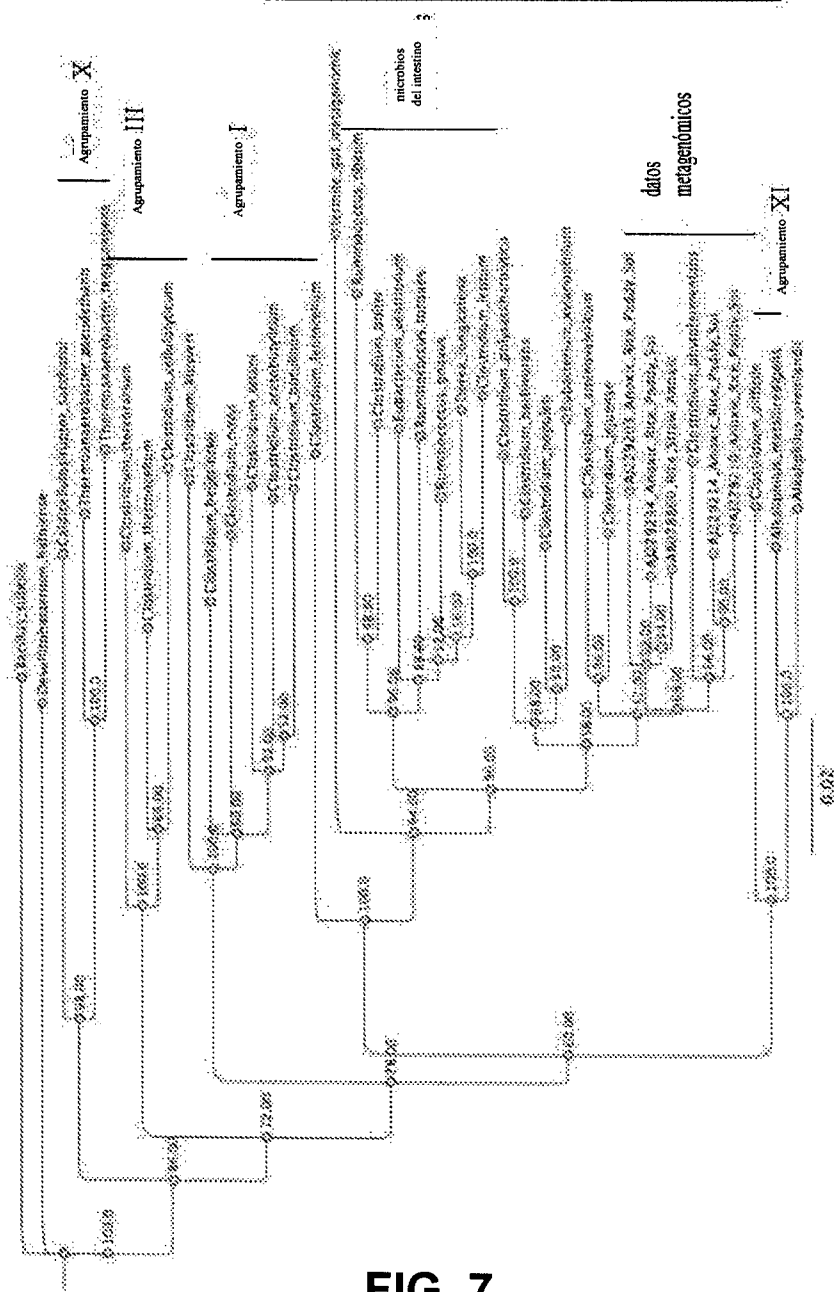


FIG. 7

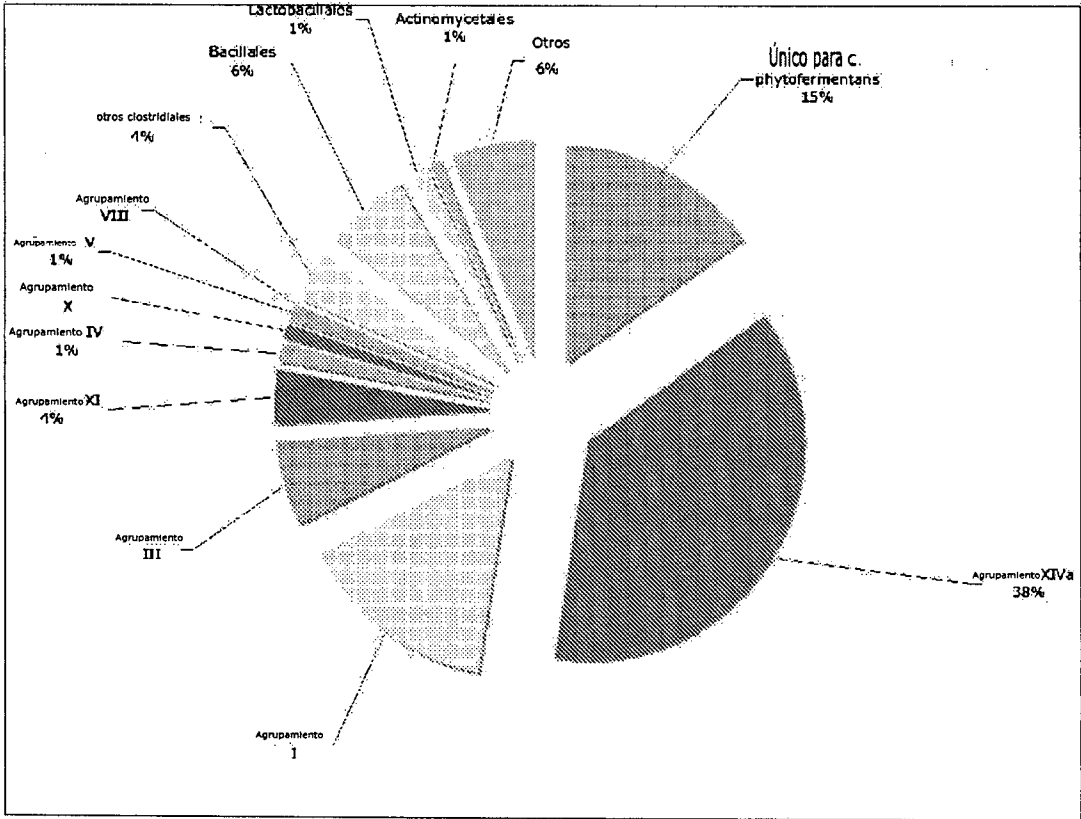
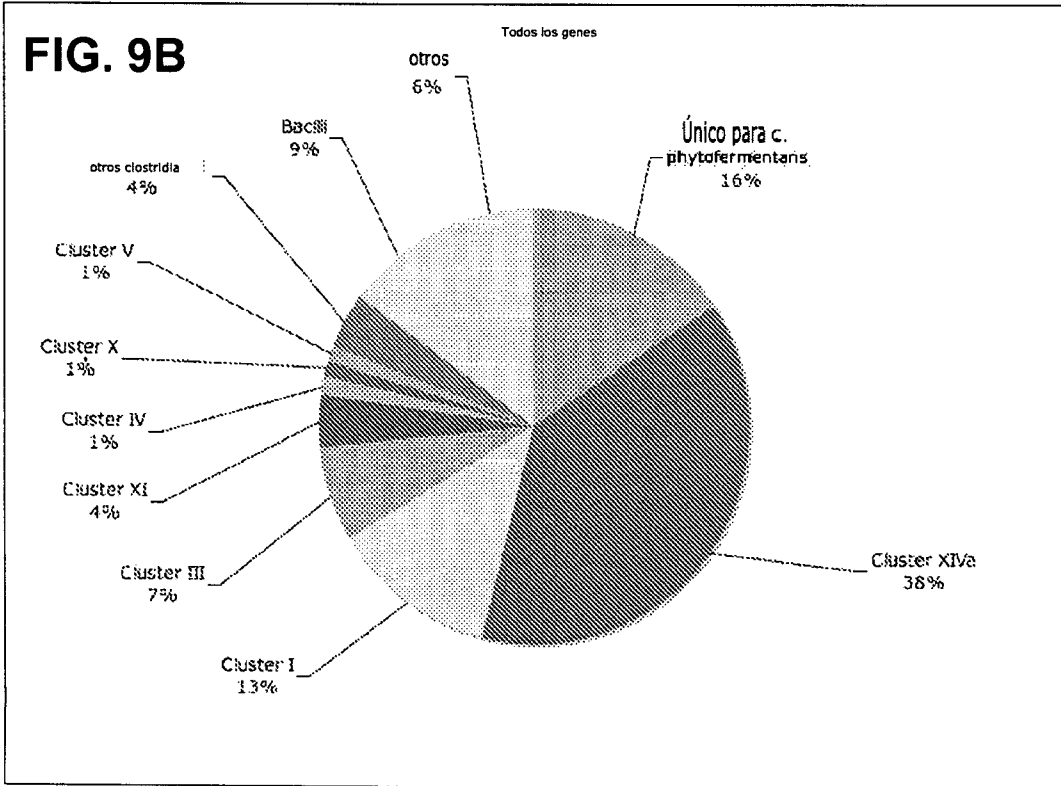
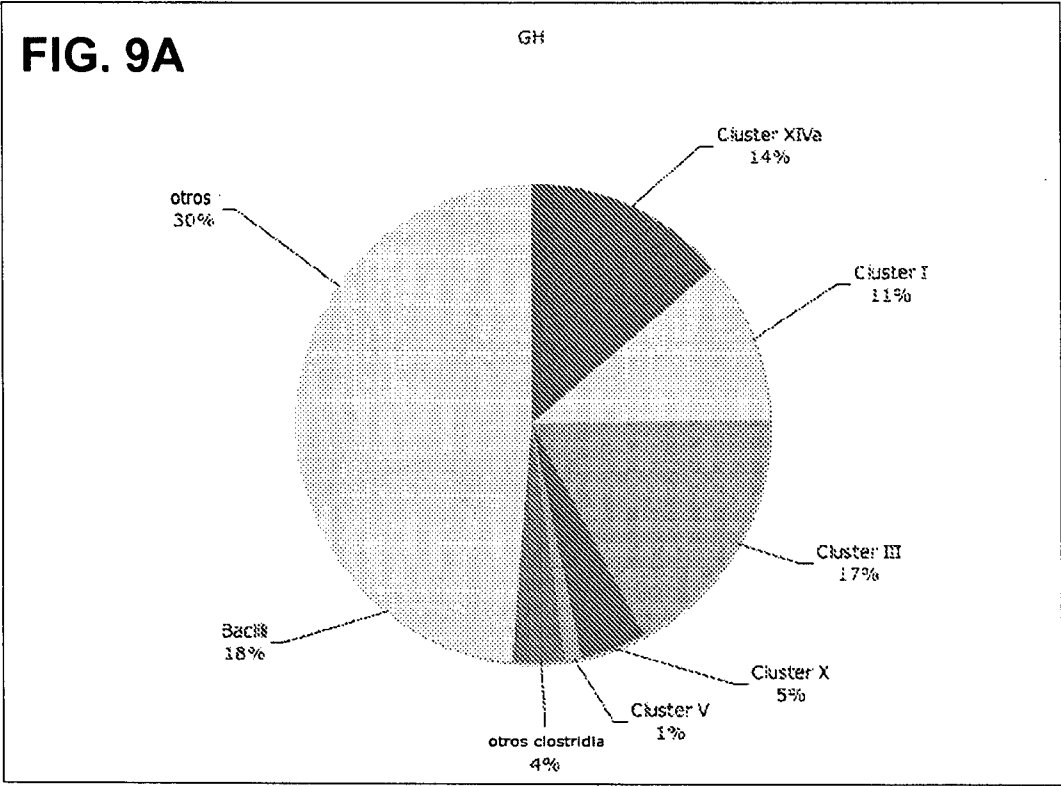


FIG. 8



Cluster=Agrupamiento

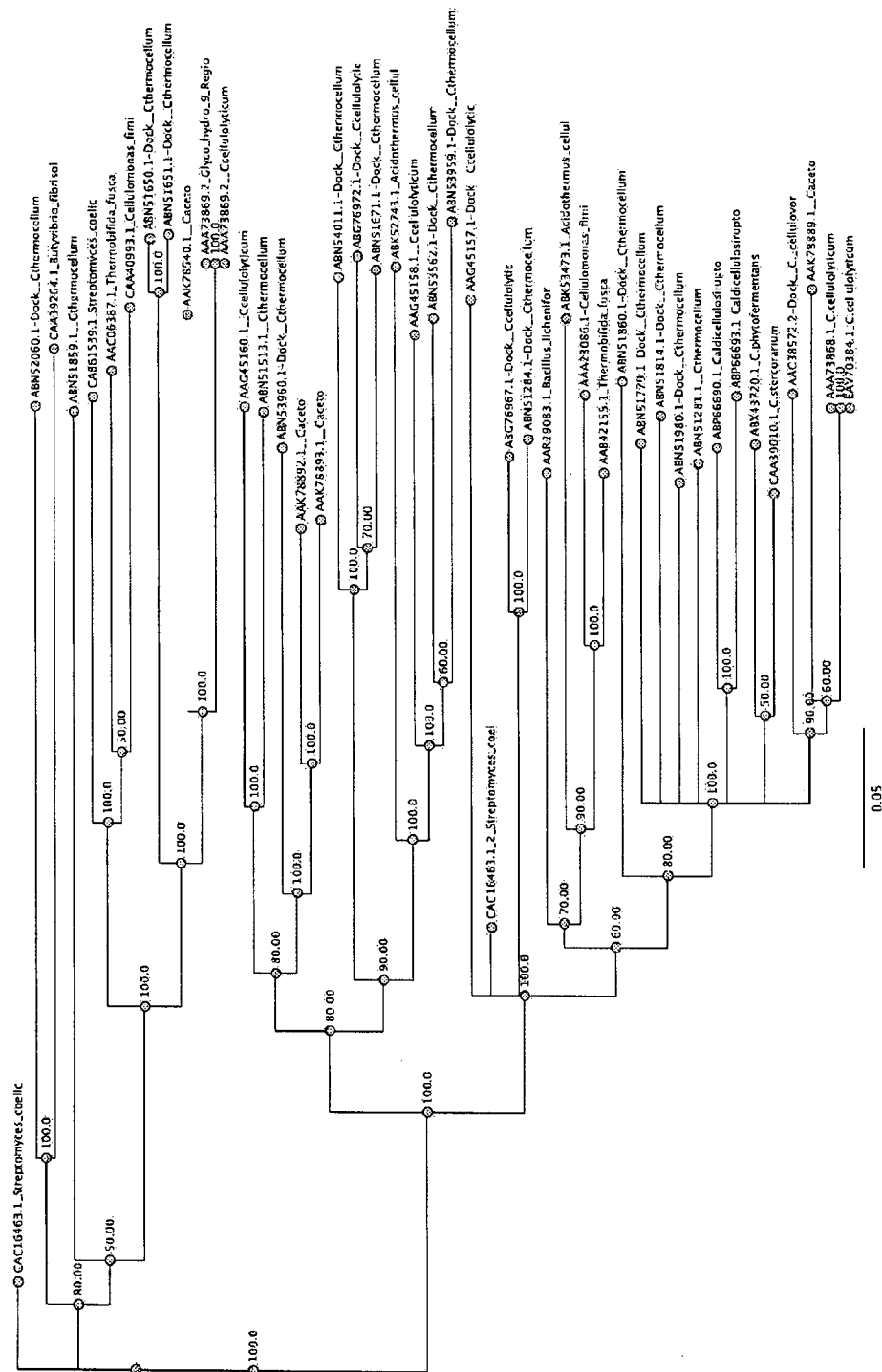


FIG. 10

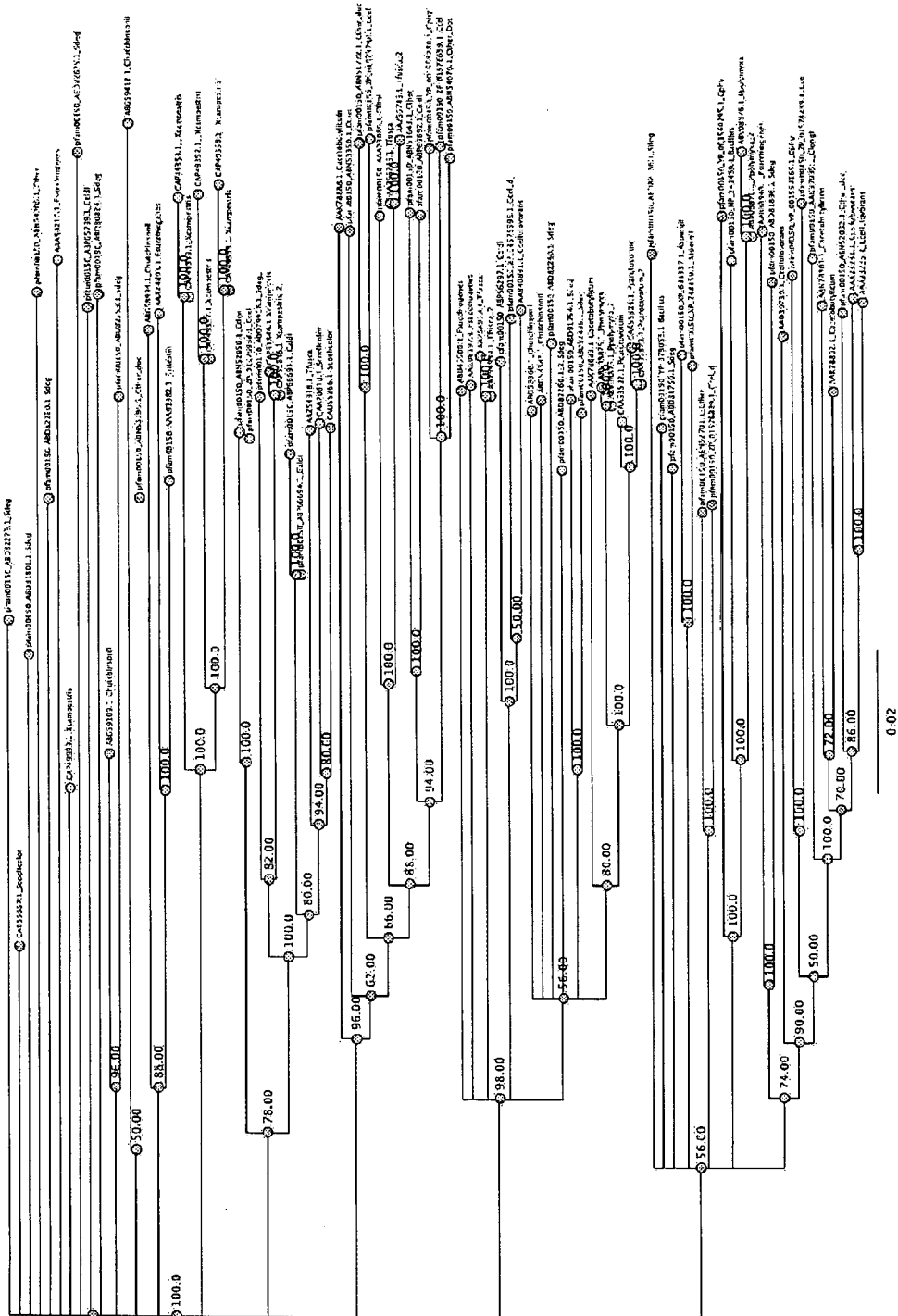


FIG. 11

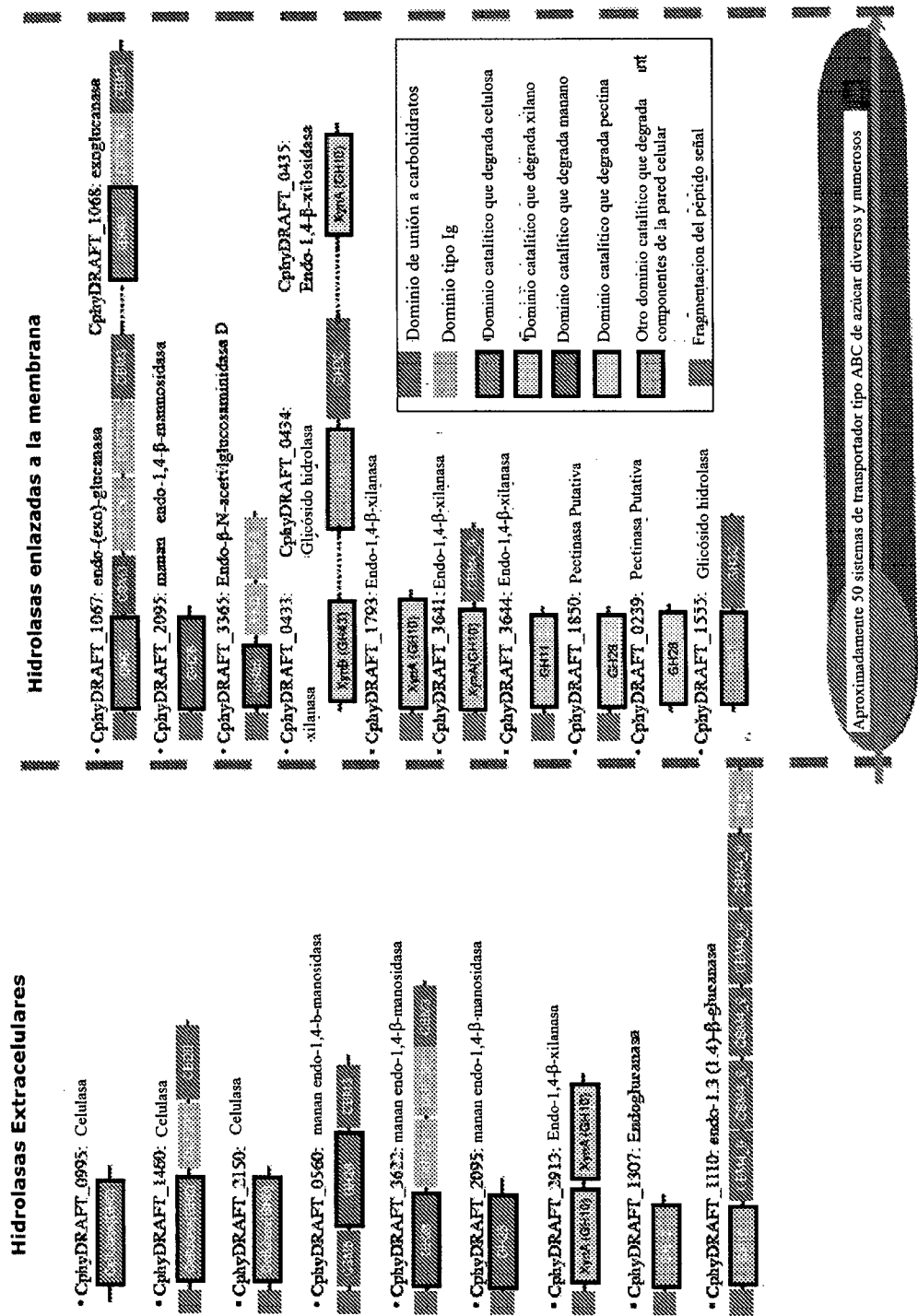


FIG. 12

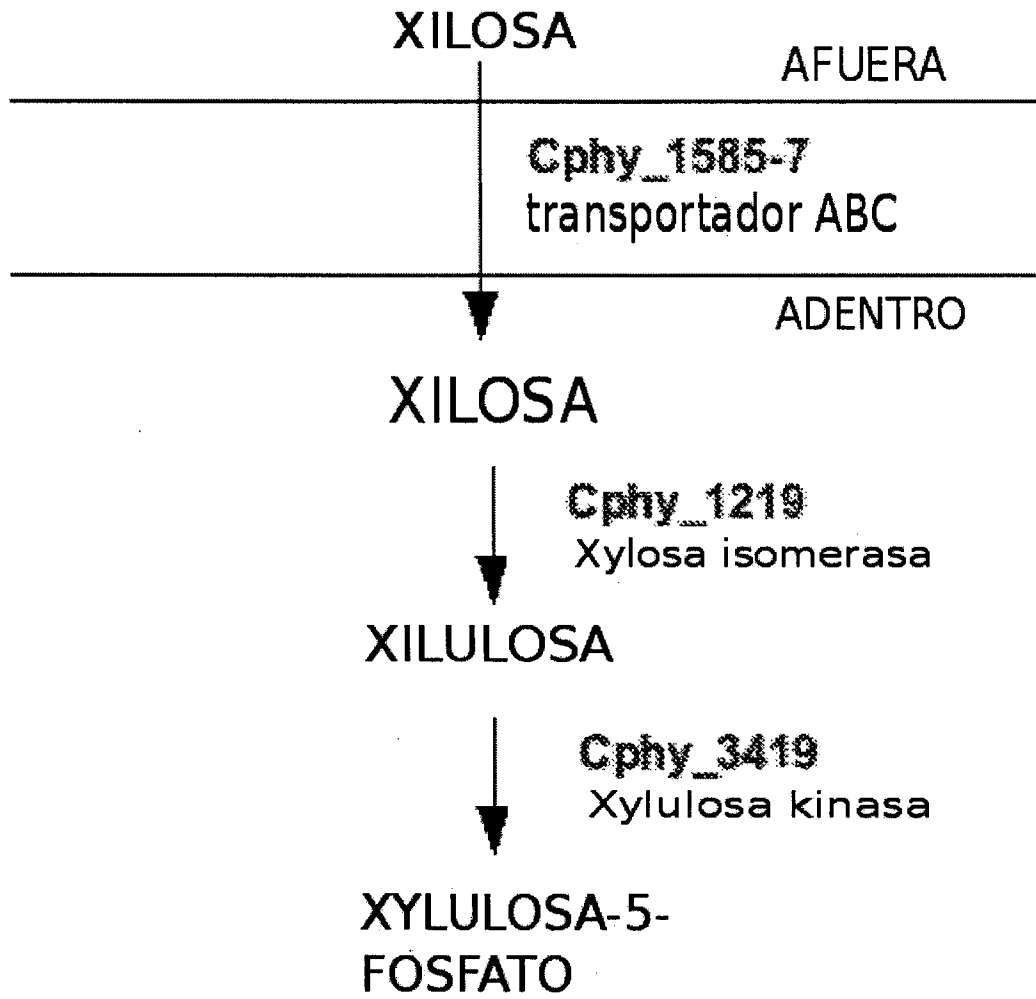


FIG. 13

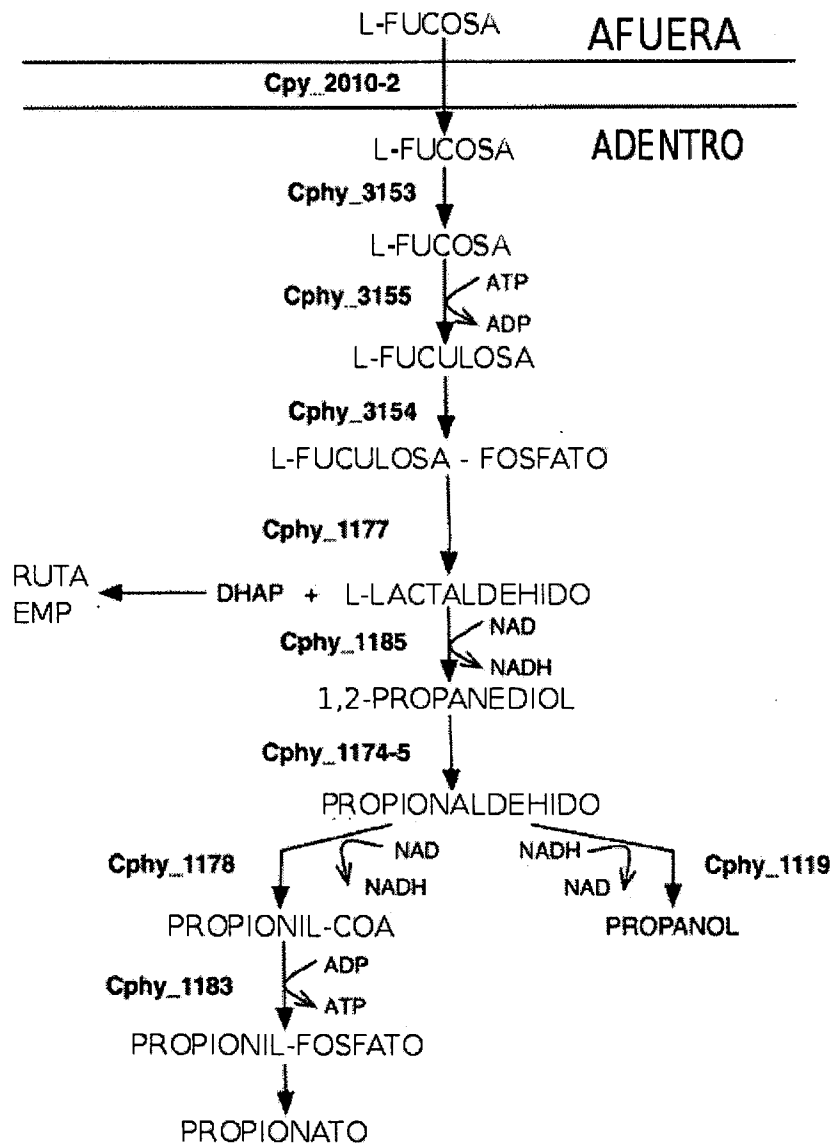


FIG. 14

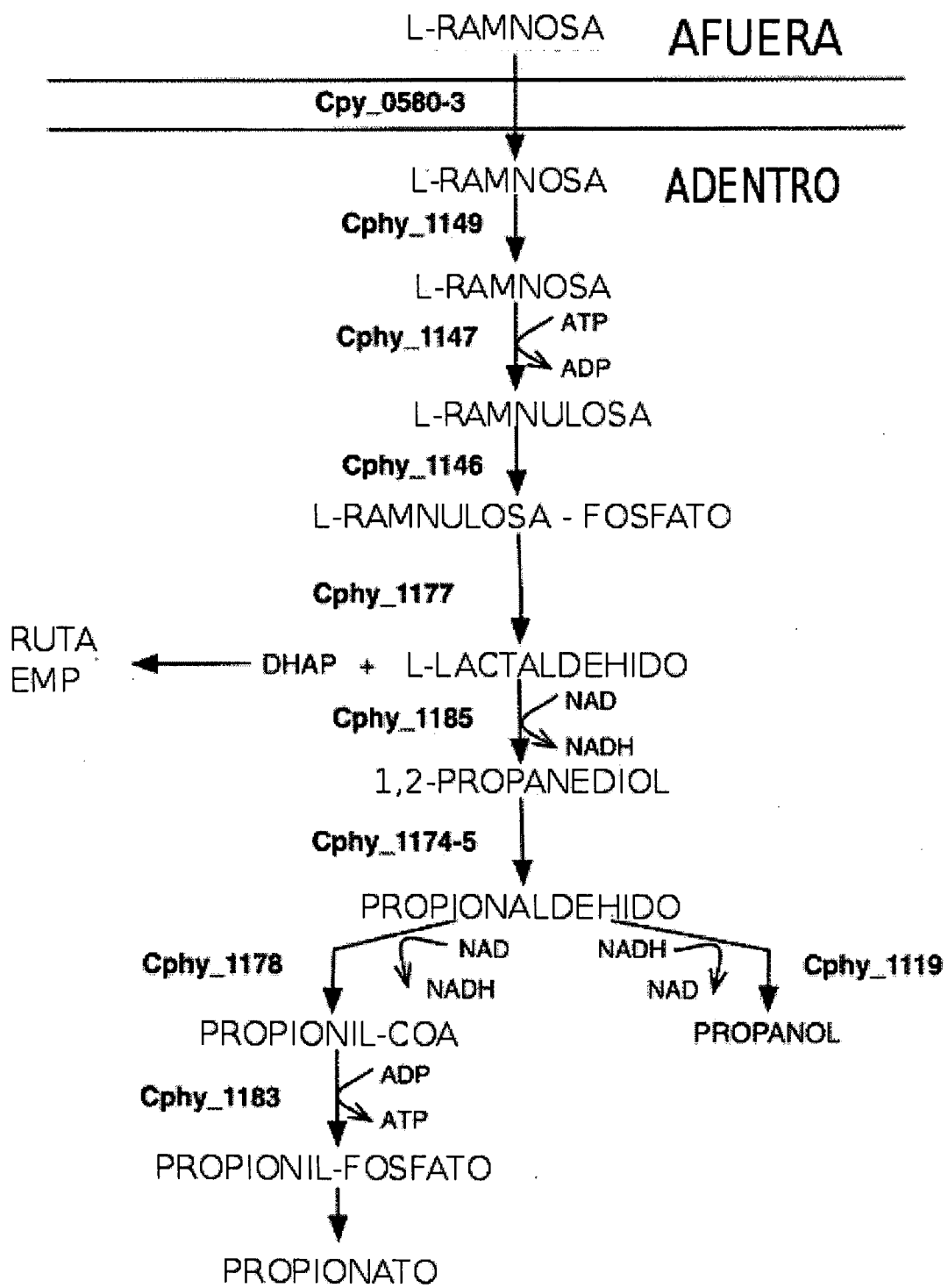


FIG. 15

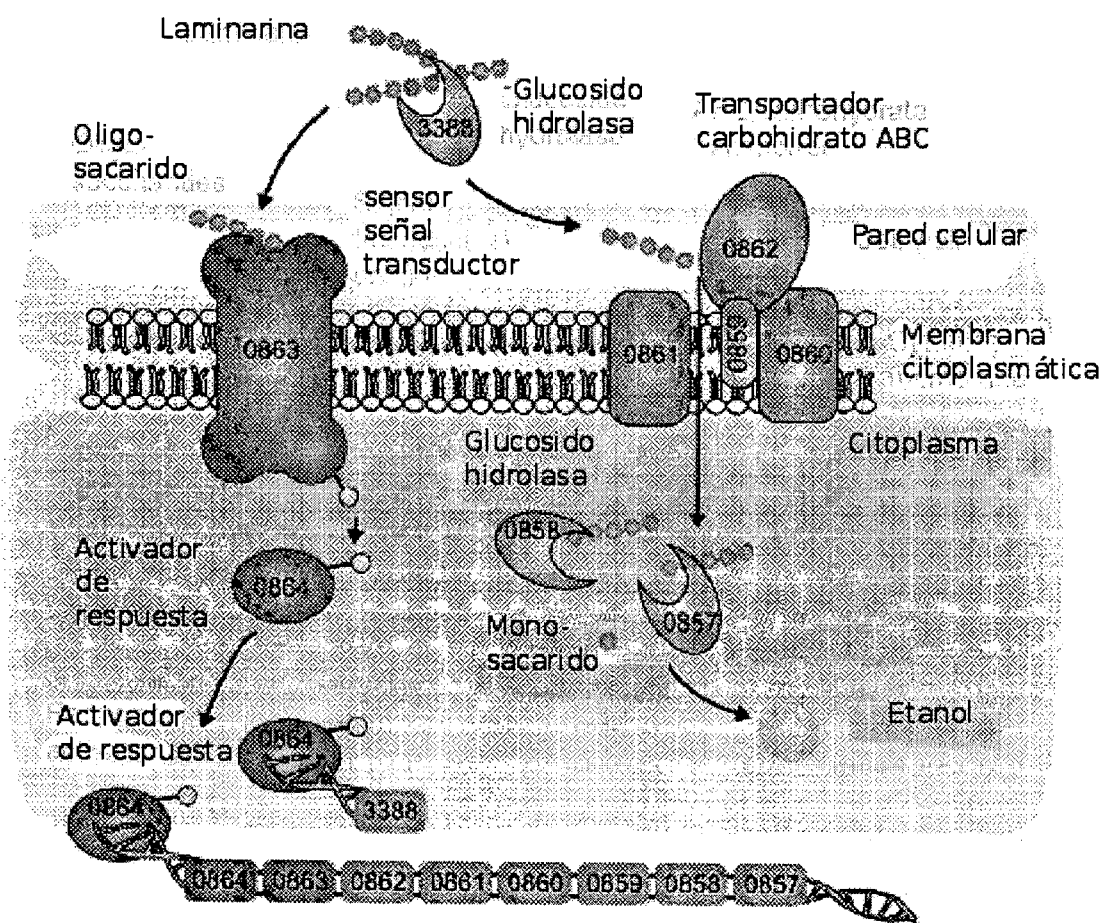


FIG. 16

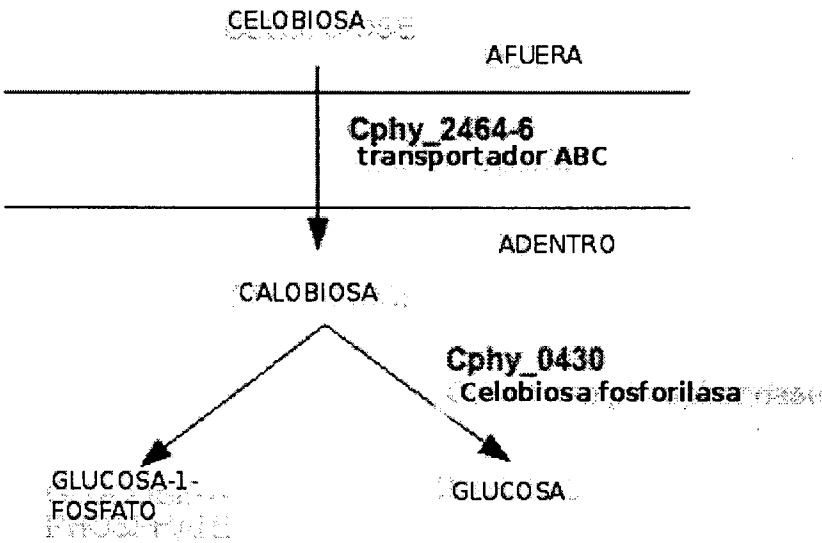


FIG. 17

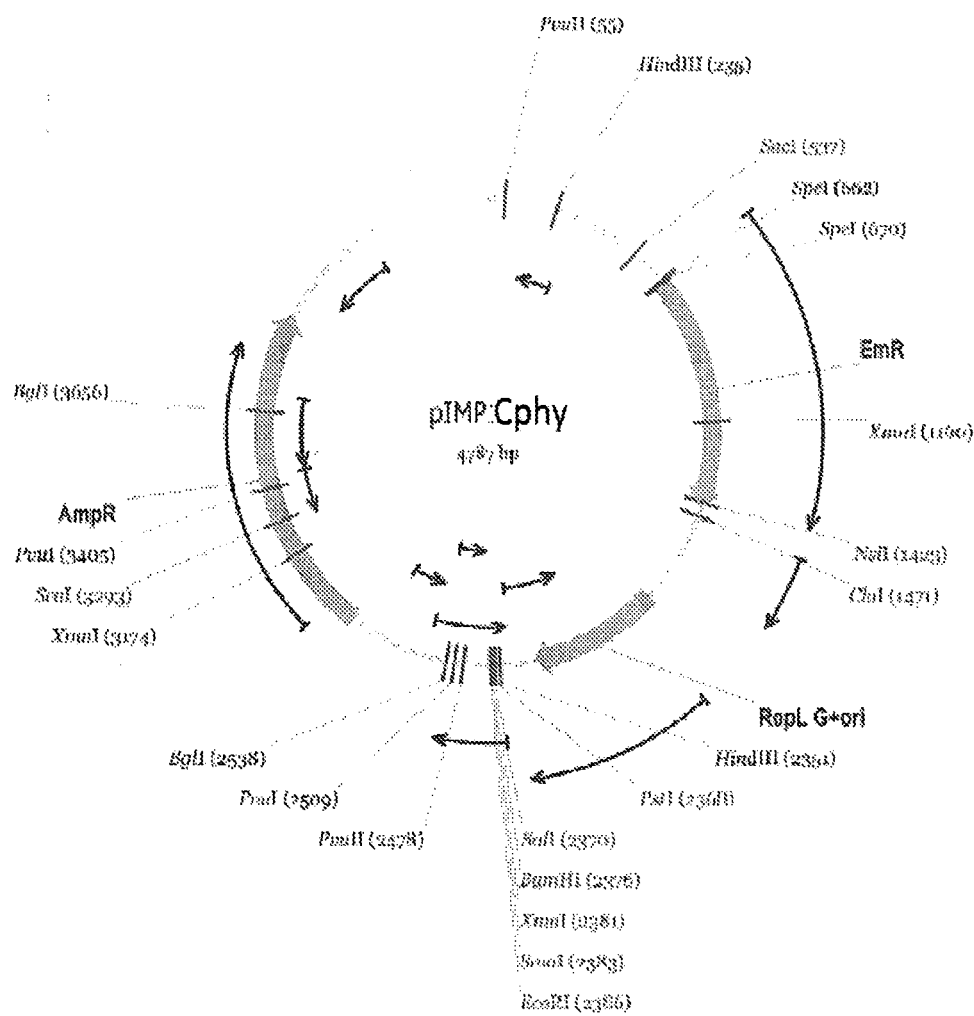


FIG. 18

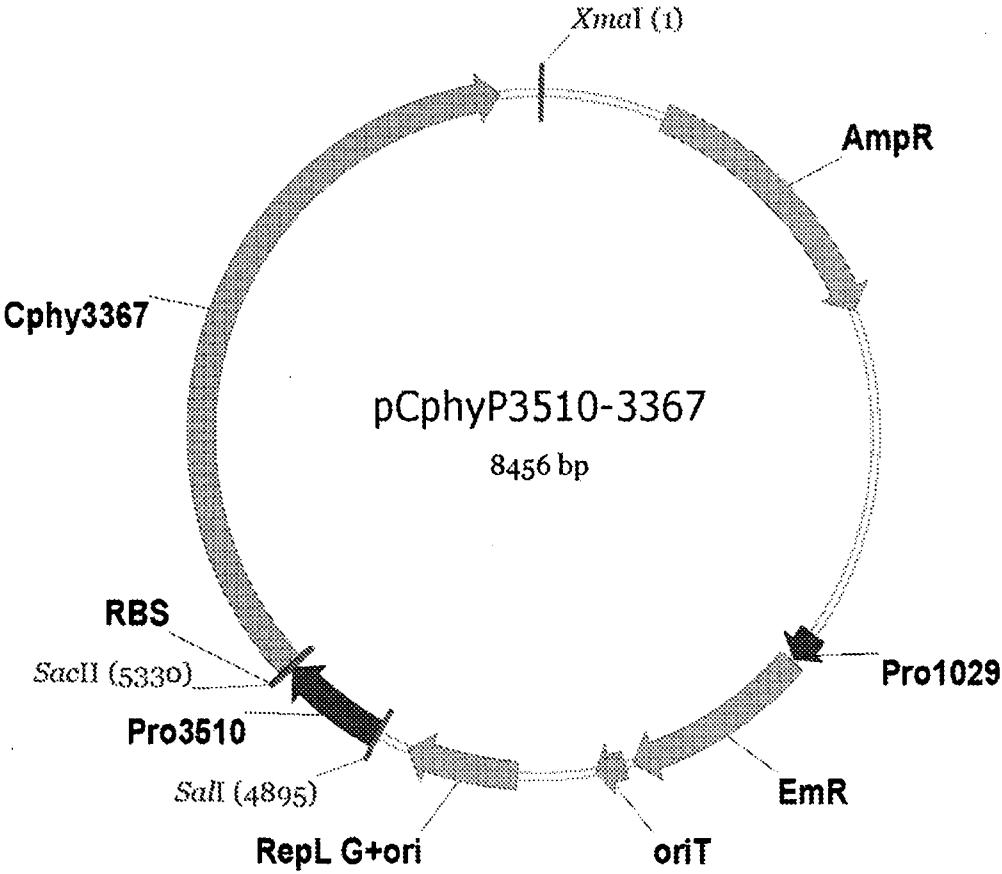


FIG. 19

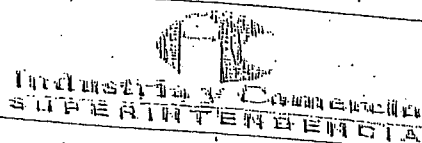
26

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-010649- -00000-0000

Fecha: 2011-01-31 16:12:57 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eje: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 205



LISTA DE CHEQUEO
ADMISIÓN A TRÁMITE - NUEVAS CREACIONES

PATENTE DE INVENCION ☐

MODELO DE UTILIDAD ☐

Art 33 Declación 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 Declación 486/00, Circular Única

1 ☒

Indicación que se solicita una PCT

30 ☒

Copia de la solicitud en español, tal como fue presentada inicialmente (capítulo descriptivo, reivindicatorio, resumen)

168 ☒

Dibujos de ser estos pertinentes

2 ☒

Comprobante de pago de las tasas establecidas (de ser el caso formato de descuento)

Completa ☒

Incompleta ☐

DISEÑO INDUSTRIAL ☐

(Art. 119 Declación 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. 82 Declación 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 ☐



Bogotá D.C.,
2020

Doctor(a)
LLOREDA RICAURTE ALICIA
Apoderado(a) y/o Representante de
QTEROS, INC.

REFERENCIA	Radicación No.	11 010649
	Solicitud internacional	PCT/US2009/051993
	Fecha inicio fase nacional	28/01/2011
	Trámite	11
	Evento	378
	Actuación	430
	Oficio No.	

06089

Como resultado del examen de forma, realizado con fundamento en el artículo 38 de la Decisión 486 de la Comisión de la Comunidad Andina, artículo 27 del Tratado de Cooperación en Materia de Patentes (PCT), regla 51 bis del Reglamento y Circular Única de la Superintendencia de Industria y Comercio, se le comunica que:

1. La solicitud no contiene la copia del documento en que consta la cesión del derecho a la patente del inventor al solicitante o a su causante, conforme a lo establecido por el artículo 26 literal k) de la Decisión 486 de la Comisión de la Comunidad Andina. Cuando se trate de documentos otorgados en el extranjero, estos deberán legalizarse de conformidad con lo dispuesto por los artículos 259 y 260 del Código de Procedimiento Civil. Cuando se trate de las declaraciones a que hace referencia la Regla 4.17 ii) del Tratado de Cooperación en Materia de Patentes PCT., el alcance de las tales declaraciones se determinará según indiquen claramente una cesión de derechos del inventor al solicitante o su causante, de conformidad con lo dispuesto en el numeral 5.2.15 del Capítulo Quinto, Título X de la Circular Única de Superintendencia de Industria y Comercio.

En cumplimiento de lo dispuesto por el artículo 39 de la Decisión 486, se le requiere para que dentro del término de dos meses siguientes a la fecha de notificación del presente oficio, complete los requisitos anteriormente enunciados. En caso de no dar cumplimiento al presente requerimiento, la solicitud se considerará abandonada y perderá su prelación.

Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Única No.10 de 2001.

JOSE LUIS SALAZAR LOPEZ
Director de Nuevas Creaciones (c)

08 ABR. 2011

El anterior requerimiento fue notificado por fijación en lista de fecha _____
adpa11



Bogotá D.C.,
2020



Industria y Comercio
SUPERINTENDENCIA

200

Doctor(a)
LLOREDA RICAURTE ALICIA
Apoderado(a) y/o Representante de
QTEROS, INC.

REFERENCIA	Radicación No.	11 010649
	Solicitud internacional	PCT/US2009/051993
	Fecha inicio fase nacional	28/01/2011
	Trámite	11
	Evento	378
	Actuación	430
	Oficio No.	

06090

Teniendo en cuenta que el capítulo reivindicatorio presentado excéde de diez reivindicaciones, es necesario que cancele la tasa establecida por cada reivindicación adicional presentada, de conformidad con lo dispuesto en la resolución de tasas vigente para la fecha de presentación de la solicitud, so pena de que no sean tenidas en cuenta dentro del presente trámite administrativo.

En consecuencia, con fundamento en lo previsto por el artículo 12 del Código Contencioso Administrativo, se le requiere para que dentro del término de dos (2) meses siguientes a la fecha de notificación del presente oficio atienda el requerimiento aquí señalado. Lo anterior, con el alcance que para tal efecto establece el artículo 13 de la misma codificación, respecto de las reivindicaciones adicionales.

Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Única N° 10 de 2001.

JOSE LUIS SALAZAR LOPEZ
Director de Nuevas Creaciones (c)

08 ABR. 2011

El anterior requerimiento fue notificado por fijación en lista de fecha _____.
adpa11