



- (51) **Clasificación Internacional de Patentes:**
C12N 15/67 (2006.01) **C12N 15/70** (2006.01)
- (21) **Número de la solicitud internacional:**
PCT/CO20 15/0000 10
- (22) **Fecha de presentación internacional:**
14 de mayo de 2015 (14.05.2015)
- (25) **Idioma de presentación:** español
- (26) **Idioma de publicación:** español
- (71) **Solicitantes:** **UNIVERSIDAD DE LOS ANDES** [CO/CO]; Carrera 1A No. 18A-10, Bogotá, D.C. (CO). **CORPORACIÓN CORPOGEN** [CO/CO]; Carrera 5#66 A 34, Bogotá, D.C. (CO).
- (72) **Inventores:** **MONGUI, Alvaro**; Calle 57 #46-34 Blq 7 Apto 613, Bogotá, D.C. (CO). **DEL PORTILLO OBANDO, Patricia**; Carrera 16 #134-81 Apto 203, Bogotá, D.C. (CO). **RESTREPO RESTREPO, Silvia**; Carrera 12 #124-30 Apto 313, Bogotá, D.C. (CO). **HOWARD, Armando Junca**; Carrerra 3 #19-75, Chía, Cundinamarca (CO).
- (74) **Mandatario:** **LLOREDA, Ricaurte, Alicia**; Calle 72 N°5-83, Bogotá, D.C. (CO).
- (81) **Estados designados** (a menos que se indique otra cosa, para toda clase de protección nacional admisible): AE,

AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Estados designados** (a menos que se indique otra cosa, para toda clase de protección regional admisible):
ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), euroasiática (AM, AZ, BY, KG, KZ, RU, TJ, TM), europea (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Publicada:

- con informe de búsqueda internacional (Art. 21(3))
- con reivindicaciones modificadas (Art. 19(1))
- con la parte de lista de secuencias de la descripción (Regla 5.2(a))

(54) **Title:** NOVEL TRANSPOSON THAT PROMOTES THE FUNCTIONAL EXPRESSION OF GENES IN EPISOMAL DNA, AND METHOD FOR INCREASING DNA TRANSCRIPTION IN THE FUNCTIONAL ANALYSIS OF METAGENOMIC LIBRARIES

(54) **Título :** NUEVO TRANSPOSÓN QUE PROMUEVE LA EXPRESIÓN FUNCIONAL DE GENES EN ADNS EPISOMALES Y UN MÉTODO PARA AUMENTAR LA TRANSCRIPCIÓN DE ADN EN ANÁLISIS FUNCIONALES DE LIBRERÍAS METAGENÓMICAS

(57) **Abstract:** A novel transposon (TnC_T7) has been developed to partially supply the transcriptional machinery during the functional analysis of genomic/metagenomic libraries. This transposon was designed and constructed such that it can be integrated randomly into any episomal DNA, allowing the inducible expression of the adjacent DNA regions in both directions. In general, this genetic tool includes a kanamycin resistance gene, two T7 bidirectional promoters and the T7 RNA-polymerase-coding gene, the latter under the control of an arabinose-inducible promoter (PBAD). Experimental validation confirmed the potential of TnC_T7 for use in functional genomic/metagenomic studies in order to partially overcome the limitations of the bacterial hosts, which prevent same from recognising the majority of foreign genes in DNA libraries.

(57) **Resumen:** Se ha desarrollado un nuevo transposón (TnC_T7) para suplir parcialmente la maquinaria transcripcional durante el análisis funcional de librerías genómicas/metagenómicas. Este transposón fue concebido y construido para tener la habilidad de integrarse de manera aleatoria dentro de cualquier ADN episomal, permitiendo la expresión inducible de las regiones de ADN adyacentes en ambas direcciones. En general, esta herramienta genética incluye un gen de resistencia a kanamicina, dos promotores bidireccionales T7 y el gen codificante de la T7RNA polimerasa, éste último bajo la regulación de un promotor inducible con arabinosa (PBAD). La validación experimental confirmó el potencial de TnC_T7 para ser usado en estudios genómicos/metagenómicos funcionales con el fin de superar parcialmente las limitaciones de los hospederos bacterianos, los cuales previenen que éstos reconozcan la mayoría de los genes foráneos de las librerías de ADN.



NUEVO TRANSPOSÓN QUE PROMUEVE LA EXPRESIÓN FUNCIONAL DE GENES EN ADNs EPISOMALES Y UN MÉTODO PARA AUMENTAR LA TRANSCRIPCIÓN DE ADN EN ANÁLISIS FUNCIONALES DE LIBRERÍAS METAGENÓMICAS

Campo de la invención

La presente invención se relaciona con el desarrollo de transposones para suplir parcialmente la maquinaria transcripcional durante el análisis funcional de librerías genómicas/metagenómicas, y por lo tanto, para aumentar la identificación de nuevos compuestos con potencial biotecnológico para librerías de ADN genómicas/metagenómicas, superando parcialmente las limitaciones de los hospederos bacterianos, que previenen que éstos reconozcan la mayoría de los genes foráneos de las librerías de ADN.

Antecedentes de la invención

Actualmente se estima que tan sólo alrededor del uno por ciento de todos los microorganismos presentes en ambientes naturales pueden ser cultivados en condiciones estándar de laboratorio, por lo que todavía es desconocido el gran potencial de compuestos novedosos, actividades enzimáticas y reguladores genéticos útiles para la industria, derivados de la mayor proporción de organismos de la biosfera terrestre.

En vista de lo anterior, es decir, debido a las limitaciones para caracterizar una alta proporción de las enzimas y metabolitos producidos por bacterias no cultivables, la metagenómica ha emergido como una aproximación alternativa al análisis microbiológico convencional. Esta estrategia se basa en la extracción del ADN total a partir de una muestra ambiental (mezcla genómica de la comunidad microbiana conocida como metagenoma) y su clonación subsecuente en una bacteria fácilmente cultivable. Con esta aproximación se han construido librerías genómicas/metagenómicas y se han usado para identificar aislados bacterianos capaces de producir enzimas y metabolitos novedosos. Luego se realizan análisis funcionales basados en la expresión heteróloga del ADN foráneo, lo cual se refleja en un rasgo particular (fenotipo)

expresado por algunos aislados bacterianos de la librería genómica/metagenómica.

Por lo tanto, los análisis funcionales de la mayoría de las librerías de ADN dependen de la expresión eficiente de los genes heterólogos en el hospedero bacteriano para lograr la identificación de funciones derivadas de genes conocidos-desconocidos o de grupos genéticos complejos. El éxito de un ensayo funcional dado dependerá en gran medida del método de detección y de la actividad de interés, pudiéndose distinguir tres tipos generales de análisis: 1) Detección directa de actividad, donde la expresión de una enzima particular o metabolito se usa para identificar los clones bacterianos; 2) Detección modulada, la cual involucra la expresión de genes requeridos para el crecimiento bacteriano bajo condiciones selectivas; y 3) Inducción por sustrato, como una estrategia que promueve la expresión génica en presencia de un sustrato dado.

Sin embargo, independiente del tipo de detección, los análisis funcionales son frecuentemente problemáticos debido al hecho de que la identificación de los fenotipos deseados depende de muchos factores, como son el sistema vector-hospedero seleccionado, el tamaño del gen de interés (individual o del grupo de genes), su abundancia en la fuente metagenómica, el método de detección usado y la eficiencia de la expresión heteróloga de genes en el hospedero seleccionado.

En la mayoría de los estudios genómicos/metagenómicos realizados, *E. coli* ha sido el hospedero preferido para la construcción de librerías y la realización de análisis funcionales debido en gran parte al entendimiento actual de la genética molecular de esta bacteria y a su amplio uso por décadas como modelo en áreas como la microbiología y la biología molecular. Además, el desarrollo e implementación de técnicas sofisticadas de modificación genética, junto con su simple manipulación, rápido crecimiento, facilidad de procesamiento y versatilidad en un amplio rango de herramientas genéticas, han hecho de *E. coli* el hospedero preferido en experimentos de biología molecular.

A pesar de estas ventajas, la maquinaria genética de *E. coli* puede ser incapaz de realizar correctamente la expresión génica de ADNs foráneos. Los

metagenomas son mezclas complejas de genomas de un amplio rango de microorganismos e incluso una librería de un solo genoma podría ser obtenida de un microorganismo distantemente relacionado. Gabor *et al.* cuantificaron la probabilidad teórica de *E. coli* de expresar genes derivados de fragmentos clonados al azar de 32 genomas procariontes completos (*Env. Microbiol.* 2004, 6:879-886). Esto fue determinado *in silico* basado en la presencia de señales funcionales de *E. coli* sobre dichos genomas y la longitud de los insertos genómicos. Usando tres modelos teóricos de expresión de genes, se encontró que en promedio 40% de las actividades enzimáticas codificadas por los genomas procarióticos podían ser accesibles a la maquinaria de la *E. coli*. Esto significa que una porción significativa de genes de estos genomas (60%) todavía sería incompatible para las maquinarias transcripcionales y traduccionales del sistema de expresión bacteriano.

Un estudio más reciente para determinar la habilidad de *E. coli* de transcribir globalmente diferentes genes, tanto procariontes como eucariontes, fue evaluado por microarreglos y RT-PCR (Warren RL, *et al.*; *Genome Res.* 2008, 18:1798-1805). Se observó que la *E. coli* fue capaz de transcribir alrededor de la mitad de los genes de *Haemophilus influenzae*, una menor proporción de genes de *P. aeruginosa* y solo un número mínimo de genes humanos. Adicionalmente, los genes que mostraron niveles de expresión significativamente mayores en *E. coli* tuvieron regiones promotoras relacionadas a los sitios de reconocimiento de la subunidad sigma-70 de la ARN polimerasa bacteriana, resaltando la selectividad de la maquinaria transcripcional del hospedero durante los primeros pasos de la expresión del ADN foráneo.

Por lo tanto, la selectividad de la maquinaria transcripcional es evidente durante las etapas iniciales de la expresión del ADN foráneo, lo que implica que cualquier estrategia encaminada a incrementar la expresión génica en estudios genómicos/metagenómicos funcionales, independiente del hospedero bacteriano seleccionado, tiene que superar esta limitación inicial.

Se han reportado diferentes estrategias para mejorar la expresión heteróloga de genes de ADN genómico/metagenómico. Así, por ejemplo, el uso de

hospederos alternativos, ya sea que la librería de ADN sea construida simultáneamente en algunos de ellos o que sea transferida de un hospedero a otro, ha mostrado ser exitoso en el incremento de la expresión génica. Esta estrategia está usualmente asociada con el desarrollo de vectores de expresión novedosos que puedan ser mantenidos de forma estable en más de un sistema bacteriano.

Un ejemplo de desarrollo (pero no de implementación funcional) de vectores de amplio espectro de hospedero o vectores "shuttle" fue el pSR44 (divulgado por Aakvik T. et al; *FEMS Microbiol Lett* 2009, 296:149-58). Este vector puede ser inducido desde un número bajo de copias hasta un número alto de copias con L-arabinosa y contiene un origen de transferencia RK2 para permitir la conjugación en hospederos adicionales como *Pseudomonas fluorescens* y *Xanthomonas campestris*. Pese a que en años recientes se ha incrementado el repertorio de vectores "shuttle", todavía pocos estudios han confirmado la versatilidad de estas herramientas genéticas para incrementar la expresión heteróloga de genes en análisis funcionales de librerías metagenómicas.

Otra modificación en vectores usados para la construcción de librerías genómicas/metagenómicas es la adición de promotores adyacentes a los sitios de múltiple clonaje con el fin de aumentar la transcripción del ADN foráneo. Por razones obvias, la efectividad de esta estrategia está más restringida a librerías construidas con insertos pequeños de ADN. Este es el caso del plásmido pJOE930, que al tener dos promotores-lac convergentes e inducibles a ambos lados de un sitio simétrico de múltiple clonaje permitió la identificación de un gran número de aislados bacterianos activos, expresando enzimas lipolíticas, amilasas, fosfatasas y dioxigenasas (Lámmle K, et al.; *J. Biotechnol.* 2007, 127:575-92).

La solicitud de patente WO 2012/069668 se relaciona con el desarrollo de vectores y cepas como sistemas de expresión, ofreciendo la posibilidad de identificar genes de interés que no son expresados en bacterias que hospedan la librería metagenómica, y permitiendo de esta forma la detección de las funciones codificadas, las cuales, de otra manera permanecerían silenciadas y no podrían ser detectadas. Específicamente, dicha solicitud de patente revela

la inclusión del promotor derivado del fago T7 en vectores de tipo cósmido y fósido para promover la transcripción de insertos genómicos/metagenómicos como resultado de la expresión de la T7RNA polimerasa (T7RNAP) desde el hospedero. El éxito de esta estrategia está basado en la alta procesividad y eficiencia de la T7RNAP para transcribir genes, pero se restringe al análisis de las regiones flanqueantes de los insertos genómicos/metagenómicos.

Otra aproximación es el uso de elementos de ADN móviles o transposones, los cuales han sido ampliamente usados en una variedad de estudios genéticos avanzados, tal como mutagénesis, secuenciación (US2014/0162897), manipulación genómica, transgénesis, terapia génica y modulación funcional de la expresión de genes (Ivics Z. et al.; *Nat. Methods* 2009, 6:415-22).

Una de las maquinarias de transposición mejor caracterizadas es la del bacteriófago Mu, pues en contraste con la relativa complejidad del mecanismo de transposición *in vivo* de este fago, que involucra un número importante de factores auxiliares, se han observado substancialmente menos condiciones para las reacciones de transposición *in vitro*. Así, los componentes mínimos de reacción para la transposición de tipo Mu incluyen el buffer de reacción, la transposasa MuA purificada, el mini transposón Mu y el ADN de interés (ADN blanco). Estos parámetros han mostrado ser suficientes para eventos de transposición eficientes con bajo sesgo de inserción sobre múltiples ADNs blanco, haciendo de la implementación del transposón Mu una herramienta ideal y ajustable en diferentes campos de investigación.

En términos de aplicaciones en biología molecular el transposón Mu ha facilitado el análisis de secuenciación, la detección de polimorfismos y la determinación precisa de interacciones proteicas. En el campo de la ingeniería de proteínas el transposón Mu ha sido básicamente usado para generar proteínas truncadas con el fin de caracterizar actividades enzimáticas diferenciales. A nivel genómico los transposones Mu han promovido ampliamente eventos de mutagénesis y transgénesis, enfocados a disminuir o aumentar la expresión funcional de genes, respectivamente, en diferentes organismos.

Específicamente, Leggewie C. *et al.* (*J. Biotechnol.* 2006, 123:281-7) revelan la construcción del transposón MuExpress que de forma aleatoria se integra *in vitro* en librerías existentes de cromosomas artificiales bacterianos (BACs, por sus siglas en inglés) o en librerías de cósmidos, favoreciendo la expresión inducible de sus regiones flanqueantes en ambas direcciones y permitiendo la secuenciación bidireccional de los respectivos clones a partir de sitios únicos de unión de los iniciadores.

Dicho transposón MuExpress fue desarrollado como una herramienta genética para abordar la dificultad de la transcripción génica al interior de insertos largos de ADN de librerías metagenómicas. Teóricamente este transposón incrementa el nivel de transcripción de insertos de ADN porque incluye en cada uno de sus extremos una región promotora T7 de lectura hacia el exterior. Sin embargo, un análisis detallado del diseño y construcción original del transposón MuExpress reveló un error importante que hace inviable el reconocimiento de una de las dos regiones promotoras T7 por la T7RNAP.

Otro transposón conocido por insertar de forma aleatoria un único promotor T7, pero derivado del sistema de transposición Tn5, es el EZ-Tn5 <T7/KAN-2> (Epicentre-Illumina, Madison, WI, USA). Sin embargo, ni el MuExpress ni el EZ-Tn5 <T7/KAN-2>, los cuales están basados en la alta procesividad de transcripción de la T7RNAP, reportaron suficiente evidencia de inserción de los transposones ni su relación con la mejora en la expresión génica. Adicionalmente, los transposones MuExpress y EZ-Tn5 <T7/KAN-2> dependen de hospederos bacterianos que expresen la T7RNAP (por ejemplo *E. coli* BL21 DE3, Invitrogen), lo cual restringe ampliamente su uso en ensayos funcionales, especialmente con ADN metagenómico, ya que la mayoría de los kits de construcción de librerías dependen de otras cepas bacterianas especializadas (por ejemplo *E. coli* Epi300, Epicentre-Illumina, Madison, WI, USA).

Con base en la versatilidad de los transposones Mu en la investigación en biología-biotecnología molecular, así como en la necesidad actual de mejorar eficientemente la expresión heteróloga de genes de librerías de ADN genómico/metagenómico, se hacía necesario crear nuevas herramientas y estrategias genéticas usando éste elementó móvil de ADN.

Resumen de la invención

La presente invención involucra el diseño de un nuevo transposón Mu y los métodos para lograr una expresión eficiente de genes alojados en ADNs episomales de librerías genómicas/metagenómicas, que bajo aproximaciones tradicionales de análisis no son detectados en ensayos funcionales. La eficiencia en el uso de la invención se refleja en una proporción incrementada de aislados bacterianos que muestran el fenotipo deseado, comparado con la proporción de aislados bacterianos que pueden ser identificados en los análisis funcionales originales. El primer aspecto de la invención se basa en el desarrollo secuencial de plásmidos para la construcción del nuevo transposón Mu.

En una realización, la invención se dirige al desarrollo de un gen sintético (Tn_A), el cual es una secuencia de ADN artificial resultado de la combinación específica de ciertos elementos de ADN, que comprende:

- (i) una secuencia promotora T7,
- (ii) un sitio de reconocimiento de repetición invertida para la transposasa MuA,
- (iii) múltiples sitios flanqueantes de reconocimiento para endonucleasas de restricción.

En otra realización, la invención se dirige al desarrollo del plásmido pUC57_Tn, el cual es un vector artificial resultado de la combinación específica de ciertos elementos de ADN, que comprende:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) una secuencia promotora T7 en una orientación específica,
- (iii) un sitio de reconocimiento de repetición invertida para la transposasa MuA en una orientación específica.

En otra realización, la invención se dirige al desarrollo del plásmido pUC57_Tn_kanAB, el cual es un vector artificial resultado de la combinación específica de ciertos elementos de ADN, que comprende:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) una secuencia promotora T7 de lectura hacia el exterior,
- (iii) un sitio de reconocimiento de repetición invertida para la transposasa MuA,
- (iv) un marcador de selección diferente de aquel que está localizado en el vector esqueleto.

En otra realización, la invención se dirige al desarrollo del plásmido pBAD18-Cm_t7nap, el cual es un vector artificial resultado de la combinación específica de ciertos elementos de ADN, que comprende:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) un gen que codifica la T7RNA polimerasa regulado por un promotor inducible.

En otra realización, la invención se dirige al desarrollo del plásmido pUC57_Tn_kanAB_t7, el cual es un vector artificial resultado de la combinación específica de ciertos elementos de ADN, que comprende:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) una secuencia promotora T7 de lectura hacia el exterior,
- (iii) un sitio de reconocimiento de repetición invertida para la transposasa MuA,
- (iv) un marcador de selección diferente de aquel que está localizado en el vector esqueleto,

- (v) un gen que codifica la T7RNA polimerasa regulado por un promotor inducible.

En otra realización, la invención se dirige al desarrollo del plásmido pUC57_TnC_T7 el cual es un vector artificial resultado de la combinación específica de ciertos elementos de ADN, que comprende:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) un marcador de selección diferente de aquel localizado en el vector esqueleto,
- (iii) un gen que codifica la T7RNA polimerasa bajo un promotor inducible,
- (iv) dos promotores flanqueantes T7 de lectura hacia el exterior,
- (v) dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA.

En otra realización, la invención se dirige al desarrollo del fósido F076_GFP el cual es un vector artificial resultado de la combinación específica de ciertos elementos de ADN, que comprende:

- (i) un fósido esqueleto con orígenes de replicación de bajo a alto número de copias y un marcador de selección,
- (ii) un inserto de ADN metagenómico,
- (iii) un gen que codifica una variante de la proteína verde fluorescente (GFP) con un sitio de unión al ribosoma (RBS) corriente arriba, ambos localizados específicamente dentro del inserto de ADN metagenómico.

En otra realización, la presente invención está dirigida de forma amplia a un método para aumentar la transcripción de ADN, como paso inicial para la expresión foránea de genes, que comprende:

- (i) generar librerías de transposición de ADN episomal, como resultado de la inserción aleatoria del transposón TnC_T7 purificado, donde dicho ADN episomal incluye plásmidos, fósidos, cósmidos o BACs,

(ii) introducir las librerías de transposición de ADN episomal de (i) dentro de células hospederas,

(iii) expresar la T7RNA polimerasa codificada en el transposón TnC_T7, para proporcionarle a las poblaciones de células hospederas bacterianas una colección diversa de transcritos de ARN derivados del ADN episomal,

(iv) analizar dicha población de células hospederas bacterianas para identificar los aislados bacterianos que expresan un gen reportero o cualquier otra función deseada.

En otra realización, la presente invención incluye ocho plásmidos, los cuales corresponden a vectores artificiales resultado de inserciones aleatorias del transposón TnC_T7 en el plásmido pKR-C12, cada uno comprendiendo:

- (i) un vector esqueleto con su propio marcador de selección y un gene reportero silenciado,
- (ii) una inserción diferencial del transposón TnC_T7 a lo largo de la secuencia de ADN de (i).

Breve descripción de las figuras

Las siguientes figuras forman parte de la presente descripción y se incluyen para demostrar adicionalmente ciertos aspectos de ésta. La invención puede entenderse mejor a través de la referencia a una o más de éstas figuras, en combinación con la descripción detallada de las realizaciones específicas aquí presentadas.

La FIG. 1 muestra la estructura de la secuencia de ADN Tn_A, la cual comprende la secuencia R1-R2 de unión para la transposasa MuA, correspondiente a uno de los sitios de reconocimiento de repetición Invertida, y a la región promotora T7 en la misma molécula de ADN, pero en la hebra opuesta. La longitud del gen Tn_A es de 138 pb e incluye sitios flanqueantes de reconocimiento para las siguientes endonucleasas de restricción: EcoRI, BglII, AscI y 8amHI.

La FIG. 2 esquematiza la clonación de la secuencia de ADN Tn_A en el vector pUC57 para generar el plásmido pUC57_Tn. La clonación directa fue lograda usando los sitios de restricción EcoRI y *Bam*HI en el gen y en el plásmido pUC57, como un requerimiento para los pasos subsiguientes con el fin de construir el plásmido que alberga el transposón TnC_T7. AmpR, gen de resistencia a ampicilina; M13 fwd, sitio de hibridación para el iniciador M13 directo; M13 rev, sitio de hibridación para el iniciador M13 reverso; ori, origen de replicación ColE1/pMB1/pBR322/pUC de alto número de copias; sitio de unión CAP, sitio de unión de proteína de activador de catabolito.

La FIG. 3 describe la clonación en pUC57_Tn del gen de resistencia a la kanamicina (*Kan*) incluyendo su promotor, para generar el plásmido pUC57_Tn_kanAB. Primero se amplificó *Kan* en dos pasos, con el fin de reemplazar el sitio de restricción *Bgl*II por *Spe*I. El producto ligado se insertó en los sitios de restricción *Asc*I y *Bam*HI de pUC57_Tn. AmpR, gen de resistencia a ampicilina; M13 fwd, sitio de hibridación para el iniciador M13 directo; M13 rev, sitio de hibridación para el iniciador M13 reverso; ori, origen de replicación ColE1/pMB1/pBR322/pUC de alto número de copias; sitio de unión CAP, sitio de unión de proteína de activador de catabolito; NeoR/KanR, gen de resistencia a neomicina y kanamicina; FRT, sitio de escisión pero no integración mediado por FLP.

La FIG. 4 muestra la clonación de la secuencia que codifica la T7RNA polimerasa (T7RNAP) en el sitio de restricción único *Kpn*I (posterior a la reparación de los extremos) del plásmido pBAD18-Cm (Guzmán LM, et al.; *J. Bacteriol.* 1995, 177:4121-30), para generar el plásmido pBAD18-Cm_t7na. La secuencia codificante de la T7RNAP se localiza corriente abajo del promotor inducible con arabinosa (PBAD) y corriente arriba de los terminadores transcripcionales *rrnB* T1 y T2. CmR, gen de resistencia a cloranfenicol; f1 ori, origen de replicación del bacteriófago f1; ori, origen de replicación ColE1/pMB1/pBR322/pUC de alto número de copias; bom, base de la región de movilidad para pBR322; araC, proteína regulatoria de L-arabinosa; promotor araBAD, promotor del operón de L-arabinosa de *E. coli*.

La FIG. 5 representa la detección del péptido AA en un ensayo de western blot, como resultado de la expresión de la T7RNAP derivada de diferentes extractos de cultivos bacterianos. 1, control positivo de la expresión del péptido AA a partir del clon bacteriano *E. coli* BL21 DE3 transformado con el plásmido pET28a_AA y suplementado con kanamicina e IPTG; 2, clon bacteriano *E. coli* TOP10 transformado con los plásmidos pET28a_AA y pBAD18-Cm_t7rnep y suplementado con kanamicina, cloranfenicol y D-glucosa; 3-6, clones bacterianos *E. coli* TOP10 transformados con los plásmidos pET28a_AA y pBAD18-Cm_t7rnep y suplementados con kanamicina, cloranfenicol y L-arabinosa.

La FIG. 6 esquematiza la clonación del promotor inducible con arabinosa y de la secuencia codificante de la T7RNAP, como un solo amplicón, en el sitio único de restricción >Ascl de pUC57_Tn_kanAB, generando el plásmido pUC57_Tn_kanAB_t7. El amplicón P_BAD_T7RNAP se localiza entre las secuencias de ADN del Tn_A y del gen de resistencia a kanamicina. AmpR, gen de resistencia a ampicilina; M13 fwd, sitio de hibridación para el iniciador M13 directo; M13 rev, sitio de hibridación para el iniciador M13 reverso; ori, origen de replicación ColE1/pMB1/pBR322/pUC de alto número de copias; sitio de unión CAP, sitio de unión de proteína de activador de catabolito; NeoR/KanR, gen de resistencia a neomicina y kanamicina; FRT, sitio de escisión pero no integración mediado por FLP; T7RNAP, secuencia codificante de la T7RNA polimerasa; promotor araBAD, promotor del operón de L-arabinosa de *E. coli*.

La FIG. 7 esquematiza la clonación del segundo extremo del transposón (Tn_B) en pUC57_Tn_kanAB_t7 para generar el plásmido pUC57_TnC_T7. Tn_B incluyó el otro sitio R1-R2 de unión para la transposasa MuA, la segunda región promotora T7 de la construcción final del transposón TnC_T7 y dos sitios flanqueantes de restricción HindIII, que permiten su clonación en el vector blanco. Se resaltan además los sitios de restricción BglII que son necesarios para liberar el transposón TnC_T7 del plásmido. AmpR, gen de resistencia a ampicilina; M13 fwd, sitio de hibridación para el iniciador M13 directo; M13 rev, sitio de hibridación para el iniciador M13 reverso; ori, origen de replicación ColE1/pMB1/pBR322/pUC de alto número de copias; sitio de unión CAP, sitio

de unión de proteína de activador de catabolito; NeoR/KanR, gen de resistencia a neomicina y kanamicina; FRT, sitio de escisión pero no integración mediado por FLP; T7RNAP, secuencia codificante de la T7RNA polimerasa; promotor araBAD, promotor del operón de L-arabinosa de *E. coli*.

La FIG. 8 representa a escala las regiones estructurales del transposón TnC_T7. Cada uno de los dos sitios de unión para la transposasa MuA está adyacente a un promotor T7 individual. En cursiva se resaltan los sitios de restricción más representativos en la construcción del transposón. La distancia entre los dos sitios de restricción *Bgl*/U es de 4,575 pb. KanR, gen de resistencia a kanamicina; promotor araBAD (promotor inducible por arabinosa (PBAD)), promotor del operón de L-arabinosa de *E. coli*; T7RNAP, secuencia codificante de la T7RNA polimerasa.

La FIG. 9 representa el principio del método aquí revelado, dirigido a aumentar la transcripción de ADN como el paso inicial de la expresión de genes foráneos. Específicamente, el método comprende el uso del transposón TnC_T7, un transposón Mu, para suplir parcialmente la maquinaria transcripcional durante los análisis funcionales de librerías genómicas/metagenómicas. Este transposón fue concebido y construido para tener la habilidad de integrarse de forma aleatoria dentro de cualquier ADN episomal, permitiendo la expresión inducible de las regiones adyacentes de ADN en ambas direcciones. A, B y C muestran ejemplos de eventos de transposición del TnC_T7 sobre ADN blanco y la habilidad en cada caso de aumentar la expresión génica como resultado de su inserción específica. En caso de que haya un gen de interés (cajas negras con flechas blancas mostrando la orientación de lectura) en un ADN blanco particular, la inserción aleatoria del transposón TnC_T7 puede promover la transcripción del ADN (flechas punteadas), lo cual conducirá eventualmente a la expresión de proteínas particulares (círculos negros) y a la detección del fenotipo deseado.

La FIG. 10 representa la detección inicial de clones bacterianos que expresan GFP, con base en la inserción aleatoria del transposón TnC_T7 en pKR-C12 (1-15). La detección de la fluorescencia fue realizada por espectrofotometría y expresada en términos de Unidades Relativas de Fluorescencia (RFUs). Como

controles negativo (-) y positivo (+) de fluorescencia se incubó *E. coli* Epi300 pKR-C12 en ausencia y presencia de *N*-(3-oxododecanoil)-L-homoserina lactona (3-oxo-C12-HSL) 5 μ M (Sigma-Aldrich, Saint Louis, MO, USA), respectivamente.

La FIG. 11 esquematiza las inserciones del transposón TnC_T7 en pKR-C12 y la detección de fluorescencia de aislados bacterianos seleccionados post-transposición. A, diagrama a escala del plásmido pKR-C12. El fragmento del plásmido resaltado muestra la localización exacta de las inserciones del transposón (asociadas a los clones bacterianos 1-8) localizadas en la cercanía del gen codificante de GFPmut3* (GFP). Las flechas punteadas orientadas de arriba hacia abajo representan los sitios de inserción del transposón en los cuales el gen *t7rnap* está localizado en la hebra sentido del ADN blanco, mientras que las flechas de abajo hacia arriba representan el mismo gen localizado en la hebra antisentido del ADN. pBBR1, origen de replicación del plásmido; LasR, regulador transcripcional; P-O, promotor Lac y sistema operador Lac; GFP, proteína verde fluorescente; CmR, gen de resistencia a cloranfenicol; GmR gen de resistencia a gentamicina; CAP, sitio de unión de proteína de activador de catabolito; *, dos inserciones independientes de TnC_T7 en la misma posición. B, detección de fluorescencia por espectrofotometría expresada en términos de RFUs/Densidad Óptica 600nm (GFP/OD); 1-8, aislados de *E. coli* Epi300 alojando los plásmidos con las inserciones del transposón mostradas en A; (-) control negativo de la expresión de GFP; (+) control positivo de la expresión de GFP. C, detección de fluorescencia por el detector de Sistema de Imágenes IVIS; 1-8, aislados de *E. coli* Epi300 alojando los plásmidos con las inserciones del transposón mostradas en A.

La FIG. 12 representa un fósido derivado de pCC2FOS, el cual incluye un inserto de ADN metagenómico de suelo (F076) y en donde la secuencia codificante de GFPmut3* fue clonada, para generar el fósido F076_GFP. El sitio de restricción *Ascl* se usó para la clonación de la secuencia codificante de GFPmut3* dentro del inserto metagenómico del fósido F076. M13 fwd, sitio de hibridación para el iniciador M13 directo; M13 rev, sitio de hibridación para el iniciador M13 reverso; sitio de unión CAP, sitio de unión de proteína de

activador de catabolito; CmR, gen de resistencia a cloranfenicol; oriV, origen de replicación para el plásmido F bacteriano; ori2, origen secundario de replicación para el plásmido F bacteriano; repE, proteína de iniciación de replicación para el plásmido bacteriano F; incC, región de incompatibilidad del plásmido F bacteriano; sopA y B, proteínas de partición del plásmido F bacteriano; sopC, región de partición similar al centromero del plásmido F bacteriano; loxP, sitio de recombinación mediado por Cre.

La FIG. 13 esquematiza la detección de fluorescencia de aislados bacterianos de *E. coli* transformados con F076_GFP post transposición con TnC_T7 y la localización de las inserciones del transposón a lo largo de la secuencia del fósido. A, detección de fluorescencia usando el Sistema de Imágenes IVIS para aislados de *E. coli* Epi300 transformados con la reacción de transposición de TnC_T7 en F076_GFPmut3*. Se incluyeron los controles negativo (-) y positivo (+) de la expresión de GFP, correspondientes a los aislados bacterianos 2 y 4 de *E. coli* Epi300, respectivamente, de la FIG. 11C. La fluorescencia basal de fondo se corrigió con base en la señal de auto fluorescencia del control negativo usado. B, revalidación de la expresión de GFP para aislados bacterianos que alojan el fósido F076_GFPmut3*, usando el Sistema de Imágenes IVIS después de la recuperación de los aislados en agar LB con cloranfenicol 20 pg/mL, kanamicina 40 µg/mL y L-arabinosa 0.2%. C, diagrama a escala de F076_GFPmut3* con los sitios identificados de inserción del transposón TnC_T7. GFP, secuencia codificante de la proteína verde fluorescente. Las cajas y las flechas blancas representan los genes y secuencias regulatorias del esqueleto fosmídico pCC2FOS. Las flechas pequeñas segmentadas orientadas de arriba hacia abajo representan los sitios de inserción del transposón en el cual el gen *t7rnap* está localizado en la hebra sentido del ADN blanco, mientras que las flechas de abajo hacia arriba representan el mismo gen localizado en la hebra antisentido del ADN. Se resalta la secuencia del inserto de ADN metagenómico (MI), así como los ORFs con longitud mayor a 150 codones localizados en la hebra sentido del ADN fosmídico. El recuadrado sombreado incluye las inserciones de TnC_T7 que promovieron la expresión GFP. CmR, gen de resistencia a cloranfenicol.

La FIG. 14 esquematiza la degradación del 4-nitrofenil butirato a partir de los extractos de clones bacterianos. BL21, clon bacteriano *E. coli* BL21 DE3 usado como control negativo; LipN, clon bacteriano *E. coli* BL21 transformado con el plásmido pET100_LipN (clon bacteriano amablemente proporcionado por Luis Peña - Biotecnología Molecular, CorpoGen, Bogotá, Colombia) usado como un control positivo; 14gF2, *E. coli* Epi300 transformado con el fósido pCC2FOS_14gF2; 1-15, aislados bacterianos de *E. coli* Epi300 transformados con el fósido pCC2FOS_14gF2 post transposición con TnC_T7.

Descripción detallada de la invención

Definiciones

A menos de que se defina o describa específicamente lo contrario en alguna parte de este texto, los siguientes términos y descripciones relacionados con la invención deben ser entendidas como se describen a continuación.

Como se usa en la presente memoria descriptiva, "ADN artificial" significa una secuencia de ADN diferente de cualquiera encontrada en la naturaleza o producida por un proceso no natural, como resultado de técnicas *in vitro* o síntesis de ADN en fase sólida.

Como se usa en la presente memoria descriptiva, "gen sintético" significa un fragmento de ADN sintetizado en el laboratorio por la combinación de nucleótidos, sin secuencias de ADN pre-existentes. En particular, el término se refiere a una molécula de ADN de doble hebra completamente sintética.

Como se usa en la presente memoria descriptiva, "sitios de reconocimiento" se refiere a localizaciones sobre una molécula de ADN que contienen secuencias de nucleótidos específicas, las cuales son reconocidas por proteínas o enzimas específicas.

Como se usa en la presente memoria descriptiva, "endonucleasas de restricción" significa enzimas que cortan moléculas de ADN de doble hélice en sitios de reconocimiento específicos. El término referido en la presente invención se relaciona con endonucleasas de restricción o enzimas que reconocen específicamente secuencias de ADN de 6-8 nucleótidos, en las

cuales la secuencia de nucleótidos de una de las hebras del ADN se lee en orden inverso a aquella de la hebra de ADN complementario (palindrómicas).

Como se usa en la presente memoria descriptiva, "elemento transponible" o "transposón" se refiere a una secuencia de ADN o a un segmento génico capaz de moverse de un genoma o posición genética para ser insertado en otra posición (por ejemplo otro genoma, cromosoma, ADN episomal). La definición mencionada incluye sólo elementos transponibles o transposones que están basados en moléculas de ADN intermediarias y que requieren de la actividad enzimática de proteínas particulares, denominadas como transposasas, para moverse a lo largo de diferentes posiciones genéticas.

Como se usa en la presente memoria descriptiva, "transposición" o "reacción de transposición" se refiere a la reacción en la cual el transposón es insertado dentro de un ADN blanco en sitios aleatorios, a través de la actividad catalítica de una transposasa.

Como se usa en la presente memoria descriptiva, "repetición invertida" significa una secuencia identificada en los extremos 5' o 3' de los transposones que son reconocidos específicamente por transposasas.

Como se usa en la presente memoria descriptiva, "transposasa" pretende significar una enzima que tiene la capacidad de reconocer y unirse a un extremo de un transposón o a las secuencias finales de un transposón en una reacción de transposición, para promover la movilización del transposón.

Como se usa en la presente memoria descriptiva, "inserciones de transposón" significa las localizaciones específicas donde el transposón es insertado dentro de un ADN blanco, como resultado de la reacción de transposición realizada por una transposasa específica.

Como se usa en la presente memoria descriptiva, "ADN blanco" o "vector blanco" significa un ADN de doble cadena que es adecuado para ser modificado usando técnicas de biología molecular. En esta invención, la definición está asociada con secuencias de ADN episomal que incluyen sitios específicos de reconocimiento para endonucleasas de restricción o que pueden

ser modificadas por transposasas como resultado de la inclusión de elementos de ADN transponibles.

Como se usa en la presente memoria descriptiva, "transcripción de ADN" significa el proceso de sintetizar una copia de ARN a partir de una molécula de ADN. Éste corresponde al primer paso de expresión génica y es llevado a cabo por una enzima especializada, una ARN polimerasa.

Como se usa en la presente memoria descriptiva, "promotor" se refiere a una región en una secuencia de ADN sobre la cual una ARN polimerasa específica puede unirse (por ejemplo el promotor T7 es reconocido únicamente por la T7RNAP), con el fin de empezar el proceso de transcripción de ADN.

Como se usa en la presente memoria descriptiva, "promotor inducible" significa que el reconocimiento del promotor por la ARN polimerasa y por lo tanto la actividad transcripcional puede ser controlada por la ausencia o presencia de factores químicos o físicos. Para propósitos de la presente patente, si un promotor es inducido por un factor específico esto llevará a la síntesis de una proteína específica (por ejemplo la T7RNAP).

Como se usa en la presente memoria descriptiva, "gen constitutivo" o "gen constitutivamente expresado" significa un gen que se transcribe continuamente a un nivel relativamente constante. Este término implica que un promotor constitutivo regula la transcripción de ADN para el gen (por ejemplo el gen de resistencia a kanamicina) y por lo tanto la expresión constante de la proteína resultante.

Como se usa en la presente memoria descriptiva, "lectura hacia al exterior" se refiere a la dirección de la transcripción del ADN desde un promotor específico, el cual está localizado particularmente dentro de una secuencia de ADN definida, (como puede ser un transposón) y que se localiza en los extremos 5' o 3' del segmento de ADN. En este caso, la lectura hacia el exterior se restringe al proceso de transcripción de ADN desde el promotor mencionado, en el cual la síntesis de ARN comienza desde el transposón, pero se extiende principalmente hacia el ADN adyacente al mismo.

Como se usa en la presente memoria descriptiva, "vector" se refiere a una molécula circular de doble cadena de ADN usada como un vehículo para transportar artificialmente ADN foráneo al interior de una célula bacteriana blanco.

Como se usa en la presente memoria descriptiva, "vector artificial" se refiere a cualquier ADN artificial como un vector, que es capaz de auto replicarse dentro de una célula bacteriana y por lo tanto ser mantenido de forma estable dentro de la célula bacteriana hospedera.

Como se usa en la presente memoria descriptiva, "auto replicación" y "episomal" se refieren a la capacidad de un vector o un vector artificial de no estar integrados dentro del ADN genómico de un cierto hospedero celular, sino de replicarse automáticamente en una célula hospedera y, por lo tanto, estar presente cuando la célula hospedera crece y se divide. En particular, este término asume la permanencia del vector o el vector artificial por varias generaciones de crecimiento dentro de la célula hospedera. Este término incluye plásmidos, fósidos, cósmidos y cromosomas bacterianos artificiales (BACs).

Como se usa en la presente memoria descriptiva, "origen de replicación" se refiere a secuencias particulares en ADNs episomales en los cuales se inicia la replicación, con base en el reclutamiento de proteínas involucradas en la replicación de ADN.

Como se usa en la presente memoria descriptiva, "transformación" significa el proceso de introducción de material genético nuevo específicamente en células bacterianas. En la presente invención el término mencionado está asociado a la introducción de vectores, vectores artificiales o vectores artificiales modificados dentro de células bacterianas.

Como se usa en la presente memoria descriptiva, "marcador de selección" se refiere a un gen localizado dentro de la bacteria (a nivel genómico o episomal) que confiere una característica para selección artificial. Este término está asociado a genes de resistencia a antibióticos (por ejemplo el gen de

resistencia a cloranfenicol) localizados en vectores o vectores artificiales para la selección de aislados de bacteria después de la transformación.

Como se usa en la presente memoria descriptiva, los términos "corriente arriba" y "corriente abajo" son empleados para diferenciar posiciones relativas en secuencias de ADN o ARN. La posición corriente arriba es una posición hacia el extremo 5' de otro segmento de ácido nucleico (por ejemplo un promotor, gen, sitio de restricción, etc.) en una hebra sencilla de ADN o en una molécula de ARN. La posición corriente abajo es una posición hacia el 3' de otro segmento de ácido nucleico en una hebra sencilla de ADN o en una molécula de ARN.

Como se usa en la presente memoria descriptiva, "ADN metagenómico" se refiere a la totalidad del ADN genómico asociado a microbios, aislado de muestras complejas como ambientes naturales abiertos (por ejemplo suelo, agua) o a partir de microbiomas de organismos multicelulares (por ejemplo humanos).

Como se usa en la presente memoria descriptiva, "inserto" o "inserto de ADN" significa una parte o fragmento o secuencia de ADN que es insertado por técnicas de biología molecular dentro de un vector o un vector artificial para su subsiguiente selección, manipulación o expresión dentro de un organismo hospedero.

Como se usa en la presente memoria descriptiva, "sitio de unión a ribosoma" se refiere a una secuencia de ARN en la cual los ribosomas pueden unirse para iniciar el proceso de síntesis de proteínas (traducción) dentro de la célula hospedera u organismo, como parte del proceso de expresión de proteínas.

Como se usa en la presente memoria descriptiva, "expresión de gen foráneo" significa la totalidad del proceso por el cual la información de un gen particular es usada para sintetizar un producto, que para los propósitos de la presente invención significa sintetizar una proteína. En el mencionado término, "foráneo" significa que el gen evaluado pertenece a un organismo diferente de aquel usado para promover la expresión del gen.

Como se usa en la presente memoria descriptiva, "gen reportero" significa un gen cuya expresión en un hospedero bacteriano puede ser fácilmente monitoreada o detectada. En el contexto de la presente invención, el gen reportero codifica una variante de la proteína verde fluorescente (GFP).

Como se usa en la presente memoria descriptiva, "gen silenciado" se refiere a un gen que es incapaz de expresar la proteína asociada a su secuencia codificante, ya sea por impedimentos durante el proceso de transcripción o de traducción dentro de la célula hospedera.

Una realización de la invención aquí revelada se dirige al diseño y desarrollo de un gen sintético (Tn_A), el cual es una secuencia artificial de ADN de doble hebra de 138 pares de bases (pb) resultado de la combinación específica de ciertos elementos de ADN. Este fragmento de ADN incluye un sitio de unión MuA correspondiente al sitio de reconocimiento de repetición invertida para la transposasa, una secuencia promotora T7 que permite la interacción específica de la T7RNAP y los siguientes sitios flanqueantes de reconocimiento para endonucleasas de restricción: EcoRI, *Bgl*W, *Asc*I y *Sam*HI. El diseño del gen sintético mencionado fue pensado para localizar el sitio de unión MuA y el promotor T7 en hebras de ADN diferentes (FIG. 1), con el fin de alcanzar la actividad deseada en la construcción final del transposón.

En una realización, la presente invención incluye el plásmido pUC57_Tn, que corresponde un vector artificial de 2,795 pb que puede ser mantenido dentro de una célula hospedera bacteriana y donde dicho plásmido es el resultado de una combinación específica de ciertos elementos de ADN (FIG. 2). pUC57_Tn tiene un vector esqueleto con el origen de replicación ColE1/pMB1/pBR322/pUC y un gen de resistencia a ampicilina como marcador de selección después de su transformación en la célula hospedera. pUC57_Tn es el resultado de clonar el gen sintético Tn_A en los sitios únicos EcoRI y *Sam*HI, después de un tratamiento con enzimas de restricción del vector esqueleto y del gen sintético. Por lo tanto, el vector pUC57_Tn en la presente realización incluye un promotor T7 de lectura hacia el exterior y un sitio de reconocimiento de repetición invertida para la transposasa MuA, ambos localizados de manera específica.

Una realización de la invención incluye el plásmido pUC57_Tn_kanAB, el cual es un vector artificial de una longitud total de 3,983 pb que puede ser replicado de manera episomal en una célula hospedera bacteriana. El plásmido pUC57_Tn_kanAB resulta de la combinación de ciertos elementos de ADN (FIG. 3), ya que éste incluye las mencionadas regiones ADN del pUC57_Tn más un gen de resistencia adicional. En la presente invención, el plásmido pUC57_Tn_kanAB es el resultado de la clonación del gen de resistencia a kanamicina, incluyendo su promotor y su señal de terminación de la transcripción en los sitios de restricción *AscI* y *BamHI* de pUC57_Tn. Para la presente realización, la clonación del gen de resistencia a kanamicina fue llevada a cabo de una manera específica con el fin de asegurar una adecuada adición de las regiones de ADN subsiguientes.

En una realización, la presente invención es orientada al desarrollo del plásmido pBAD18-Cm_t7nap (FIG. 4), el cual es un vector artificial que puede ser mantenido de forma estable dentro de una célula hospedera bacteriana. Este vector tiene un tamaño total de 8,738 pb y es el resultado de la combinación de los siguientes elementos de ADN: origen de replicación ColE1/pMB1/pBR322/pUC, un gen de resistencia a cloranfenicol y la secuencia codificante de la T7RNAP clonada en el sitio único de restricción *KpnI*. Adicionalmente, el pBAD18-Cm_t7nap descrito en esta invención tiene la secuencia codificante de la T7RNAP localizada corriente abajo del promotor inducible con arabinosa y corriente arriba de los terminadores transcripcionales rrnB T1 y T2.

Una realización de la invención se relaciona con el plásmido pUC57_Tn_kanAB_t7, el cual es un vector artificial de 7,097 pb (FIG. 6), resultado de la combinación específica y orientación de ciertos elementos de ADN y que puede mantenerse dentro de una célula hospedera bacteriana. El plásmido pUC57_Tn_kanAB_t7 incluye los elementos de ADN mencionados del pUC57_Tn_kanAB, los cuales son un origen de replicación ColE1/pMB1/pBR322/pUC, un gen de resistencia a ampicilina, el gen sintético Tn_A y el gen de resistencia a kanamicina, más el promotor inducible con arabinosa y la secuencia codificante de la T7RNAP del vector pBAD18-Cm_t7nap. El vector artificial pUC57_Tn_kanAB_t7 resulta específicamente de

la clonación del promotor inducible con arabinosa y la secuencia codificante de la T7RNAP en el sitio único de restricción *AscI* de pUC57_Tn_kanAB, entre las secuencias de ADN de Tn_A y de resistencia a kanamicina en el vector artificial.

Una realización de la invención aquí revelada está dirigida al desarrollo del plásmido pUC57_TnC_T7 (FIG. 7), el cual es un vector artificial resultado de la combinación específica de ciertos elementos de ADN y que puede mantenerse en una célula hospedera bacteriana. Este vector artificial incluye todos los elementos estructurales del vector pUC57_Tn_kanAB_t7 más el segundo extremo del transposón, denotado aquí como Tn_B. El plásmido pUC57_TnC_T7 tiene una longitud total de 7,240 pb y tiene clonado específicamente la región Tn_B en el sitio único de restricción *HindIII* de pUC57_Tn_kanAB_t7. Como consecuencia, el vector artificial pUC57_TnC_T7 tiene dos promotores T7 flanqueantes de lectura hacia el exterior, así como dos sitios de reconocimiento de repeticiones invertidas para la transposasa MuA. Por lo tanto, pUC57_TnC_T7 tiene clonado la secuencia completa del transposón TnC_T7, la cual a su vez puede ser liberada por medio del tratamiento con la enzima de restricción *BglII*.

En una realización la presente invención se relaciona con el desarrollo del fósido F076_GFP (FIG. 11), el cual es un vector artificial que comprende 45,619 pb. Este vector resulta de combinar de una manera específica un esqueleto fosmídico, un inserto de ADN metagenómico y la secuencia codificante de un gen reportero. El esqueleto fosmídico corresponde al vector comercial pCC2FOS (Epicentre-Illumina, Madison, WI, USA). El inserto metagenómico resulta de la clonación aleatoria de ADN metagenómico en el pCC2FOS. El gen reportero corresponde a la secuencia codificante de GFP incluyendo un sitio de unión a ribosoma (RBS), ambos fragmentos de ADN siendo clonados como un solo amplicón en un sitio único de restricción *AscI* localizado en el inserto metagenómico.

Una realización de la invención revelada es un método para aumentar la transcripción de ADN, incluyendo, pero no limitado a la expresión de genes foráneos (FIG. 9), el cual comprende:

- (i) Generar librerías de ADN basadas en la transposición aleatoria o en inserciones de transposón en ADN episomal. Las librerías de ADN basadas en transposición pueden ser obtenidas a partir de secuencias purificadas de ADN episomal, como plásmidos, fósquidos, cósmidos o BACs, que tienen insertos únicos de ADN o a partir de conjuntos de secuencias de ADN episomal cada una teniendo un inserto de ADN diferente.
- (ii) Introducir una o más de dichas librerías de ADN episomal basadas en transposición de (i) dentro de células hospederas bacterianas por métodos estándar de transformación.
- (iii) Inducir la expresión de la T7RNA polimerasa en los aislados bacterianos resultantes transformados con las secuencias de ADN episomal basadas en transposición. La expresión específica de la T7RNA polimerasa a partir de cada inserción del transposón TnC_T7 proporciona una colección diversa de transcriptos de ADN o secuencias de ARN en la población de células bacterianas resultantes.
- (iv) Analizar dicha población de células hospederas bacterianas para identificar los aislados bacterianos específicos que expresan, pero no limitándose a, un gen reportero que codifica GFP. La expresión del gen reportero, como cualquier otro tipo de fenotipo bajo estudio usando el método revelado en la presente invención, se asocia con las secuencias específicas de ARN generadas en la bacteria analizada, lo cual a su vez se correlaciona con inserciones específicas del transposón TnC_T7 en las librerías originales de ADN episomal.

En una realización, la presente invención incluye ocho plásmidos correspondientes a vectores artificiales que resultan de inserciones aleatorias del transposón en pKR-C12, siendo este un plásmido que incluye un gen reportero silenciado que codifica para GFP. Los plásmidos incluidos en esta realización están caracterizados por tener un transposón TnC_T7 diferencial insertado en el plásmido blanco original y por tener la misma longitud total. La localización específica de la inserción del transposón en cada caso define la

eficiencia del respectivo aislado bacteriano transformado para expresar el gen reportero.

Ejemplos

Ejemplo 1: Expresión inducible de la T7RNAP a partir del promotor P_{BAD}

Un vector artificial para la expresión recombinante de la T7RNAP fue generado al clonar la secuencia codificante de la proteína mencionada en el sitio de múltiple clonaje del plásmido pBAD18-Cm. Para esto, la secuencia codificante de la T7RNAP se amplificó con Accuzyme como polimerasa de alta fidelidad (Bioline, Londres, Reino Unido), usando como molde ADN genómico purificado de la cepa BL21 de *E. coli* (Invitrogen-Life Technologies, Carisbad, CA, USA) y los iniciadores proporcionados en Seq-ID1 y Seq-ID2. Por otro lado, el vector pBAD18-Cm (Guzman LM, et al.; *J. Bacteriol.* 1995, 177:4121-30) fue linealizado por restricción enzimática con *Kpn*I y sus extremos de ADN reparados con la T4 ADN polimerasa (New England Biolabs, Ipswich, MA, USA). Después de la purificación del amplicón de PCR y del vector, se llevó a cabo una reacción de ligación y transformación en *E. coli* TOP10 (Invitrogen-Life Technologies, Carisbad, CA, USA), de acuerdo con métodos estándar conocidos en el estado del arte. La orientación correcta del inserto fue verificada por digestión con enzimas de restricción, PCR de colonia y/o secuenciación del ADN plasmídico aislado a partir de los clones bacterianos resultantes (FIG. 4).

En un paso siguiente, un clon de *E. coli* TOP10 alojando el pBAD18-Cm_t7rnap fue transformado con un plásmido que incluye la secuencia codificante del péptido AA localizado corriente abajo de un promotor T7 (pET28a_AA). La selección de los aislados bacterianos resultantes que incluyen ambos plásmidos fue realizada usando los marcadores de selección correspondientes para ambos vectores.

Para evaluar la expresión de la secuencia codificante de la T7RNAP a partir del promotor P_{BAD} del vector pBAD18-Cm_t7rnap, cultivos de células bacterianas fueron inducidos con isopropil β-D-l-tiogalactopiranosido (IPTG) o L-arabinosa, dependiendo del hospedero receptor final de los plásmidos mencionados (ya

sea *E. coli* BL21 o *E. coli* TOP10, respectivamente). La detección del péptido AA por anticuerpos anti-poli-histidina en ensayos de western blot fue llevada a cabo empleando extractos bacterianos totales, evidenciando las condiciones en las cuáles la T7RNAP pudo ser exitosamente expresada (FIG. 5).

Ejemplo 2: Instrucción para la clonación de la secuencia del transposón
TnC_T7

Con el fin de construir el plásmido que alberga la secuencia completa del transposón TnC_T7 se realizaron los siguientes pasos:

La Seq-ID3 fue diseñada para incluir el sitio R1-R2 de unión para la transposasa MuA, una región promotora T7 y los sitios para las enzimas de restricción EcoRI, *Bgl*II, *Asc*I y *Sam*HI (FIG.1). La secuencia de ADN resultante (Tn_A) fue sintetizada (Genscript, Piscataway, NJ, USA) y fue subsecuentemente clonada en los sitios únicos de restricción EcoRI y *fam*HI del plásmido pUC57. La orientación correcta de la inserción del Tn_A se verificó por digestiones con enzimas de restricción y/o secuenciación de ADN. El plásmido resultante se denota aquí como pUC57_Tn y se proporciona en Seq-ID4 (FIG. 2).

El gen de resistencia de kanamicina, incluyendo su promotor, se amplificó por PCR en dos reacciones independientes desde el plásmido pKD4 (*Datsenko KA, et al.; Proc. Natl. Acad. Sci. U. S. A. 2000, 97:6640-5*). Con el fin de reemplazar el sitio de restricción *Bgl*II en la secuencia de ADN ensamblada, los fragmentos amplificados resultantes del gen de resistencia a kanamicina fueron ligados después de la digestión enzimática con *Spe*I. La secuencia resultante del gen de resistencia al antibiótico se proporciona en Seq-ID5 (el sitio de restricción *Spe*I se muestra subrayado).

La secuencia de 1,214 pb proporcionada en Seq-ID5 fue digerida con *Asc*I y *Bam*HI, purificada y clonada en pUC57_Tn (Seq-ID4), después de la digestión del vector con las mismas enzimas de restricción. La orientación correcta del inserto se verificó en ADN plasmídico aislado de clones bacterianos resultantes por medio de digestión con enzimas de restricción, PCR de colonia y/o por

secuenciación de ADN del constructo final, el cual se denota aquí como el vector pUC57_Tn_kanAB (FIG. 3).

La secuencia codificante de la T7RNAP y el promotor inducible con arabinosa fueron amplificados con Accuzyme como polimerasa de alta fidelidad (Bioline, Londres, Reino Unido), usando ADN purificado del plásmido pBAD18-Cm_t7rnep (FIG. 4) como molde y los iniciadores proporcionados en Seq-ID6 and Seq-ID7 (los sitios de restricción *AscI* se muestran subrayados y no se incluyeron nucleótidos adicionales en las secuencias de los iniciadores para permitir una apropiada digestión de la enzima de restricción sobre el producto de PCR resultante). El amplicón de 3,122 pb, correspondiente a la secuencia pBAD_T7RNAP (Seq-ID8), fue insertado en el sitio único de restricción *AscI* de pUC57_Tn_kanAB (FIG. 3), generando el plásmido pUC57_Tn_kanABJ7 (FIG. 6). La orientación correcta del inserto se verificó por digestión con enzimas de restricción, PCR de colonia y/o por secuenciación de ADN, en ADN plasmídico aislado a partir de los clones bacterianos resultantes transformados con la correspondiente reacción de ligación.

La clonación del segundo extremo del transposón, denotado aquí como Tn_B se llevó a cabo en el vector pUC57_Tn_kanAB_t7 (FIG. 6) para generar el plásmido pUC57_TnC_T7 (FIG. 7). El Tn_B fue amplificado por PCR con Accuzyme como polimerasa de alta fidelidad (Bioline, Londres, Reino Unido), usando el ADN del plásmido pUC57_Tn (FIG. 2) como molde y los iniciadores proporcionados en Seq-ID9 y Seq-ID10 (los sitios de restricción *HindIII* se muestran subrayados y 3 nucleótidos adicionales en los extremos 5' de los iniciadores fueron incluidos para permitir una adecuada digestión enzimática sobre el producto de PCR correspondiente). El amplicón de 155 pb (Seq-ID11) y el plásmido pUC57_Tn_kanAB_t7 fueron ligados después de la digestión con la enzima *HindIII*. La orientación correcta de la inserción de Tn_B fue verificada por digestión con enzimas de restricción y/o por secuenciación de ADN.

Como resultado, el plásmido pUC57_TnC_T7 aloja el transposón TnC_T7, el cual en consecuencia incluye dos sitios flanqueantes R1-R2 de unión para la transposasa MuA, dos regiones promotoras T7, el gen de resistencia a kanamicina y la secuencia codificante de la T7RNAP bajo la regulación del

promotor P_{B^AD} (FIG. 8). El diseño final del plásmido pUC57_TnC_T7 permite liberar el transposón por restricción con *Bgl*W, haciendo que esté listo para reacciones *in vitro* con la transposasa MuA y con cualquier ADN episomal blanco. El realizar esta restricción enzimática con *Bgl*W ha mostrado ser crucial para generar los nucleótidos 5' sobresalientes requeridos para un ensamble y estabilidad eficientes del transpososoma Mu, así como para realizar las reacciones de transferencia de hebra {Savilahti H, et al.; *EMBO J.* 1995, 14:4893-903).

Ejemplo 3: Identificación de células bacterianas que expresan GFP como resultado de la transposición de TnC_T7 en ADN plasmídico

Para evaluar si el transposón TnC_T7 puede aumentar la expresión de genes en ADN episomal, se llevaron a cabo eventos de transposición con el transposón TnC_T7 en el plásmido sensor pKR-C12 (Riedel K, et al.; *Microbiology.* 2001, 147:3249-62), el cual es incapaz de expresar GFP en *E. coli* porque este hospedero bacteriano carece del sistema de *quorum sensing* necesario para dicha expresión (Riedel K, et al.; *Microbiology.* 2001, 147:3249-62). Por lo tanto y para efectos de la presente invención, la expresión de GFP a partir de pKR-C12 en *E. coli* sólo es posible si el proceso de transcripción empieza a partir de cualquiera de los promotores T7 proporcionados por TnC_T7.

El plásmido pKR-C12 purificado fue empleado como ADN episomal blanco para reacciones de transposición *in vitro* de TnC_T7 con la enzima transposasa MuA (Thermo Scientific, Waltham, MA, USA), siguiendo las recomendaciones del fabricante. Las reacciones resultantes fueron transformadas en la cepa bacteriana *E. coli* Epi300 (Epicentre-Illumina, Madison, WI, USA), de acuerdo con métodos estándar conocidos en el arte, empleando gentamicina y kanamicina como marcadores de selección. Los clones de *E. coli* Epi300 post-transposición de TnC_T7 en pKR-C12 fueron crecidos independientemente en medio LB hasta que alcanzaron una Densidad Óptica (OD)_{600nm} de 0.4 e inducidos con L-arabinosa 0.2% por 5 horas adicionales a 30°C. Los ensayos de detección de fluorescencia por espectrofotometría fueron llevados a cabo en un Synergy Microplate Reader (BioTek, Winooski, VT, USA). Cada cultivo

bacteriano fue evaluado en placas negras de poliestireno de 96 pozos con fondo claro (Sigma-Aldrich, Saint Louis, MO, USA) y analizado con una longitud de onda de excitación a 474 nm y emisión a 515 nm. Como resultado de este tipo de ensayos, clones bacterianos expresando GFP resultantes de la transposición de TnC_T7 en pKR-C12 fueron finalmente identificados (FIG. 10). Se analizaron los clones bacterianos post-transposición para localizar los sitios de inserción del TnC_T7 en pKR-C12, llevando a cabo análisis de secuenciación por Sanger desde iniciadores hibridados en la secuencia del transposón.

Alternativamente, ensayos de detección de fluorescencia en clones bacterianos post-transposición fueron realizados después de crecer las bacterias a 37°C por 14-16 horas en platos de LB-agar suplementado con L-arabinosa al 0.2% y los correspondientes marcadores de selección. En este caso, la expresión de GFP fue analizada usando el Sistema de Imágenes *in vivo* IVIS 200 (PerkinElmer, Waltham, MA USA) con los filtros de excitación y emisión de GFP y 15 s de exposición de luminiscencia (FIG. 11).

En consecuencia, el transposón TnC_T7 tuvo la habilidad de iniciar la transcripción de genes en ADN plasmídico y su validación como una herramienta genética en una cepa de *E. coli* diferente de BL21 indicó que la expresión de la T7RNAP ocurrió desde su gen correspondiente localizado dentro del transposón. Los plásmidos resultantes post-transposición en pKR-C12, obtenidos a partir de los clones bacterianos 1-8 mostrados en la FIG. 11, exhibieron patrones diferenciales de expresión de GFP dependiendo de la inserción específica del transposón TnC_T7.

Ejemplo 4: Instrucción para la clonación de la secuencia codificante de GFP en contexto metagenómico

Se generó un fósrido que incluye la secuencia codificante de GFP dentro de su inserto de ADN metagenómico, por la clonación de dicha secuencia en un sitio único de restricción. Para esto, la secuencia codificante de GFP (también denotada como *gfp*) fue amplificada, incluyendo un RBS corriente arriba, con Accuzyme como polimerasa de alta fidelidad (Bioline, Londres, Reino Unido), usando ADN purificado del plásmido pKR-C12 como molde y los iniciadores

proporcionados en Seq-ID12 y Seq-ID13 (se subrayan los sitios de restricción *AscI* y no se incluyeron nucleótidos adicionales en las secuencias de los iniciadores para permitir una adecuada digestión enzimática de restricción sobre el producto de PCR resultante). El amplicón de *gfp* de 918 pb (Seq-ID14) fue introducido dentro del inserto de ADN de un clon metagenómico. Por ejemplo, el ADN fosmídico purificado de un clon metagenómico que alberga un inserto de ADN perteneciente a una muestra de suelo fue linearizado por restricción enzimática y usado para insertar el amplicón de *gfp* descrito anteriormente. Por lo tanto, se digirió el aislado original de ADN fosmídico, denotado aquí como pCC2FOS_F076 con la enzima de restricción *AscI* y se ligó con el amplicón *gfp*, para generar el fósido F076_GFP (Seq-ID15; FIG. 12). La precisa orientación del inserto se verificó por digestión con enzimas de restricción, PCR de colonia y/o por secuenciación de ADN del constructo final, a partir de ADN fosmídico aislado de los clones bacterianos resultantes transformados con la correspondiente reacción de ligación.

Ejemplo 5: Identificación de células bacterianas que expresan GFP como resultado de la transposición de TnC_T7 en ADN fosmídico

Para evaluar la capacidad del transposón TnC_T7 de aumentar la expresión de genes en ADN fosmídico, se llevaron a cabo eventos de transposición en el fósido F076_GFP (FIG. 12). Por lo tanto, se usó el fósido F076_GFP purificado como ADN episomal blanco para reacciones de transposición *in vitro* de TnC_T7 con la enzima transposasa MuA (Thermo Scientific, Waltham, MA, USA), siguiendo las recomendaciones del fabricante. Las reacciones resultantes se transformaron en la cepa bacteriana *E. coli* Epi300 (Illumina Inc., San Diego, CA, USA), de acuerdo con métodos estándar conocidos en el arte, empleando kanamicina como marcador de selección.

Los ensayos de detección de fluorescencia en clones bacterianos post-transposición en F076_GFP se llevaron a cabo después de crecer las bacterias a 37°C por 14-16 horas en platos de LB-agar suplementado con L-arabinosa al 0.2% y el correspondiente marcador de selección. La expresión de GFP fue evaluada usando el Sistema de Imágenes *in vivo* IVIS 200 (PerkinElmer, Waltham, MA USA) con los filtros de excitación y emisión de GFP y 15 s de

exposición de luminiscencia (FIG. 13). Se llevaron a cabo análisis de secuenciación por Sanger desde los iniciadores de hibridación sobre la secuencia del transposón, para identificar los sitios de inserción de TnC_T7 en F076_GFP, a partir de los clones bacterianos resultantes post-transposición. En consecuencia, la validación en el uso del transposón TnC_T7 para iniciar la transcripción génica en ADN fosmídico se logró siguiendo los procedimientos aquí descritos.

Ejemplo 6: Identificación de células bacterianas que expresan actividad lipolítica como resultado de la transposición de TnC_T7

Se evaluó el aumento de otras actividades enzimáticas, diferentes de la expresión de GFP, en clones derivados de la metagenómica usando el transposón TnC_T7.

Por ejemplo, el vector fosmídico pCC2FOS_14gF2 aislado de una librería metagenómica construida con ADN derivado de suelo se empleó para detectar actividad lipolítica, ya que previamente se había identificado por análisis *in silico* un potencial sitio activo de lipasa (InterProScan: IPR002168) sobre el inserto del ADN metagenómico secuenciado.

Se usó el fósido pCC2FOS_14gF2 como ADN episomal blanco para reacciones de transposición *in vitro* de TnC_T7, como se describe en los ejemplos 3 y 5, ya que los ensayos funcionales previos para evaluar la degradación de tributirina en medio LB-agar o de degradación de 4-nitrofenil butirato (Sigma-Aldrich, Saint Louis, MO, USA) del clon metagenómico (alojando pCC2FOS_14gF2) no exhibieron diferencias significativas comparadas con la línea base para el control negativo de actividad lipolítica (*E. coli* Epi300 pCC2FOS).

Las reacciones de transposición de TnC_T7 sobre pCC2FOS_14gF2 fueron transformadas en la cepa bacteriana *E. coli* Epi300 y seleccionadas con cloranfenicol y kanamicina. Clones post-transposición fueron crecidos independientemente en medio LB hasta que alcanzaron una OD_{600nm} de 0.4 e inducidos con L-arabinosa al 0.2% por 5 horas adicionales a 37°C. Los cultivos bacterianos resultantes fueron normalizados por OD y sus respectivos pellets

lavados con solución tampón Tris-HCl y re-suspendidos en 1/5 de su volumen original en el tampón de Tris. Se obtuvieron extractos bacterianos completos luego de la lisis de células usando un Mini-Beadbeater-96 (Biospec Products, Bartlesville, OK, USA) y purificación por filtración. Siguiendo los métodos descritos, los ensayos funcionales con un set de clones de *E. coli* Epi300 pCC2FOS_14gF2 post-transposición mostraron incrementos significativos en actividad lipolítica por la degradación del 4-nitrofenil butirato (FIG. 11), al cuantificar por absorbancia a 410 nm en un NanoDrop 2000 (Thermo Scientific, Waltham, MA, USA) (FIG. 14).

REIVINDICACIONES

1. Un transposón para suministrar parcialmente la maquinaria transcripcional del hospedero durante análisis funcionales de librerías genómicas/metagenómicas, el transposón caracterizado porque comprende:

- (i) un marcador de selección diferente de aquel que está localizado en el vector esqueleto,
- (ii) un gen que codifica la T7RNA polimerasa bajo un promotor inducible,
- (iii) dos promotores flanqueantes T7 de lectura hacia el exterior,
- (iv) dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA.

2. El transposón de acuerdo con la reivindicación 1, donde el marcador de selección diferente de aquel que está localizado en el vector esqueleto es el gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional.

3. El transposón de acuerdo con la reivindicación 1, donde el promotor inducible para el gen que codifica la T7RNA polimerasa es el promotor inducible con arabinosa (PBAD)-

4. El transposón de acuerdo con la reivindicación 1, donde cada uno de los dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA se encuentra adyacente a un promotor individual T7 de lectura hacia el exterior.

5. Un vector artificial que contiene el transposón de las reivindicaciones 1 a 4, caracterizado porque dicho vector es el resultado de la combinación específica de los siguientes elementos de ADN:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) un marcador de selección diferente de aquel que está localizado en el vector esqueleto,

- (iii) un gen que codifica la T7RNA polimerasa bajo un promotor inducible,
- (iv) dos secuencias flanqueantes promotoras 17 de lectura hacia el exterior,
- (v) dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA.

6. El vector artificial de acuerdo con la reivindicación 5, el cual tiene un vector esqueleto con el origen de replicación ColE1/pMB1/pBR322/pUC y un gen de resistencia a ampicilina como marcador de selección.

7. El vector artificial de acuerdo con la reivindicación 5, donde el marcador de selección diferente de aquel que está localizado en el vector esqueleto es el gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional.

8. El vector artificial de acuerdo con la reivindicación 5, donde el promotor inducible para el gen que codifica la T7RNA polimerasa es el promotor inducible con arabinosa (PBAD)-

9. El vector artificial de acuerdo con la reivindicación 5, donde cada uno de los dos sitios de reconocimiento de repetición invertida para la transposasa MuA se encuentra adyacente a un promotor individual T7 de lectura hacia el exterior.

10. El vector artificial de acuerdo con las reivindicaciones 5 a 9, donde dicho vector puede ser mantenido en una célula hospedera bacteriana.

11. Una secuencia artificial de ADN para localizar un sitio de unión MuA y un promotor 17 en hebras diferentes de ADN, caracterizada porque comprende:

- (i) una secuencia promotora 17,
- (ii) un sitio de reconocimiento de repetición invertida para la transposasa MuA,

(iii) múltiples sitios flanqueantes de reconocimiento para endonucleasas de restricción.

12. La secuencia artificial de ADN de la reivindicación 11, donde dicho sitio de unión MuA corresponde a un sitio de reconocimiento de repetición invertida para la transposasa.

13. La secuencia de ADN artificial de la reivindicación 11, que incluye la secuencia promotora T7 para la interacción específica con la T7RNAP y los siguientes sitios flanqueantes de reconocimiento para endonucleasas de restricción: EcoRI, *Bgl*II, *Asc*I and *Sam*HI.

14. Un vector artificial resultado de la combinación específica de los siguientes elementos de ADN:

(i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,

(ii) una secuencia promotora T7 con una orientación específica, y

(iii) un sitio de reconocimiento de repetición invertida para la transposasa MuA en una orientación específica.

15. El vector artificial de la reivindicación 14, el cual tiene el cual tiene un vector esqueleto con el origen de replicación ColE1/pMB1/pBR322/pUC y un gen de resistencia a ampicilina para su selección después de la transformación en el hospedero bacteriano.

16. Un vector artificial resultado de la combinación específica de los siguientes elementos de ADN:

(i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,

(ii) una secuencia promotora T7 de lectura hacia el exterior,

(iii) un sitio de reconocimiento de repetición invertida para la transposasa MuA,

(iv) un marcador de selección diferente de aquel que está localizado en el vector esqueleto.

17. El vector artificial de la reivindicación 16, donde el marcador de selección diferente de aquel que está localizado en el vector esqueleto es el gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional.

18. Un vector artificial resultado de la combinación específica de los siguientes elementos de ADN:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) un gen que codifica la T7RNA polimerasa regulado por un promotor inducible.

19. El vector artificial de la reivindicación 18, el cual tiene un vector esqueleto con un origen de replicación ColE1/pMB1/pBR322/pUC, un gen de resistencia a cloranfenicol y la secuencia codificante de la T7RNAP clonada en el sitio único de restricción *Kpn*I.

20. El vector artificial de las reivindicaciones 18 y 19, donde la secuencia que codifica la T7RNAP está localizada corriente abajo del promotor inducible con arabinosa y corriente arriba de los terminadores transcripcionales *rmB T1* y *T2*.

21. Un vector artificial resultado de la combinación específica de los siguientes elementos de ADN:

- (i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección,
- (ii) una secuencia promotora T7 de lectura hacia el exterior,
- (iii) un sitio de reconocimiento de repetición invertida para la transposasa MuA,

(iv) un marcador de selección diferente de aquel que está localizado en el vector esqueleto.

(v) un gen que codifica la T7RNA polimerasa regulado por un promotor inducible.

22. El vector artificial de la reivindicación 21, el cual tiene un origen de replicación ColE1/pMB1/pBR322/pUC, un gen de resistencia a ampicilina y la secuencia codificante de la T7RNAP.

23. El vector artificial de las reivindicaciones 21 y 22, donde la secuencia que codifica la T7RNAP se localiza corriente abajo del promotor inducible con arabinosa.

24. El vector artificial de la reivindicación 21, donde el marcador de selección diferente de aquel que está localizado en el vector esqueleto es el gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional.

25. Un método para la construcción del transposón de acuerdo con las reivindicaciones 1 a 4, caracterizado porque comprende los pasos:

a) construir una secuencia artificial de ADN para localizar un sitio de unión MuA y un promotor T7 en hebras diferentes de ADN, donde la secuencia artificial de ADN comprende una secuencia promotora T7, un sitio de reconocimiento de repetición invertida para la transposasa MuA, y múltiples sitios flanqueantes de reconocimiento para endonucleasas de restricción;

b) construir un primer plásmido que comprende un vector esqueleto con el origen de replicación ColE1/pMB1/pBR322/pUC y un gen de resistencia a ampicilina para su selección después de la transformación del hospedero bacteriano, y donde el plásmido está construido por la clonación de la secuencia artificial de ADN del paso a) en los sitios únicos de restricción EcoRI y SamHI, posterior al tratamiento con enzimas de restricción tanto del vector esqueleto como de la secuencia artificial de ADN;

- c) construir un segundo plásmido que comprende las mencionadas regiones de ADN del primer plásmido [paso b)], más un gen de resistencia adicional, y donde dicho segundo plásmido está construido por la clonación del gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional, dentro de los sitios de restricción *AscI* y *BamHI* del primer plásmido [paso b)];
- d) construir un tercer plásmido, el cual comprende el origen de replicación ColE1/pMB1/pBR322/pUC, un gen de resistencia a cloranfenicol y la secuencia codificante de la T7RNAP, esta última clonada en el sitio único de restricción *KpnI*, y donde dicho tercer plásmido tiene la secuencia codificante de la T7RNAP localizada corriente abajo del promotor inducible con arabinosa y corriente arriba de los terminadores transcripcionales *rrnB T1* y *T2*;
- e) construir un cuarto plásmido, como resultado de la combinación específica de componentes del segundo y tercer plásmidos, y donde dicho cuarto plásmido resulta de la clonación del promotor inducible con arabinosa y la secuencia que codifica la T7RNAP en el sitio único de restricción *AscI* del segundo plásmido entre la secuencia artificial del paso a) y la secuencia de ADN de resistencia a kanamicina en el vector artificial;
- f) construir un quinto plásmido, el cual comprende todos los elementos estructurales del cuarto plásmido más un segundo extremo del transposón, donde dicho segundo extremo del transposón está clonado específicamente en el sitio único de restricción *HindIII* y
- g) liberar el transposón a través del tratamiento de la enzima de restricción *fliCII* sobre el quinto plásmido.

26. Un vector artificial resultado de la combinación específica de los siguientes elementos de ADN:

- (i) un fósido esqueleto con orígenes de replicación de bajo a alto número de copias y un marcador de selección,

(ii) un inserto de ADN metagenómico,

(iii) un gen que codifica una variante de la proteína verde fluorescente (GFP) con un sitio de unión al ribosoma (RBS) corriente arriba, ambos localizados específicamente dentro del inserto de ADN metagenómico.

27. El vector artificial de la reivindicación 26, donde el esqueleto fosmídico corresponde al vector pCC2FOS, el inserto metagenómico resultó de la clonación aleatoria del ADN metagenómico en pCC2FOS y el gen reportero corresponde a la secuencia que codifica GFP (incluyendo corriente arriba un sitio de unión a ribosoma (RBS)), este último clonado como un solo amplicón en el sitio único de restricción *AscI* del inserto metagenómico.

28. Un vector artificial resultado de la inserción aleatoria del transposón de las reivindicaciones 1 a 4, que comprende:

(i) un vector esqueleto con su propio marcador de selección y un gen reportero silenciado,

(ii) una inserción diferencial del transposón de las reivindicaciones 1 a 4 a lo largo de la secuencia de ADN de (i).

29. El vector artificial de la reivindicación 28, el cual corresponde a uno de los ocho plásmidos resultantes de las inserciones aleatorias del transposón TnC_T7 en el plásmido pKR-C12.

30. Un método para aumentar la transcripción de ADN como paso inicial de la expresión de genes foráneos, caracterizado porque comprende:

(i) generar librerías de transposición de ADN episomal, como resultado de la inserción aleatoria del transposón TnC_T7 purificado, donde dicho ADN episomal incluye plásmidos, fósidos, cósmidos o BACs,

(ii) introducir dichas librerías de transposición de ADN episomal de (i) dentro de células hospederas,

(iii) expresar la T7RNA polimerasa codificada en el transposón TnC_T7, para generar poblaciones de células hospederas bacterianas con colecciones diversas de transcriptos derivados de ADN episomal,

- (iv) analizar dichas poblaciones de células hospederas bacterianas para identificar los aislados bacterianos que expresan un gen reportero o cualquier otra función deseada.

REIVINDICACIONES MODIFICADAS
recibidas por la oficina Internacional el 12 de septiembre 2016 (12.09.2016)

1. Un transposón para suministrar parcialmente la maquinaria transcripcional del hospedero durante análisis funcionales de librerías genómicas/metagenómicas, el transposón caracterizado porque comprende:

(i) un marcador de selección diferente de aquel que está localizado en el vector esqueleto, el cual es el gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional;

(ii) un gen que codifica la T7RNA polimerasa bajo un promotor inducible, donde dicho promotor inducible para el gen que codifica la T7RNA polimerasa es el promotor inducible con arabinosa (PBAD);

(iii) dos promotores flanqueantes T7 de lectura hacia el exterior; y

(iv) dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA,

donde cada uno de los dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA se encuentra adyacente a un promotor individual T7 de lectura hacia el exterior.

2. Un vector artificial que contiene el transposón de la reivindicación 1, caracterizado porque dicho vector es el resultado de la combinación específica de los siguientes elementos de ADN:

(i) un vector esqueleto con un origen de replicación de alto número de copias y un marcador de selección, donde el vector esqueleto tiene origen de replicación ColE1/pMB1/pBR322/pUC y un gen de resistencia a ampicilina como marcador de selección;

(ii) un marcador de selección diferente de aquel que está localizado en el vector esqueleto, el cual es el gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional;

(iii) un gen que codifica la T7RNA polimerasa bajo un promotor inducible, donde dicho promotor inducible es el promotor inducible con arabinosa (PBAD);

(iv) dos secuencias flanqueantes promotoras 11 de lectura hacia el exterior; y

(v) dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA,

donde cada uno de los dos sitios de reconocimiento de repetición invertida para la transposasa MuA se encuentra adyacente a un promotor individual 11 de lectura hacia el exterior.

3. El vector artificial de acuerdo con la reivindicación 2, donde dicho vector puede ser mantenido en una célula hospedera bacteriana.

4. Una secuencia artificial de ADN para localizar un sitio de unión MuA y un promotor 11 en hebras diferentes de ADN de conformidad con el transposón de la reivindicación 1, caracterizada porque comprende:

(i) una secuencia promotora T7;

(ii) un sitio de reconocimiento de repetición invertida para la transposasa MuA; y

(iii) múltiples sitios flanqueantes de reconocimiento para endonucleasas de restricción,

donde dicho sitio de unión MuA corresponde a un sitio de reconocimiento de repetición invertida para la transposasa, y donde dicha secuencia de ADN artificial incluye la secuencia promotora 11 para la interacción específica con la T7RNAP y los siguientes sitios flanqueantes de reconocimiento para endonucleasas de restricción: EcoRI, BglII, AscI and BamHI.

5. Un método para la construcción del transposón de acuerdo con la reivindicación 1, caracterizado porque comprende los siguientes pasos:

a) construir una secuencia artificial de ADN para localizar un sitio de unión MuA y un promotor 11 en hebras diferentes de ADN, donde la secuencia

artificial de ADN comprende una secuencia promotora T7, un sitio de reconocimiento de repetición invertida para la transposasa MuA, y múltiples sitios flanqueantes de reconocimiento para endonucleasas de restricción;

- b) construir un primer plásmido que comprende un vector esqueleto con el origen de replicación ColE1/pMB1/pBR322/pUC y un gen de resistencia a ampicilina para su selección después de la transformación del hospedero bacteriano, y donde el plásmido está construido por la clonación de la secuencia artificial de ADN del paso a) en los sitios únicos de restricción *EcoRI* y *Bam*YW, posterior al tratamiento con enzimas de restricción tanto del vector esqueleto como de la secuencia artificial de ADN;
- c) construir un segundo plásmido que comprende las mencionadas regiones de ADN del primer plásmido [paso b)], más un gen de resistencia adicional, y donde dicho segundo plásmido está construido por la clonación del gen de resistencia a kanamicina, incluyendo tanto su promotor como su señal de terminación transcripcional, dentro de los sitios de restricción *Asa* y *Bam*H\ del primer plásmido [paso b)];
- d) construir un tercer plásmido, el cual comprende el origen de replicación ColE1/pMB1/pBR322/pUC, un gen de resistencia a cloranfenicol y la secuencia codificante de la T7RNAP, esta última clonada en el sitio único de restricción *Kpn*I, y donde dicho tercer plásmido tiene la secuencia codificante de la T7RNAP localizada corriente abajo del promotor inducible con arabinosa y corriente arriba de los terminadores transcripcionales *rrnB* T1 y T2;
- e) construir un cuarto plásmido, como resultado de la combinación específica de componentes del segundo y tercer plásmidos, y donde dicho cuarto plásmido resulta de la clonación del promotor inducible con arabinosa y la secuencia que codifica la T7RNAP en el sitio único de restricción *Asd* del

segundo plásmido entre la secuencia artificial del paso a) y la secuencia de ADN de resistencia a kanamicina en el vector artificial;

f) construir un quinto plásmido, el cual comprende todos los elementos estructurales del cuarto plásmido más un segundo extremo del transposón, donde dicho segundo extremo del transposón está clonado específicamente en el sitio único de restricción *HindIII*; y

g) liberar el transposón a través del tratamiento de la enzima de restricción *BglII* sobre el quinto plásmido.

6. Un vector artificial resultado de la inserción aleatoria del transposón de la reivindicación 1, que comprende:

(i) un vector esqueleto con su propio marcador de selección y un gen reportero silenciado; y

(ii) una inserción diferencial del transposón de las reivindicaciones 1 a 4 a lo largo de la secuencia de ADN de (i).

7. El vector artificial de la reivindicación 6, el cual corresponde a uno de los ocho plásmidos resultantes de las inserciones aleatorias del transposón TnC_T7 de acuerdo con la reivindicación 1 en el plásmido pKR-C12.

8. Un método para aumentar la transcripción de ADN como paso inicial de la expresión de genes foráneos, caracterizado porque comprende:

(i) generar librerías de transposición de ADN episomal, como resultado de la inserción aleatoria del transposón TnC_T7 purificado de acuerdo con la reivindicación 1, donde dicho ADN episomal incluye plásmidos, fósidos, cósmidos o BACs; e,

(ii) introducir dichas librerías de transposición de ADN episomal de (i) dentro de células hospederas.

(iii) expresar la T7RNA polimerasa codificada en el transposón TnC_T7, para generar poblaciones de células hospederas bacterianas con colecciones diversas de transcriptos derivados de ADN episomal,

(iv) analizar dichas poblaciones de células hospederas bacterianas para identificar los aislados bacterianos que expresan un gen reportero o cualquier otra función deseada.

FIG. 1

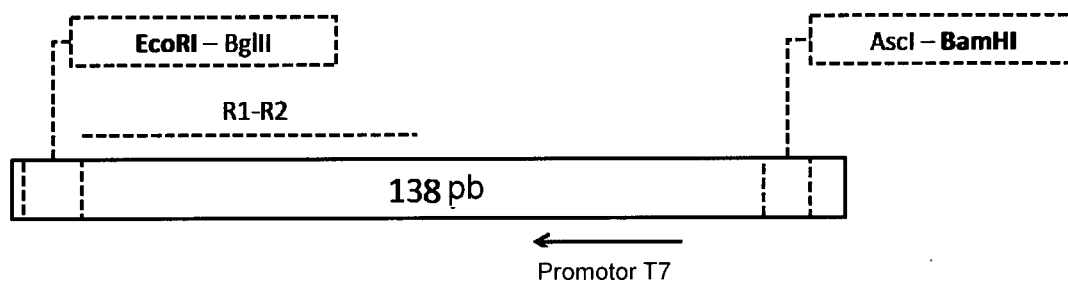


FIG. 2

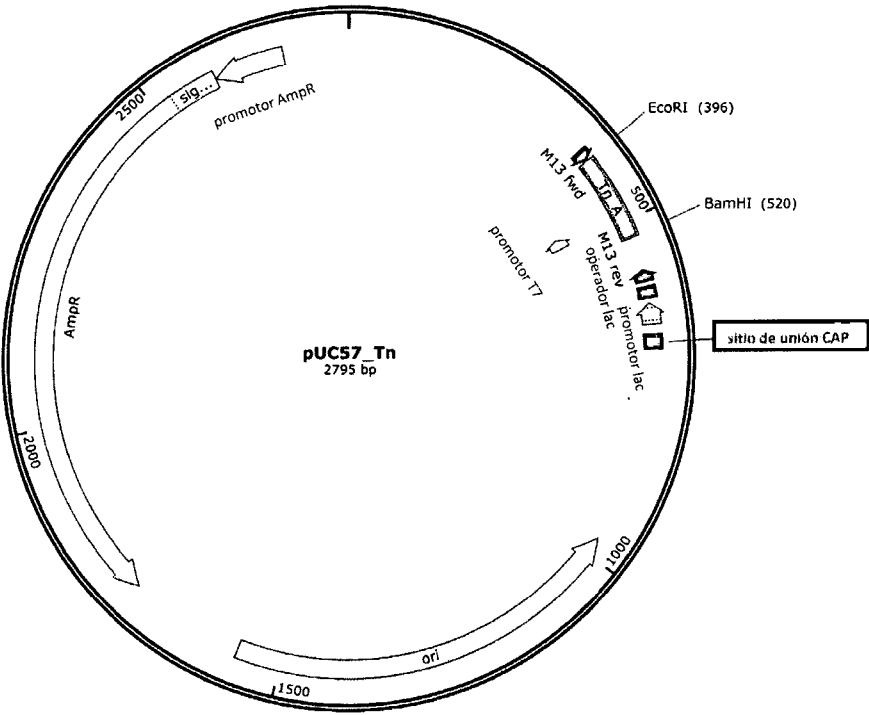


FIG. 3

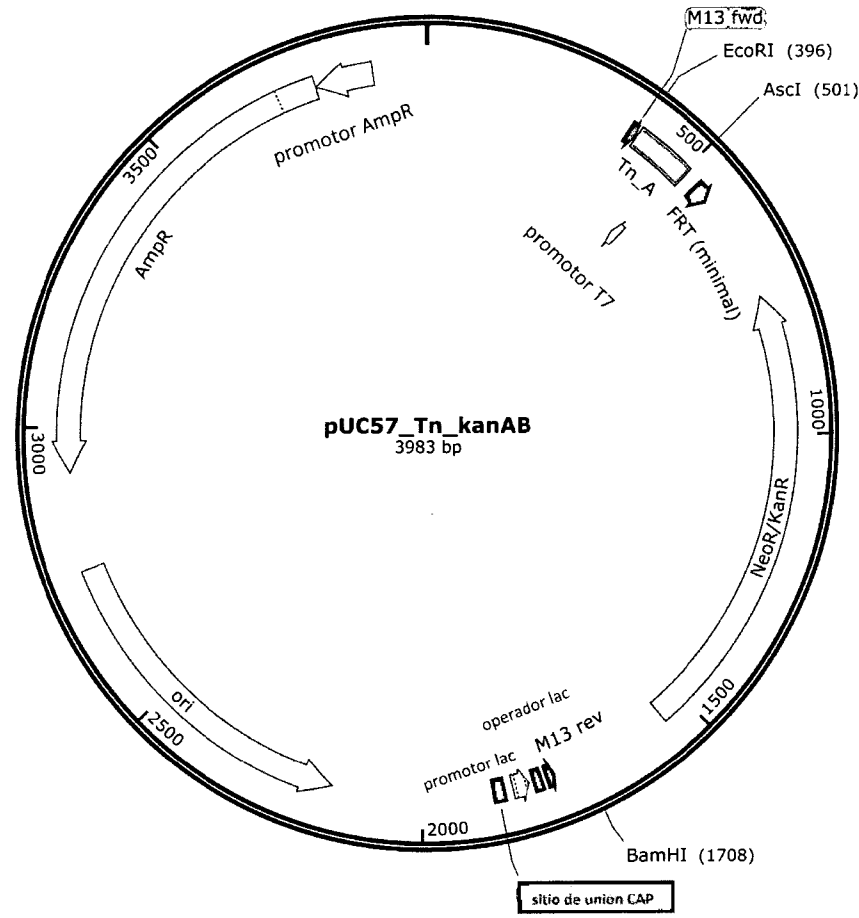


FIG. 4

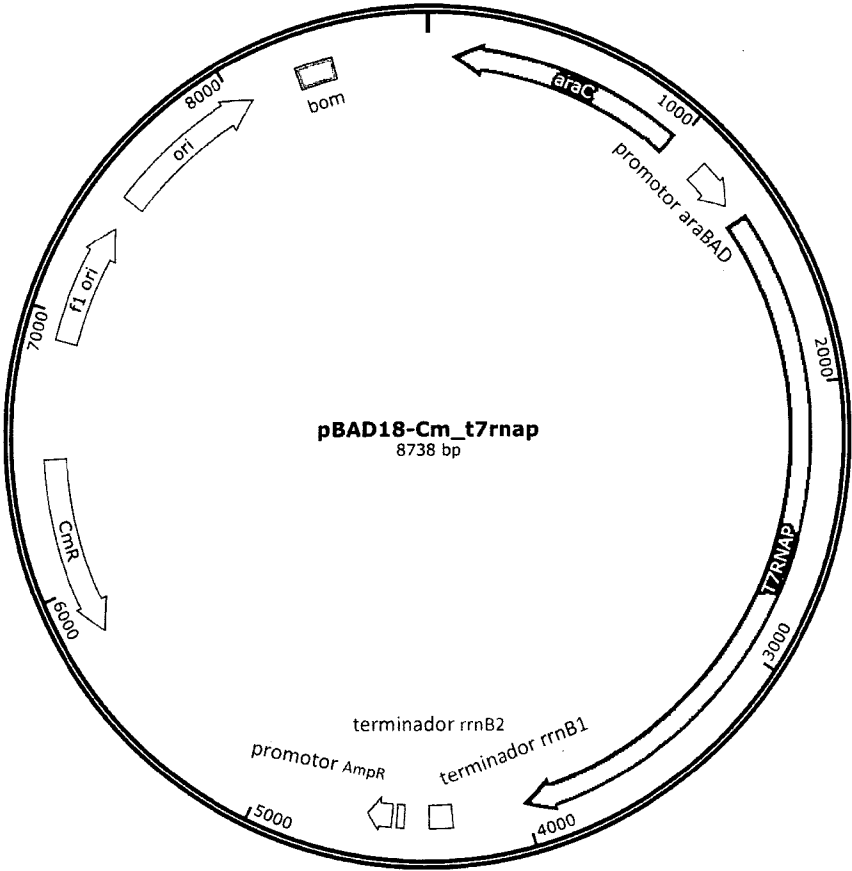


FIG. 5

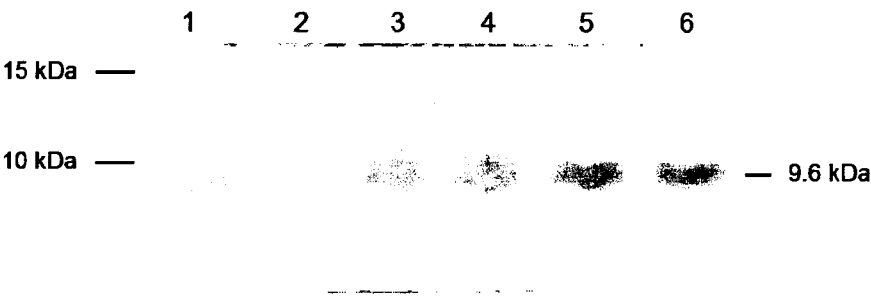


FIG. 6

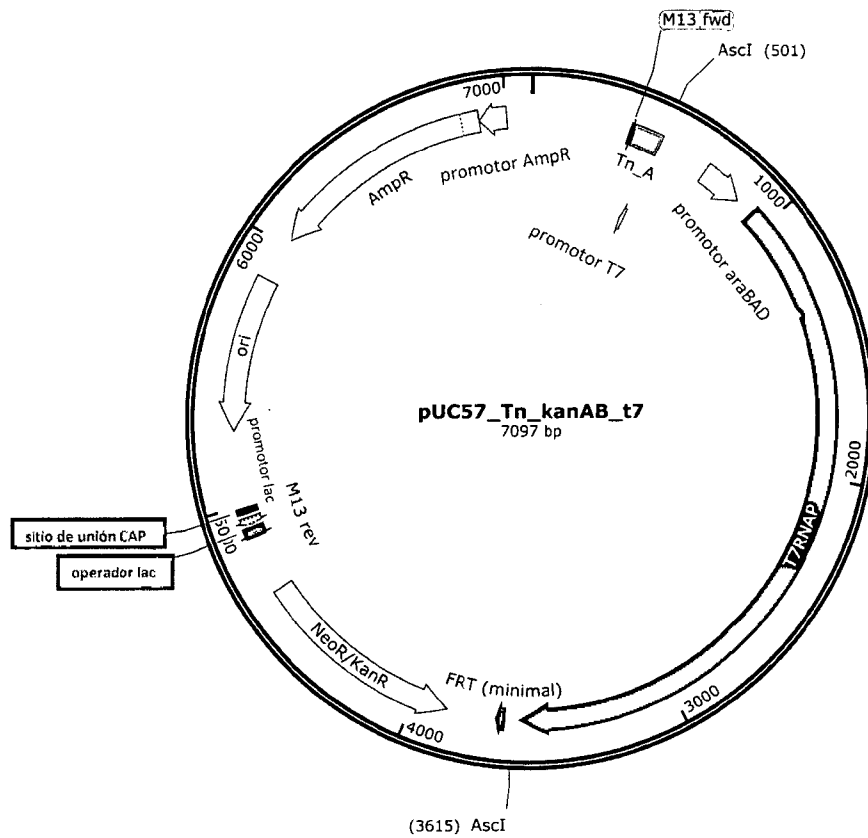


FIG. 7

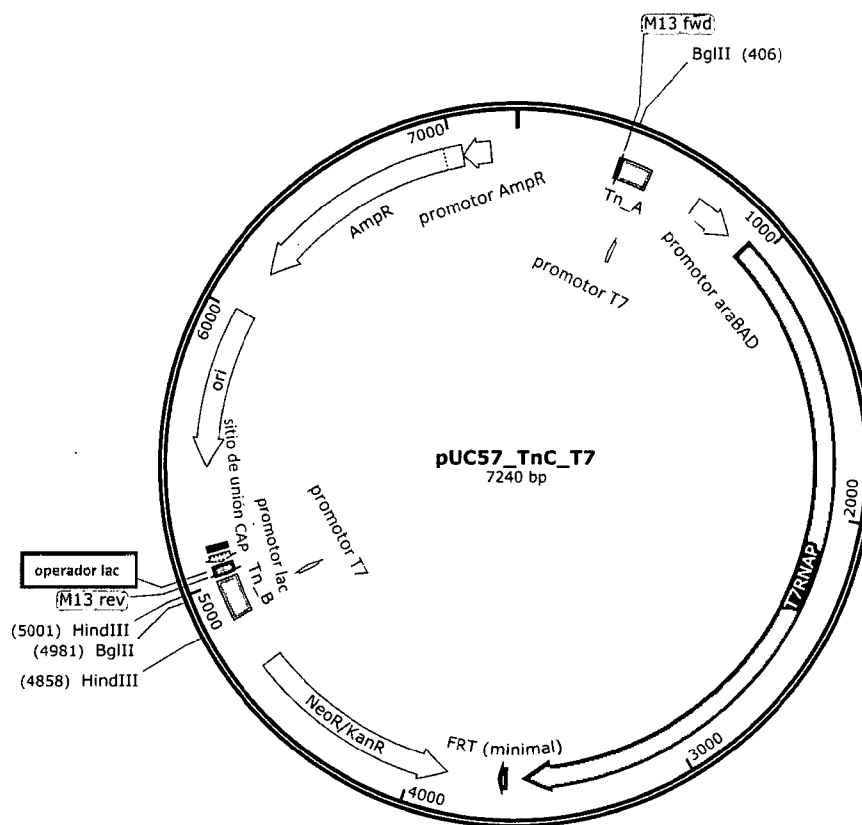


FIG. 8

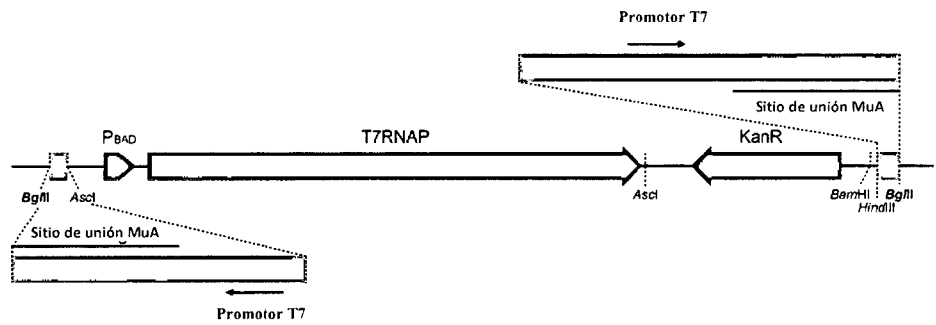


FIG. 9

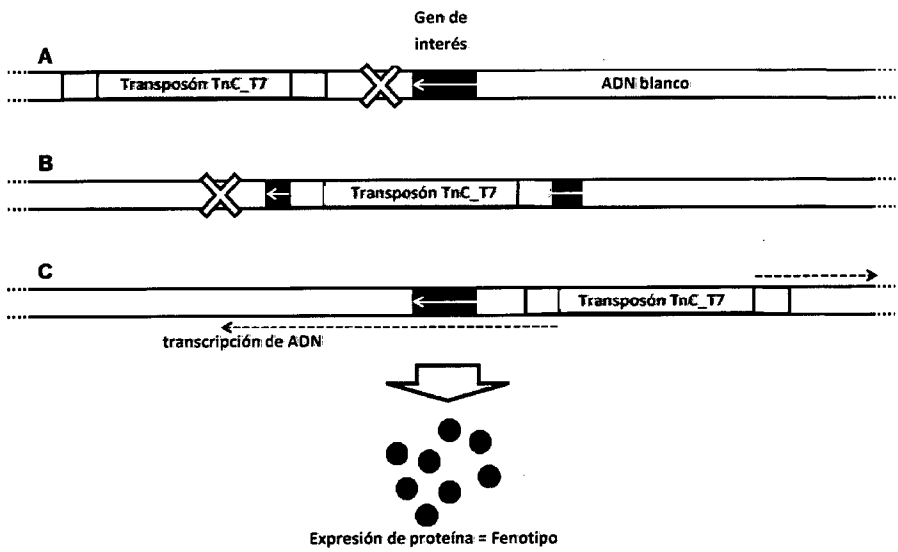


FIG. 10

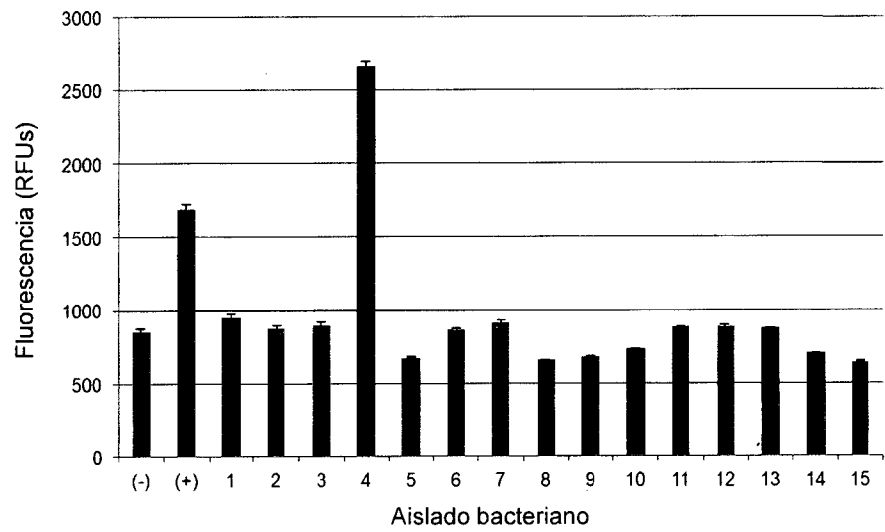


FIG. 11

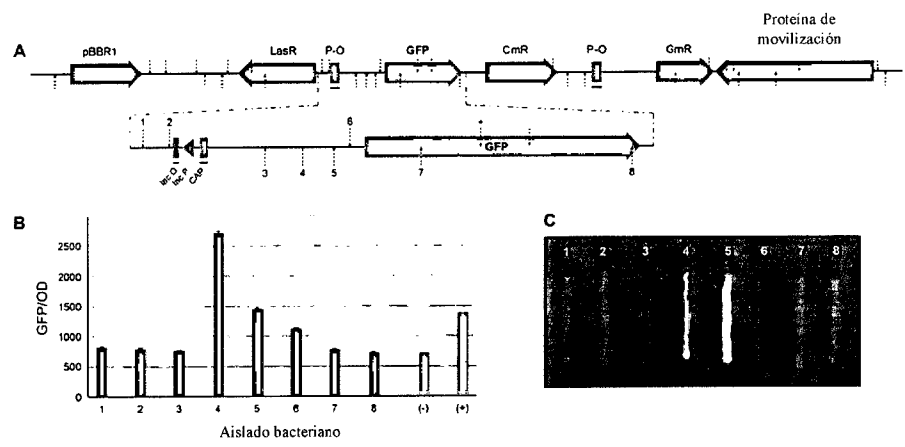


FIG. 12

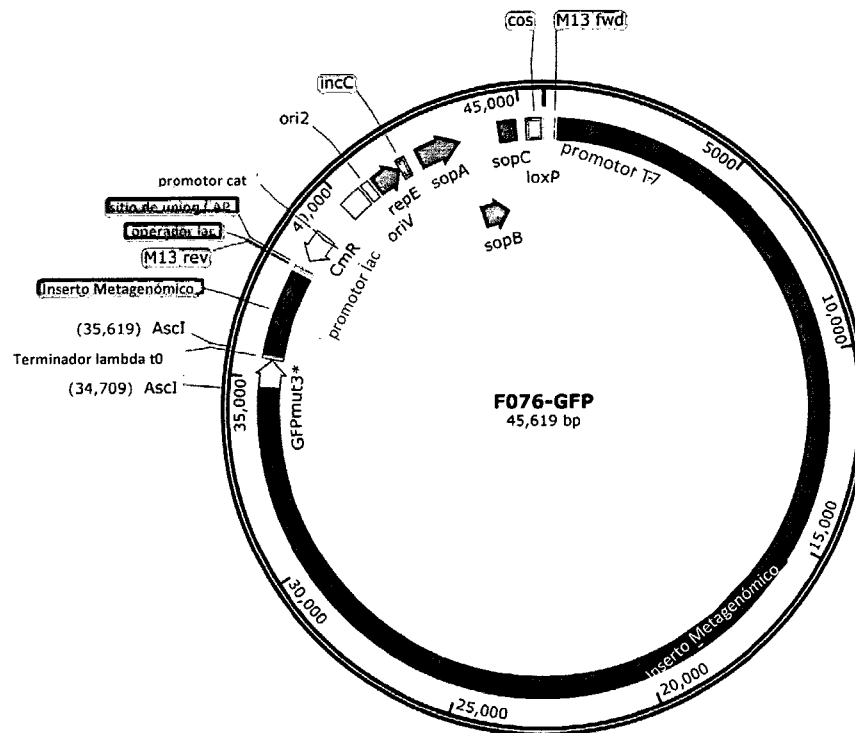


FIG. 13

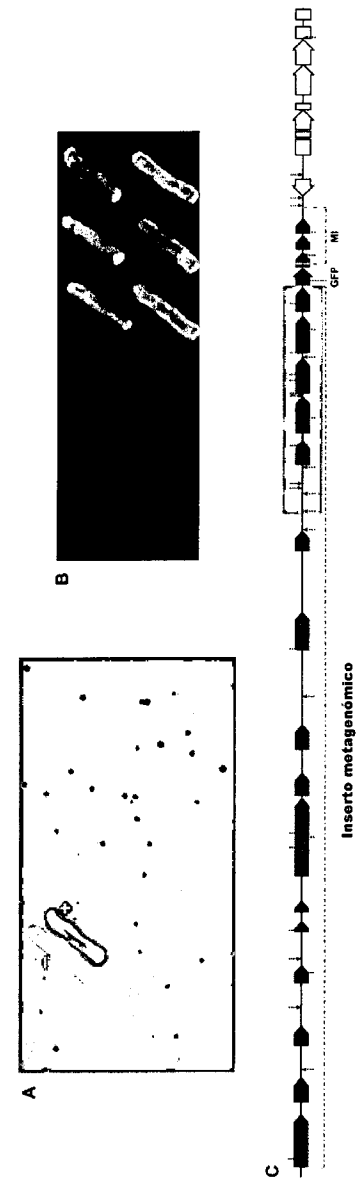
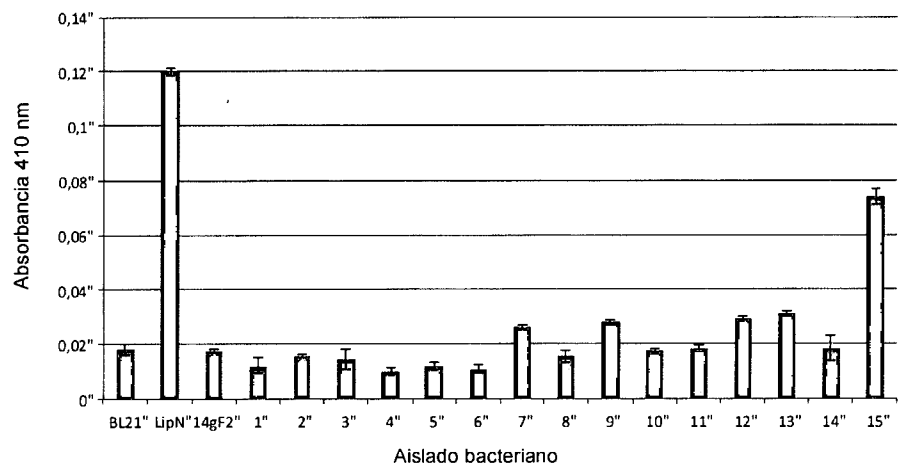


FIG. 14



INTERNATIONAL SEARCH REPORT

International application No.

PCT/CO20 15/000010

A. CLASSIFICATION OF SUBJECT MATTER

(CIP): C12N15/67, C12N15/70 (2015.01)

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)

(CIP): C12N15/67, C12N15/70

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practicable, search terms used)

STN, DIALOG, NCBI, PATENTSCOPE, ESPACENET, USPTO, GOOGLE SCHOLAR

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MONGUI, A. et al. Construction of a Transposon Capable of Enhancing Functional Expression of Genes in Episomal DNAs. ASM2014, 114th General Meeting, American Society for Microbiology, May 17-20, 2014, Boston, Massachusetts Poster Session, Presentación: H-1217. [retrieved from internet 14.12.2015] retrieved from internet : <URL: http://www.asmonlineeducation.com/php/asm2014abstracts/data/index.htm > the whole document	1-10, 21-24, 28, 29
X	LEGGEWIE, C. et al. A novel transposon for functional expression of DNA libraries, Journal of Biotechnology, 2006, Vol. 123 (3): 281-287. ISSN 0168-1656. the whole document	11-17
Y	cited in the application	1-10, 21-24, 28, 29

☒ 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 another 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

30 December 2015 30/12/2015

Date of mailing of the international search report

28 January 2016 28/01/2016

Name and mailing address of the TSA/

INAPI, Av. Libertador Bernardo O'Higgins 194, Piso 17, Santiago, Chile

Facsimile No.

Authorized officer

GARRIDO GAMBOA, Carolina

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CO20 15/00001 0

C [Continuation] DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCKINNEY, J. et al. Tightly Regulated Gene Expression System in Salmonella enterica Serovar Typhimurium. Journal of Bacteriology. 2002, vol. 184(21):6056-6059. DOI:10.1128/JB.184.21.6056-6059.2002. the whole document	18-20
A	LÄMMLE, K. et al. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning, Journal of Biotechnology, 2007, Vol. 127(4): 575-592, ISSN 0168-1656. the whole document cited in the application	
A	TERRÓN-GONZALÉZ, L. et al. Heterologous viral expression systems in fosmid vectors increase the functional analysis potential of metagenomic libraries. Scientific Reports 2013, 3: 1107. doi: 10.1038/srep01107 the whole document	
A	CULLIGAN E. P. et al. Metagenomics and novel gene discovery: Promise and potential for novel therapeutics. Virulence. 2014;5 (3):399-412. doi: 10.4161/viru.27208. pages 402-403	
A	US6159736 A (WISCONSIN ALUMNI RESEARCH FOUNDATION) 12 December 2000 the whole document	
A	WO2010049807 A2 (GENE BRIDGES GMBH) 06 May 2010 the whole document	
A	WO2000078977 A1 (AVENTIS PHARMA INC.) 28 December 2000 the whole document	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CO2015/000010

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item I.c 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 sequence listing:

a. ☐ forming part of the international application as filed:

☐ in the form of an Annex C/ST.25 text file.

☐ on paper or in the form of an image file.

b. ☒ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

☒ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).

☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).

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 forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CO20 15/000010

Box No. II Observations where certain claims were found insearchable (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.: **26-27**
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:

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:

see extra sheet

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 additional fees, this Authority did not invite payment of additional fees.
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.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CO2015/000010

US6159736 (A)	12-12-2000	AT282693 (T)	15-12-2004
		AU758960 (B2)	03-04-2003
		AU6057399 (A)	10-04-2000
		CA2343000 (C);	30-03-2000
		CA2343000 (A1)	30-03-2000
		CN100420747 (C);	24-10-2001
		CN1319135 (A)	24-10-2001
		DE69922032 (T2)	01-12-2005
		DK1 115856 (T3)	21-02-2005
		EP1 115856 (A1);	18-07-2001
		EP1 115856 (B1)	18-07-2001
		ES2228170 (T3)	01-04-2005
		JP4464563 (B2);	24-09-2002
		JP2002531062 (A)	24-09-2002
		JP201 006881 3 (A)	02-04-2010
		JP201 3 165731 (A)	29-08-2013
		PL346772 (A1);	25-02-2002
		PL1 97408 (B1)	25-02-2002
		RU2237715 (C2)	10-10-2004
		US6294385 (B1)	25-09-2001
		WO0017343 (A1)	30-03-2000
WO201 0049807 (A2)	06-05-2010	WO201 0049807 (A3)	06-05-2010
		CN102264905 (A)	30-1 1-201 1
WO2000078977 (A1)	28-12-2000	AU5496600 (A)	09-01-2001
		BR001 1765 (A)	05-03-2002
		CA2375821 (A1)	28-12-2000
		EP1 196613 (A1)	17-04-2002
		JP2003503029 (A)	28-01-2003
		MXPA01012406 (A)	04-06-2002
		NO20016153 (A)	15-02-2002
		NZ515502 (A)	27-08-2004
		US6818441 (B1)	16-11-2004
		ZA2001 09441 (A)	30-04-2003

(Continuation of Box II)

Claims 26 and 27 relate to an artificial vector formed by a backbone fosmid, particularly pCC2FOS with a metagenomic DNA insert, in which the metagenomic insert comprises a gene encoding the green fluorescent protein (GFP) with an upstream ribosomal binding site (RBS). The pCC2FOS vector is a commercial vector designed to generate metagenomic libraries and a metagenomic DNA insert corresponds to an undetermined sequence unless sequenced. In fact, it is most likely that each of the metagenomic samples taken from the environment will be different from one another and, as such, each metagenomic library will also be different. Consequently, it is not possible to determine if a metagenomic insert will have a sequence comprising a gene encoding the GFP and if it has an RBS. This lack of clarity is such as to render a meaningful search over the whole of the scope of these claims impossible. Consequently, no search has been carried out for claims 26 and 27.

(Continuation of Box III)

There is no single inventive concept linking all of the claims. It is possible to distinguish first a transposon, which is designed to partially provide the transcriptional machinery of the host during functional analysis of genomic/metagenomic libraries, the vector comprising same, the construction method thereof, and its use in the method in order to increase the transcription of episomal DNA; as well as an artificial vector which can be used to test the efficiency of the transposon in metagenomic inserts, which has no structural relationship with the aforementioned transposon, and which also solves a different technical problem, namely that of providing a test vector for elements that can be used to increase the transcription of silenced genes in a metagenomic context. In addition, vectors are claimed that would correspond to those generated during the design of the aforementioned transposon. Said vectors do not comprise all of the elements that define the transposon and, consequently, they do not have all of the elements required to solve the technical problem. There is no one feature common to all of these vectors that would allow them to be grouped together according to a single inventive concept, since the only shared feature is the presence of an origin of replication and a selection marker, which are elements common to vector plasmids in general.

In conclusión, a priori, the following groups can be distinguished:

Group 1 (claims 1-10, 25 and 28-30): Transposon that partially provides the transcriptional machinery of the host during functional analysis of genomic/metagenomic libraries, comprising: (i) a selection marker different from the backbone vector, (ii) a gene encoding T7 RNA polymerase under inducible promoter control, (iii) two outward-reading, flanking T7 promoters, (iv) two inverted repeat recognition flanking sites for the MuA transposase; a vector comprising the transposon; a method for constructing the transposon; and a method using the transposon in order to increase transcription of DNA for expression of foreign genes.

Group 2 (claims 11-13): Artificial DNA sequence comprising an MuA binding site, a T7 promoter in different DNA strands, and multiple flanking recognition sites for restriction endonucleases.

Group 3 (claims 14-15): Artificial vector comprising a backbone vector with an origin of replication that has a high copy number, a selection marker, a T7 promoter sequence, and an MuA transposase recognition site.

Group 4 (claims 16-17): Artificial vector comprising a backbone vector with an origin of replication that has a high copy number, a selection marker, an outward-reading T7 promoter sequence, an MuA transposase recognition site, and a selection marker different from that located in the backbone vector.

Group 5 (claims 18-20): Artificial vector comprising a backbone vector with an origin of replication that has a high copy number, a selection marker, and a gene encoding the T7 RNA polymerase regulated by an inducible promoter.

Group 6 (claims 21-24): Artificial vector comprising a backbone vector with an origin of replication that has a high copy number, a selection marker, an outward-reading T7 promoter sequence, an MuA transposase recognition site, a selection marker different from that located in the backbone vector, and a gene encoding the T7 RNA polymerase regulated by an inducible promoter.

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional N°

PCT/CO2015/000010

A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

(CIP): C12N15/67, C12N15/70 (2015.01)

De acuerdo con la Clasificación Internacional de Patentes (CIP) o según la clasificación nacional y CIP.

B. SECTORES COMPRENDIDOS POR LA BÚSQUEDA

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

(CIP): C12N15/67, C12N15/70

Otra documentación consultada, además de la documentación mínima, en la medida en que tales documentos formen parte de los sectores comprendidos por la búsqueda

Bases de datos electrónicas consultadas durante la búsqueda internacional (nombre de la base de datos y, si es posible, términos de búsqueda utilizados) **STN, DIALOG, NCBI, PATENTSCOPE, ESPACENET, USPTO, GOOGLE SCHOLAR**

C. DOCUMENTOS CONSIDERADOS RELEVANTES

Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones N°
Y	MONGUI, A. et al. Construction of a Transposon Capable of Enhancing Functional Expression of Genes in Episomal DNAs. ASM2014, 114th General Meeting, American Society for Microbiology, May 17-20, 2014, Boston, Massachusetts Poster Session, Presentación: H-1217. [Recuperado de internet 14.12.2015] Recuperado desde internet: <URL: http://www.asmonlineeducation.com/php/asm2014abstracts/data/index.htm > Todo el documento	1-10, 21-24, 28, 29
X	LEGGEWIE, C. et al. A novel transposon for functional expression of DNA libraries, Journal of Biotechnology, 2006, Vol. 123 (3): 281-287. ISSN 0168-1656. Todo el documento	11-17
Y	"Citado en la solicitud"	1-10, 21-24, 28, 29

☒ En la continuación del Recuadro C se relacionan otros documentos

☒ Los documentos de familias de patentes se indican en el Anexo

* Categorías especiales de documentos citados:	"T"	documento ulterior publicado con posterioridad a la fecha de presentación internacional o de prioridad que no pertenece al estado de la técnica pertinente pero que se cita por permitir la comprensión del principio o teoría que constituye la base de la invención.
"A" documento que define el estado general de la técnica no considerado como particularmente relevante.	"X"	documento particularmente relevante; la invención reivindicada no puede considerarse nueva o que implique una actividad inventiva por referencia al documento aisladamente considerado.
"E" solicitud de patente o patente anterior pero publicada en la fecha de presentación internacional o en fecha posterior.	"Y"	documento particularmente relevante; la invención reivindicada no puede considerarse que implique una actividad inventiva cuando el documento se asocia a otro u otros documentos de la misma naturaleza, cuya combinación resulta evidente para un experto en la materia.
"L" documento que puede plantear dudas sobre una reivindicación de prioridad o que se cita para determinar la fecha de publicación de otra cita o por una razón especial (como la indicada).	"&"	documento que forma parte de la misma familia de patentes.
"O" documento que se refiere a una divulgación oral, a una utilización, a una exposición o a cualquier otro medio.		
"P" documento publicado antes de la fecha de presentación internacional pero con posterioridad a la fecha de prioridad reivindicada.		

Fecha en que se ha concluido efectivamente la búsqueda internacional.
30/12/2015 30/diciembre/2015

Fecha de expedición del informe de búsqueda internacional
28/01/2016 28/enero/2016

Nombre y dirección postal de la Administración encargada de la búsqueda internacional
INAPI, Av. Libertador Bernardo O'Higgins 194, Piso 17, Santiago, Chile

Funcionario autorizado
GARRIDO GAMBOA, Carolina

N° de fax

N° de teléfono 56-2-28870551 56-2-28870550

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional N°

PCT/CO201 5/000010

Recuadro I Secuencia(s) de nucleótidos y/o de aminoácidos (continuación del punto I.c de la primera hoja)

1. En lo que se refiere a las secuencias de nucleótidos y/o de aminoácidos divulgadas en la solicitud internacional y necesarias para la invención reivindicada, la búsqueda se ha llevado a cabo sobre la base de una lista de secuencias:
 - a. ☐ que forma parte de la solicitud internacional tal y como se presentó:
 - ☐ en formato de archivo de texto según Anexo C/ST.2₅.
 - ☐ en formato papel o en formato de archivo de imagen.
 - b. ☒ presentada junto con la solicitud internacional de acuerdo a la Regla 13/er. 1.a) del PCT exclusivamente a los fines de la búsqueda, en formato de archivo de texto según Anexo C/ST.2₅.
 - c. ☐ Presentada con posterioridad a la fecha de presentación exclusivamente a los fines de la búsqueda internacional:
 - ☐ en formato de archivo de texto según Anexo C/ST.25 (Regla 13/er. 1.a)).
 - ☐ en formato papel o en formato de archivo de imagen (Regla 13/er. 1.b) e Instrucciones Administrativas, sección 713).
2. ☐ Además, en caso de que se haya presentado más de una versión o copia de una lista de secuencias, se ha entregado la declaración requerida de que la información contenida en las copias subsiguientes o adicionales es idéntica a la que formaba parte de la solicitud tal y como se presentó o no va más allá de lo presentado inicialmente
3. Comentarios adicionales:

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional N°

PCT/CO2015/000010

Recuadro II Observaciones cuando se estime que algunas reivindicaciones no pueden ser objeto de búsqueda (continuación del punto 2 de la primera hoja)

Este informe de búsqueda internacional no se ha realizado en relación a ciertas reivindicaciones según el Artículo 17.2a) por los siguientes motivos:

1. ☐ Las reivindicaciones N°s:
se refieren a un objeto con respecto al cual esta Administración no está obligada a proceder a la búsqueda, a saber:
2. ☒ Las reivindicaciones N°s: 26-27
se refieren a elementos de la solicitud internacional que no cumplen con los requisitos establecidos, de tal modo que no pueda efectuarse una búsqueda provechosa, concretamente:
(Ver Hoja adicional)
3. ☐ Las reivindicaciones N°s:
son reivindicaciones dependientes y no están redactadas de conformidad con los párrafos segundo y tercero de la Regla 6.4.a).

Recuadro III Observaciones cuando falta unidad de invención (continuación del punto 3 de la primera hoja)

La Administración encargada de la búsqueda internacional ha detectado varias invenciones en la presente solicitud internacional, a saber:
(Ver Hoja adicional)

1. ☐ Dado que todas las tasas adicionales requeridas han sido satisfechas por el solicitante dentro del plazo, el presente informe de búsqueda de tipo internacional comprende todas las reivindicaciones que pueden ser objeto de búsqueda.
2. ☒ Dado que todas las reivindicaciones que pueden ser objeto de búsqueda podrían serlo sin realizar un esfuerzo que justifique tasas adicionales, esta Administración no requirió el pago de tasas adicionales.
3. ☐ Dado que tan sólo una parte de las tasas adicionales requeridas ha sido satisfecha dentro del plazo por el solicitante, el presente informe de búsqueda de tipo internacional comprende solamente aquellas reivindicaciones respecto de las cuales han sido satisfechas las tasas, concretamente las reivindicaciones N°s:
4. ☐ Ninguna de las tasas adicionales requeridas ha sido satisfecha por el solicitante dentro de plazo. En consecuencia, el presente informe de búsqueda de tipo internacional se limita a la invención mencionada en primer término en las reivindicaciones, cubierta por las reivindicaciones N°s:

Indicación en cuanto a la protesta

- ☐ Se acompañó a las tasas adicionales la protesta del solicitante y, en su caso, el pago de una tasa de protesta.
- ☐ Se acompañó a las tasas adicionales la protesta del solicitante, pero la tasa de protesta aplicable no se pagó en el plazo establecido en el requerimiento.
- ☐ El pago de las tasas adicionales no ha sido acompañado de ninguna protesta.

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional N°

PCT/CO201 5/00001 0

C (continuación). DOCUMENTOS CONSIDERADOS RELEVANTES		
Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones N°
X	MCKINNEY, J. et al. Tightly Regulated Gene Expression System in Salmonella entérica Serovar Typhimurium. Journal of Bacteriology. 2002, vol. 184(21):6056-6059. DOI:10.1128/JB.184.21.6056-6059.2002. Todo el documento	18-20
A	LÁMMLE, K. et al. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning, Journal of Biotechnology, 2007, Vol. 127(4): 575-592, ISSN 0168-1656. Todo el documento "Citado en la solicitud"	
A	TERRÓN-GONZALÉZ, L. et al. Heterologous viral expression systems in fosmid vectors increase the functional analysis potential of metagenomic libraries. Scientific Reports 2013, 3: 1107. doi: 10.1038/srep01107 Todo el documento	
A	CULLIGAN E. P. et al. Metagenomics and novel gene discovery: Promise and potential for novel therapeutics. Virulence. 2014;5 (3):399-412. doi: 10.4161/viru.27208. Páginas 402-403	
A	US6159736 A (WISCONSIN ALUMNI RESEARCH FOUNDATION) 12 diciembre 2000 Todo el documento	
A	WO2010049807 A2 (GENE BRIDGES GMBH) 06 mayo 2010 Todo el documento	
A	WO2000078977 A1 (AVENTIS PHARMA INC.) 28 diciembre 2000 Todo el documento	

INFORME DE BÚSQUEDA INTERNACIONAL

Información relativa a miembros de familias de patentes

Solicitud internacional N°

PCT/CO2015/000010

Documento de patente citado en Informe de Búsqueda	Fecha de Publicación	Miembro(s) de Familia	Fecha de Publicación
US6159736 (A)	12-12-2000	AT282693 (T)	15-12-2004
		AU758960 (B2)	03-04-2003
		AU6057399 (A)	10-04-2000
		CA2343000 (C);	30-03-2000
		CA2343000 (A1)	30-03-2000
		CN100420747 (C);	24-10-2001
		CN1319135 (A)	24-10-2001
		DE69922032 (T2)	01-12-2005
		DK1115856 (T3)	21-02-2005
		EP1115856 (A1);	18-07-2001
		EP1115856 (B1)	18-07-2001
		ES2228170 (T3)	01-04-2005
		JP4464563 (B2);	24-09-2002
		JP2002531062 (A)	24-09-2002
		JP2010068813 (A)	02-04-2010
		JP2013165731 (A)	29-08-2013
		PL346772 (A1);	25-02-2002
		PL197408 (B1)	25-02-2002
		RU2237715 (C2)	10-10-2004
		US6294385 (B1)	25-09-2001
		WO0017343 (A1)	30-03-2000
WO2010049807 (A2)	06-05-2010	WO2010049807 (A3)	06-05-2010
		CN102264905 (A)	30-11-2011
WO2000078977 (A1)	28-12-2000	AU5496600 (A)	09-01-2001
		BR0011765 (A)	05-03-2002
		CA2375821 (A1)	28-12-2000
		EP1196613 (A1)	17-04-2002
		JP2003503029 (A)	28-01-2003
		MXPA01012406 (A)	04-06-2002
		NO20016153 (A)	15-02-2002
		NZ515502 (A)	27-08-2004
		US6818441 (B1)	16-11-2004
		ZA200109441 (A)	30-04-2003

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional N°

PCT/CO201 5/000010

(Continuación de Recuadro II)

Las reivindicaciones 26 y 27 se refieren a un vector artificial compuesto de un fósforo esqueleto, particularmente pCC2FOS con un inserto de ADN metagenómico, donde el inserto metagenómico comprende un gen que codifica la proteína fluorescente verde (GFP) con un sitio de unión al ribosoma (RSB) corriente arriba. El vector pCC2FOS es un vector comercial diseñado para generar librerías metagenómicas y un inserto de ADN metagenómico corresponde a una secuencia no determinada a menos que se secuencie. De hecho, lo más probable es que cada muestra metagenómica que se tome del ambiente sea distinta una de otra, por lo que cada librería metagenómica también lo será. En consecuencia, no se puede determinar si un inserto metagenómico tendrá una secuencia que comprende un gen que codifica la GFP y si tiene un RSB. Esta falta de claridad no permite realizar una búsqueda significativa para todo el espectro abarcado en dichas reivindicaciones. En consecuencia, no se ha realizado búsqueda para las reivindicaciones 26 y 27.

(Continuación de Recuadro III)

El pliego de reivindicaciones no tiene un único concepto inventivo que vincule a todas las reivindicaciones. Por una parte se distingue el transposón diseñado para suministrar parcialmente la maquinaria transcripcional del hospedero durante análisis funcionales de librerías genómicas/metagenómicas, el vector que lo comprende, su método de construcción y su uso en el método para aumentar la transcripción de ADN episomal; como también un vector artificial que permite probar la eficacia del transposón en insertos metagenómicos, el cual no tiene ninguna relación estructural con el transposón ya mencionado y, además, resuelve un problema técnico distinto que sería el contar con un vector de prueba para elementos que permitan aumentar la transcripción de genes silenciados en contexto metagenómico. Adicionalmente, se reivindican vectores que corresponderían a aquellos generados durante el diseño del transposón mencionado. Dichos vectores no comprenden todos los elementos que definen al transposón y, por lo tanto, no tienen todos los elementos necesarios para solucionar el problema técnico. Entre dichos vectores no existe una característica común que permita agruparlos bajo un concepto inventivo, ya que lo único que comparten es la presencia de un origen de replicación y un marcador de selección, los cuales son elementos comunes a plásmidos vectores en general.

En conclusión, a priori, se distinguen los siguientes grupos de invenciones:

Grupo 1 (reivindicaciones 1-10, 25 y 28-30): Transposón que suministra parcialmente la maquinaria transcripcional del hospedero durante análisis funcionales de librerías genómicas/metagenómicas, el cual comprende (i) un marcador de selección diferente al vector esqueleto, (ii) gen que codifica T7RNA polimerasa bajo promotor inducible, (iii) dos promotores flanqueantes T7 de lectura hacia el exterior, (iv) dos sitios flanqueantes de reconocimiento de repetición invertida para la transposasa MuA; vector que comprende al transposón; método de construcción del transposón; y método que utiliza al transposón para aumentar transcripción de ADN para expresión de genes foráneos.

Grupo 2 (reivindicaciones 11-13): Secuencia artificial de ADN que comprende un sitio de unión MuA, un promotor T7 en hebras diferentes de ADN y múltiples sitios flanqueantes de reconocimiento para endonucleasas de restricción.

Grupo 3 (reivindicaciones 14-15): Vector artificial que comprende un vector esqueleto con un origen de replicación de alto número de copias, un marcador de selección, una secuencia promotora T7 y un sitio de reconocimiento de la transposasa MuA.

Grupo 4 (reivindicaciones 16-17): Vector artificial que comprende un vector esqueleto con un origen de replicación de alto número de copias, un marcador de selección, una secuencia promotora T7 de lectura hacia el exterior, un sitio de reconocimiento de la transposasa MuA y un marcador de selección diferente de aquel localizado en el vector esqueleto.

Grupo 5 (reivindicaciones 18-20): Vector artificial que comprende un vector esqueleto con un origen de replicación de alto número de copias, un marcador de selección y un gen que codifica la T7RNA polimerasa regulado por un promotor inducible.

Grupo 6 (reivindicaciones 21-24): Vector artificial que comprende un vector esqueleto con un origen de replicación de alto número de copias, un marcador de selección, una secuencia promotora T7 de lectura hacia el exterior, un sitio de reconocimiento de la transposasa MuA, un marcador de selección diferente de aquel localizado en el vector esqueleto y un gen que codifica la T7RNA polimerasa regulado por un promotor inducible.