



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 324 495**

(51) Int. Cl.:

C12N 15/79 (2006.01)

C12N 15/866 (2006.01)

C12N 15/10 (2006.01)

A01K 67/027 (2006.01)

A61K 48/00 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **06818378 .9**

(96) Fecha de presentación : **06.11.2006**

(97) Número de publicación de la solicitud: **1945773**

(97) Fecha de publicación de la solicitud: **23.07.2008**

(54) Título: **Expresión recombinante de complejo de multiproteínas utilizando poligenes.**

(30) Prioridad: **08.11.2005 EP 05024271**

(73) Titular/es: **ETH Zürich**
Raemistrasse 101/Eth Transfer
8092 Zürich, CH

(45) Fecha de publicación de la mención BOPI:
07.08.2009

(72) Inventor/es: **Richmond, Timothy y**
Berger, Imre

(45) Fecha de la publicación del folleto de la patente:
07.08.2009

(74) Agente: **Isern Jara, Jorge**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Expresión recombinante de complejo de multiproteínas utilizando poligenes.

La presente invención hace referencia a un polinucleótido recombinante que codifica, cada uno de ellos dentro de un marco de lectura abierto individual (ORF), al menos dos poligenes que codifican, cada uno de ellos, al menos tres polipéptidos biológicamente activos en los que al menos dos de los polipéptidos codificados por los genes que forman parte de los poligenes son de origen no viral, al menos dos de los polipéptidos codificados por los genes que forman parte de los poligenes son capaces, cada uno de ellos, de al menos interaccionar transitoriamente con al menos otro de los polipéptidos codificados por un gen de dichos poligenes, y los genes que forman parte de cada poligen están conectados entre sí mediante un secuencia que codifica al menos un sitio de escisión por proteasas y/o por una secuencia que codifica al menos un péptido autoescindible. Son realizaciones adicionales de la presente invención un vector que contiene el polinucleótido recombinante, una célula huésped que contiene el polinucleótido recombinante y/o el vector y un animal transgénico no humano transformado con el polinucleótido recombinante y/o el vector. La presente invención también hace referencia a métodos para la producción del polinucleótido y para la fabricación de complejos de multiproteínas. Las realizaciones de la presente invención son particularmente útiles en la terapia génica, cribado de candidatos a fármaco, producción de vacunas y cristalización de complejos de multiproteínas para investigación estructural.

Un objetivo fundamental de los esfuerzos en investigación biológica en la era post-genómica es la elucidación de las redes de interacción de proteínas (interactoma). Dado que muchos de los complejos de multiproteínas identificados no se encuentran presentes en cantidades suficientes en sus células originales para poder realizar un análisis de biología molecular detallado, su estudio depende de las tecnologías recombinantes de producción de proteínas heterólogas a gran escala. Los métodos de expresión recombinante requieren una inversión desproporcionada tanto en horas de trabajo como en materiales antes de la expresión de las multiproteínas, y después de su expresión no proporcionan flexibilidad para alterar con rapidez los componentes de las multiproteínas para la realización de estudios de expresión revisados.

Existen diversas tecnologías recombinantes utilizadas actualmente para obtener complejos de multisubunidades. Por ejemplo, las proteínas se pueden expresar aisladamente en *E. coli*, ya sea en forma soluble o en forma de cuerpos de inclusión, purificar y a continuación reconstituir *in vitro* con proteínas producidas de forma similar formando complejos de multiproteínas. También pueden utilizarse células eucariotas (por ejemplo, células de mamífero o levaduras) como huéspedes en experimentos de expresión transitoria. Esta metodología es completamente dependiente de la existencia de un protocolo de reconstitución *in vitro* eficiente. Mientras que esta estrategia puede rendir resultados aceptables para sistemas más simples con tamaños de subunidad pequeños, no es generalmente aplicable para complejos de multiproteínas más complicados que contengan varias subunidades grandes (por ejemplo próximas a los complejos reguladores de las células eucariotas superiores, en particular humanas).

La co-expresión ha sido reconocida como una alternativa superior a la estrategia de la reconstitución *in vitro* descrita en el párrafo anterior. En el pasado, se han desarrollado diversos sistemas de co-expresión tanto para la expresión procariota como la eucariota. En los sistemas procariota, la co-expresión puede conseguirse generando un único plásmido que contenga todos los genes seleccionados o co-transformando diversos plásmidos que contengan uno o dos genes y diferentes marcadores de resistencia y replicones.

La co-expresión en células eucariotas se ha realizado utilizando el sistema baculovirus, inicialmente con un éxito limitado mediante la co-infección con diversos virus, y posteriormente de forma más exitosa mediante la expresión de todas las proteínas a partir de un único virus, ofreciendo diversas ventajas y eliminando diversas limitaciones existentes en los sistemas procariota (tales como tamaños de subunidad comparativamente menores, ausencia de procesamiento auténtico, dificultad de expresión de proteínas eucariotas (especialmente humanas) etc.). En el sistema baculovirus, se ha demostrado que la expresión a partir de un único virus aumenta dramáticamente el rendimiento (Berger y otros (2004) *Nature Biotech.* 22, 1583-1587; ver también Comentario ("Comment") (2004), *Nature Biotech.* 22, vii, Novedades y Opiniones ("New & Views") (2004), *Nature Biotech.* 22, 152, Aspectos destacados en investigación ("Research Highlights"), *Nature Methods* 2, 7 (2005); Bertolotti-Ciarlet y otros (2003) *Vaccine* 21, 3885-3900), a la vez que reduce de forma decisiva las necesidades logísticas especialmente para la producción a gran escala.

Una mejora importante en la expresión de multiproteínas fue la provisión del sistema modular para la generación de casetes de expresión de multigenes proporcionados por los presentes inventores, que se da a conocer en WO 2005/085456 A1 (PCT/EP2004/013381; ver también Berger y otros (2004) *ibídem*). La tecnología MultiBac descrita en WO 2005/085456 A1 (PCT/EP2004/013381) permite la generación sencilla de casetes de expresión de multigenes así como la modificación y revisión de los experimentos de expresión (Berger y otros (2004) *ibídem*).

Sin embargo, un obstáculo para la expresión exitosa y el acoplamiento *in vivo* de complejos de multisubunidades, en particular con varias (6, 7, 8 o más) subunidades (que constituyen la mayoría de los complejos de regulación génica eucariotas, por ejemplo humanos) es el hecho de que los niveles de expresión relativa de estas subunidades varían singularmente de forma significativa en base a en muchos casos mecanismos no totalmente conocidos (por ejemplo eficacia de la transcripción y traducción, estabilidad de las proteínas, estabilidad del ARNm y estructuras secundarias, etc.). Como consecuencia de ello, la subunidad que se expresa en menor cantidad en un sistema intrínsecamente desequilibrado condicionará el éxito global del experimento de producción de complejos de multisubunidades al limitar el

rendimiento total del complejo. Por ello, la maquinaria de transcripción/traducción produce cantidades en exceso de otros componentes que no son incorporados en proceso, “malgastando” recursos de transcripción/traducción celular. Los niveles individuales de expresión varían singularmente varias veces (por ejemplo hasta 10 veces o más) de unos a otros suponiendo pérdidas que son refractarias a una producción exitosa de los complejos de multisubunidades deseados, en particular en el caso de complejos de más de 4 subunidades, tales como 5, 6, 8, 10 o más subunidades (por ejemplo en el caso de muchos complejos de regulación génica eucariotas).

Los virus del súper-grupo picornavirus poseen un genoma que consta de una molécula de ARN de cadena única en orientación sentido que contiene uno o dos ORF (“marcos de lectura abiertos”) que codifican una poliproteína que comprende las proteínas virales que están conectadas entre sí mediante sitios de escisión por proteasas virales o por péptidos autoescindibles (revisado, por ejemplo en Ryan y otros (1997) *J. Gener. Virol.* 78, 699-723).

El concepto general de expresar una poliproteína a través de un virus recombinante para la producción de complejos proteína se ha aplicado en la reconstitución de un complejo TCR (receptor de célula T):CD3 (Szymczak y otros (2004) *Nature Biotech.* 5, 589-594). Los autores utilizaron dos vectores retrovirales recombinantes en los que un vector contenía las secuencias que codifican las dos subunidades del TCR mientras que el otro vector codificaba una poliproteína que comprendía las cuatro subunidades de CD3. Las subunidades se conectaron mediante secuencias de péptidos de autoescisión 2A derivados de aftovirus. Una desventaja de este enfoque es que, para reconstituir el complejo completo, deben prepararse dos vectores diferentes y son necesarias dos transfecciones.

El enfoque de la poliproteína viral se ha aplicado en sistemas de expresión baculovirus para constructos pequeños tales como IL-12 heterodimérica (Kokuho y otros (1999) *Vet. Immunol. Immunopathol.* 72, 289-302) y proteínas de fusión que contienen una señal de focalización nuclear derivada de la polihedrina de baculovirus y una proteína de interés (US 5.179.007).

Por tanto, el problema técnico que subyace en la presente invención es el de dar a conocer un nuevo sistema para una expresión mejorada de proteínas múltiples.

La solución del problema técnico especificado en el párrafo anterior se proporciona por las realizaciones definidas en las reivindicaciones.

En particular, la presente invención da a conocer un polinucleótido que codifica, cada uno de ellos dentro de un marco de lectura abierto individual (ORF), al menos dos poligenes que codifican cada uno de ellos al menos tres polipéptidos biológicamente activos en los que al menos dos de los polipéptidos codificados por los genes que forman parte de los poligenes son de origen no viral, al menos dos de los polipéptidos codificados por los genes que forman parte de los poligenes son capaces cada uno de ellos de al menos interaccionar transitoriamente con al menos otro de los polipéptidos codificados por un gen de dichos poligenes, y los genes que forman parte de cada poligen están conectados entre sí mediante una secuencia que codifica al menos un sitio de escisión por proteasas y/o por una secuencia que codifica al menos un péptido autoescindible.

El polinucleótido de acuerdo con la presente invención puede ser un polinucleótido ADN, ARN o un polinucleótido que comprende uno o más análogos de nucleótido sintético. El polinucleótido puede estar presente en forma de cadena única o doble. Son especialmente preferentes las formas ADN, en particular ADN de doble cadena. El polinucleótido de la presente invención puede producirse mediante síntesis química. Los constructos de polinucleótido preferentes de la presente invención se producen mediante tecnología génica recombinante (ver, por ejemplo, Sambrook y otros “Clonación molecular” (“Molecular Cloning”), Cold Spring Harbor Laboratory, 1989).

Un “poligen” tal como se utiliza en la presente invención es una secuencia de ácidos nucleicos que codifica al menos tres polipéptidos biológicamente activos en un ORF único. Por tanto, cada “gen” que forma parte del poligen es una secuencia de ácidos nucleicos que codifica un polipéptido, en particular una proteína o fragmento, variante, mutante o análogo de la misma, que posee una función estructural, reguladora o enzimática en particular. Preferentemente, el “gen” que codifica el polipéptido comprende la región codificadora de un ADNc que codifica la proteína estructural, reguladora o enzimática, o fragmento, variante, mutante o análogo de la misma.

Un “fragmento” del polipéptido codificado por un gen contenido en los poligenes significa una parte o región del polipéptido original, preferentemente un fragmento que conserva al menos una de las funciones de la proteína completa. Una “variante” del polipéptido codificado por un gen contenido en el poligen significa un polipéptido que es un equivalente funcional o no funcional del polipéptido original derivado de otra especie o un derivado funcional o no funcional del polipéptido original que se origina por un empalme alternativo o procesamiento post-traduccion. Un “mutante” del polipéptido codificado por un gen contenido en el poligen significa un polipéptido derivado de una proteína natural por inserción, substitución, adición y/o delección de uno o más residuos aminoácidos. Un “análogo” del polipéptido codificado por un gen contenido en el poligen significa un equivalente funcional del polipéptido original que puede incluso tener una secuencia de aminoácidos no relacionada pero que ejerce la misma función que el polipéptido del cual es un análogo.

Correspondientemente, a nivel de ácidos nucleicos, un “fragmento” de gen es una parte o región del gen original a partir del cual se deriva el “fragmento”. El gen “variante” posee una secuencia que se halla en una especie diferente comparada con el gen original, o puede codificar una variante de empalme o una versión procesada a nivel post-

traduccional del polipéptido en cuestión. El “mutante” es derivado a partir del gen original mediante inserción, sustitución, adición y/o delección de uno o más nucleótidos. El “análogo” de un gen codifica un equivalente funcional del polipéptido codificado por el gen original.

Al menos dos de los genes en los poligenes de acuerdo con la presente invención son de origen no viral. “No viral” significa que la secuencia de ácido nucleico que codifica el polipéptido (que representa una proteína funcional o un fragmento, variante, análogo o mutante de la misma) no se halla originalmente o no se deriva del genoma de un virus. En particular, las secuencias de nucleótidos comprendidas en el poligen del polinucleótido de acuerdo con la presente invención se originan de eucariotas y/o procariotas.

Así, de acuerdo con la presente invención, los genes que codifican las subunidades (tales como un complejo de multiproteínas o miembros de una vía metabólica o cualesquiera otras proteínas que al menos potencialmente interactúan al menos transitoriamente entre sí) de un ensamblaje de multisubunidades se encuentran presentes en al menos dos marcos de lectura abiertos (ORF). Las secuencias que codifican las subunidades (polipéptidos) del ensamblaje se encuentran presentes en al menos dos poligenes en los que los genes que forman parte de cada poligen están conectados entre sí mediante una secuencia (puede haber más de una) que codifica un sitio de escisión por proteasas (es decir una secuencia que comprende el sitio de reconocimiento de una proteasa) o al menos un péptido autoescindible.

De acuerdo con una realización preferente de la presente invención la proteasa o proteasas capaces de escindir el sitio de escisión codificado por la secuencia o secuencias que conectan los genes que forman parte de los poligenes son codificadas por el polinucleótido de la presente invención. De forma más preferente, el gen o genes que codifican la proteasa o proteasas forma o forman parte de la menos uno de los poligenes.

Los sitios de escisión por proteasas y péptidos autoescindibles son conocidos para los expertos (ver por ejemplo, Ryan y otros (1997) *J. Gener. Virol.* 78, 699-722; Scymczak y otros (2004) *Nature Biotech.* 5, 589-594). Son ejemplos preferentes de sitios de escisión por proteasas, los sitios de escisión por las proteasas NIa de potivirus (por ejemplo, la proteasa del virus del grabado del tabaco), proteasas HC de potivirus, proteasas P1 (P35) de potivirus, proteasas NIa de biovirus, proteasas codificadas ARN-2 de biovirus, proteasas L de aftovirus, proteasas 2A de enterovirus, proteasas 2A de rinovirus, proteasas 3C de picorna, proteasas 24K de comovirus, proteasas 24K de nepovirus, proteasa similar a 3C de RTSV (virus esférico del tungro del arroz), proteasa similar a 3C de PYVF (virus de la mancha amarilla de la chirimía), trombina, factor Xa y enteroquinasa. Debido a su elevado rigor en la escisión, son particularmente preferentes los sitios de escisión de las proteasas del TEV (virus del grabado del tabaco). Así, los genes de los poligenes de acuerdo con la presente invención están preferentemente conectados mediante una cadena de nucleótidos que comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con la forma general EXXYXQ(G/S) en la que X representa cualquier aminoácido (la escisión por TEV se produce entre Q y G o Q y S). Son más preferentes las secuencias de nucleótidos de enlace que codifican ENLYFQG y ENLYFQS, respectivamente.

Los péptidos autoescindibles (también denominados “elementos hidrolíticos que actúan en cis”, CHYSEL; ver deFelipe (2002) *Curr. Gene Ther.* 2, 355-378) preferentes derivan de péptidos 2A de potivirus y cardiovirus. Son especialmente preferentes los péptidos autoescindibles seleccionados entre los péptidos 2A derivados del FMDV (virus de la enfermedad boca-mano), virus de la rinitis equina A, virus de Thosa asigna, y el teschovirus porcino.

Al menos dos de los polipéptidos codificados por los poligenes de la presente invención son capaces de al menos interactuar transitoriamente con otro polipéptido codificado por los poligenes, o al menos se sospecha que son capaces de al menos interactuar transitoriamente con otro polipéptido codificado por un gen contenido en los poligenes. Las “interacciones” típicas que se producen entre los polipéptidos incluyen unión covalente, enlaces de hidrógeno, interacciones electrostáticas e interacciones de Van-der-Waals. Las interacciones “transitorias” son habituales entre las biomoléculas, en particular entre las proteínas, y están representadas típicamente por interacciones entre enzimas y sus sustratos, receptores y sus ligandos (agonistas o antagonistas), interacciones entre miembros de vías metabólicas e interacciones entre proteínas de complejos reguladores (por ejemplo, reguladores de genes).

Los polipéptidos codificados por las secuencias de nucleótidos que forman parte de los poligenes de la presente invención pueden ser el mismo o diferentes. Así, cada poligen presente en los constructos de la invención puede contener una o más copias de cada secuencia de nucleótidos que codifica una proteína de interés. De este modo es por ejemplo posible proporcionar constructos que sirven para la expresión óptima de las proteínas deseadas, en particular en el caso de proteínas que se expresan normalmente a diferentes niveles de expresión y/o están presentes en un ensamblaje macromolecular en diferentes estequiometrías. Por tanto, en el caso de que un polipéptido se exprese escasamente en los sistemas utilizados habitualmente, pueden integrarse dos o más copias de la secuencia de codificación correspondiente en uno o más poligenes del constructo de la invención. También puede utilizarse el mismo enfoque en caso de que el polipéptido esté presente en un dímero, trímero o multímero en un complejo deseado. De este modo, los constructos de la presente invención pueden acoplarse individualmente de acuerdo con los requisitos (niveles de expresión, estequiometría, etc.) de cualquier complejo u otro ensamblaje macromolecular que un experto en la técnica quiera expresar y/o purificar.

Es adicionalmente preferente que los genes que forman parte de los poligenes se seleccionen del grupo formado por genes que codifican miembros de complejos de multiproteínas y genes que codifican miembros de vías metabólicas. Los complejos de multiproteínas preferentes son complejos de proteínas reguladoras de genes tales como complejos de factores de transcripción, complejos de transporte tales como complejos involucrados en el transpor-

te nuclear y/o celular, complejos de plegamiento proteico, complejos receptor/ligando, complejos de reconocimiento célula-célula, complejos involucrados en la apoptosis, complejos involucrados en la regulación del ciclo celular, etc. Son miembros de vías metabólicas, por ejemplo, miembros del metabolismo de carbohidratos (como glucólisis, gluconeogénesis, ciclo del ácido cítrico, biosíntesis de glucógeno, vía de la galactosa, ciclo de Calvin, etc.), metabolismo de lípidos (como metabolismo del triacilglicerol, activación de ácidos grasos, β -oxidación de ácidos grasos (incluso cadena/cadena impar) vía de la α -oxidación, biosíntesis de ácidos grasos, biosíntesis de colesterol, etc.), metabolismo de aminoácidos (como reacciones de glutamato, ciclo de la urea de Krebs-Henseleit, vía de shikimato, biosíntesis de fenilalanina y tirosina, biosíntesis de triptófano, etc.), metabolismo energético (como fosforilación oxidativa, síntesis de ATP, fotosíntesis, metabolismo del metano, etc.) metabolismo de ácidos nucleicos (biosíntesis y degradación de purinas y pirimidinas, replicación de ADN, etc.). Los miembros de los complejos de multiproteínas y de la vía metabólica pueden tomarse de por ejemplo, <http://www.biocarta.com/genes/index.asp> y G. Michal (ed.) Vías bioquímicas (Biochemical Pathways) 1ª edición, John Wiley & Sons, Hoboken, NJ, USA, 1999, cuyo contenido se incorpora a la presente invención como referencia.

Cada poligen de acuerdo con la presente invención contiene al menos 3 genes, es decir secuencias que codifican un polipéptido biológicamente activo. Son más preferentes los poligenes que codifican 4, 5, 6 o más o incluso más proteínas. Tal como se ha mencionado anteriormente, es preferente que la proteasa o proteasas capaces de escindir los sitios de escisión por proteasas que conectan los polipéptidos esté o estén codificados por al menos uno de los poligenes.

De acuerdo con una realización preferente, el polinucleótido de la presente invención contiene al menos dos secuencias promotor que están operativamente unidas a los poligenes, siendo así capaces de controlar la expresión de los mismos. Los promotores adecuados en los constructos de la presente invención pueden seleccionarse del grupo formado por promotores de baculovirus muy tardíos polh, p10 y pXIV, promotor tardío de baculovirus vp39, promotor de baculovirus tardío/muy tardío vp30polh, promotores tempranos de baculovirus P_{cap/poth}, *pcna*, *etl*, *p35*, *da26*; CMV, SV40, UbC, EF-1 α , RSVLTR, MT, P_{DS47}, Ac5, P_{GAL} y P_{ADH}. Las secuencias de promotores pueden ser las mismas para todos los poligenes o pueden seleccionarse diferentes promotores para diferentes poligenes.

Preferentemente, cada uno de los ORF que contienen un poligen de la presente invención está flanqueado por una secuencia de finalización tal como SV40, HSVtk o BGH (hormona de crecimiento bovina).

El polinucleótido, de acuerdo con la presente invención, puede contener además secuencias reguladoras tales como secuencias incrementadoras o supresoras.

Es adicionalmente preferente que el polinucleótido de acuerdo con la presente invención contenga al menos un sitio para su integración en un vector o célula huésped. Dicho sitio de integración permitirá la conveniente incorporación genómica o transitoria del polinucleótido en vectores (tales como virus) y células huésped (por ejemplo células huésped eucariotas), respectivamente. Son preferentes los sitios para la integración genómica.

Son sitios de integración especialmente preferentes aquellos que son compatibles con la integración del polinucleótido en un virus. Más preferentemente, el sitio de integración es compatible con la integración del polinucleótido en un virus seleccionado dentro del grupo formado por adenovirus, virus asociados a adenovirus (AAV), parvovirus autónomos, virus del herpes simple (HSV), retrovirus, rhadinovirus, virus Epstein-Barr, lentivirus, virus Semliki Forest (virus del bosque Semliki) y baculovirus.

En una realización adicionalmente preferente, el sitio de integración es compatible con la integración del polinucleótido en una célula huésped eucariota que puede seleccionarse preferentemente del grupo formado por células de mamífero (como células humanas, por ejemplo HeLa, Huh7, HEK293, HepG2, KATO-III, IMR32, MT-2, células β pancreáticas, queratinocitos, fibroblastos de médula ósea, CHP212, células neurales primarias, W12, SK-N-MC, Saos-2, WI38, hepatocitos primarios, FLC3, 143TK-, DLD-1, células de vena umbilical, fibroblastos de pulmón embrionario, fibroblastos del prepucio primarios, células de osteosarcoma, células MRC5, células MG63 etc.), células porcinas (como CPL, FS-13, PK-15), células bovinas (como MDB, BT), células ovinas (como FLL-YFT), células de *C. elegans*, células de levaduras (como *S. cerevisiae*, *S. pombe*, *C. albicans*, *P. pastoris*) y células de insecto (como *S. frugiperda*, por ejemplo Sf9, Sf21, Express Sf+, células High Five H5, *D. melanogaster*, por ejemplo células S2 Schneider).

Los sitios de integración particularmente preferentes se seleccionan a partir de elementos transposón de Tn7, sitios de unión específicos de integrasa λ , y SSRs (recombinasas específicas de sitio), preferentemente el sitio específico cre-lox (LoxP) o el sitio de recombinación específico de la recombinasa FLP (FRT).

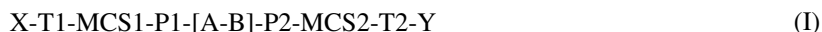
En una realización preferente de la presente invención el polinucleótido comprende adicionalmente uno o más marcadores de resistencia para seleccionar células huésped con propiedades deseadas basadas en una resistencia a sustancias de otro modo tóxicas. Son ejemplos de marcadores de resistencia adecuados los que proporcionan resistencia contra la ampicilina, cloramfenicol, gentamicina, espectinomycin y/o kanamicina.

Para su incorporación a una célula huésped procariota, el nucleótido de la presente invención comprende preferentemente un origen de replicación R6K γ condicional para hacer que la propagación sea dependiente del gen *pir* en un huésped procariota.

ES 2 324 495 T3

La inserción de los poligenes casetes de expresión en constructos dados a conocer en WO 2005/085456 A1 (PCT/EP2004/013381) da lugar a realizaciones especialmente preferentes del polinucleótido de la presente invención.

Por tanto, es preferente que el polinucleótido de la presente invención comprenda una disposición funcional según la Fórmula I:



que comprende:

(a) al menos dos casetes de expresión T1-MCS1-P1 y P2-MCS1-T2 en una disposición cabeza a cabeza, cabeza a cola o cola a cola, comprendiendo cada uno de ellos un sitio de clonación múltiple MCS1 o MCS2, flanqueado por un promotor P1 y un secuencia de finalización T1 para MCS1 y flanqueado por un promotor P2 y una secuencia de finalización T2 para MCS2.

(b) al menos un módulo de multiplicación M entre los promotores P1 y P2 que comprende al menos dos sitios de restricción A y B.

(c) al menos dos sitios de restricción X e Y flanqueando cada uno de ellos los casetes de expresión, en el que

- (i) los sitios de restricción A y X así como B e Y son compatibles, pero
- (ii) los productos de unión AY y BX no son escindibles enzimáticamente por enzimas de restricción a, b, x o y específicos de los sitios de restricción A, B, X e Y, y
- (iii) los sitios de restricción A y B así como los sitios de restricción X e Y son incompatibles,

en el que cada poligen está insertado en uno de los casetes de expresión.

Respecto a realizaciones adicionales preferentes de constructos que poseen una disposición de Fórmula (I) se hace referencia expresamente a WO 2005/085456 A1 (PCT/EP2004/013381). En particular, los sitios de restricción A y B en el módulo de multiplicación M se seleccionan del grupo formado por los sitios de restricción BstZ17I, SpeI, ClaI y NruI o sitios de restricción escindidos por isosquizómeros de los mismos. Los isosquizómeros son enzimas de restricción que poseen sitios de escisión idénticos. Adicionalmente, son ejemplos de sitios de restricción X e Y los sitios de restricción seleccionados del grupo formado por PmeI y AvrII o sitios de restricción escindidos por isosquizómeros de los mismos.

Los polinucleótidos preferentes que posee cada uno de los poligenes insertado en uno de los casetes de expresión contenidos en la Fórmula (I) comprenden las siguientes características:

- (a) los promotores P1 y P2 se seleccionan del grupo formado por polh y p10;
- (b) las secuencias de finalización se seleccionan del grupo formado por SV40 y HSVtk;
- (c) los sitios de restricción A y B en el módulo de multiplicación M se seleccionan del grupo formado por los sitios de restricción BstZ171I, SpeI, ClaI y NruI;
- (d) los sitios de restricción X e Y se seleccionan del grupo formado por los sitios de restricción PmeI y AvrII; y
- (e) los sitios para la integración de virus se seleccionan del grupo formado por un sitio para recombinasas específicas de sitio (SSRs) y un sitio para un elemento de transporte.

Respecto a la producción de polinucleótidos que poseen la disposición anterior según la Fórmula (I) se hace referencia expresa a WO 2005/085456 A1 (PCT/EP2004/013381).

La disposición de los polinucleótidos de la presente invención que codifican varios polipéptidos activos biológicamente dentro de dos o más ORFs conteniendo cada uno de ellos un poligen supone una mejora importante respecto a la clonación y expresión de genes que codifican miembros de complejos proteicos de multisubunidades: Por un lado, el ensamblaje de todos los genes de subunidades en un único ORF es habitualmente imposible o altamente dificultoso debido al inmenso tamaño y número de secuencias de codificación a acoplar. Por otro lado, el ensamblaje eficiente de varios o todos los miembros de un ensamblaje de multisubunidades cada uno de ellos presente en casetes de expresión independientes ha mostrado ser habitualmente altamente ineficiente dado que el rendimiento global del complejo viene determinado por la subunidad que se expresa menos. De acuerdo con la presente invención, las subunidades de un complejo de multiproteínas son codificadas por al menos dos poligenes (cada uno representando un ORF individual) codificando cada uno de ellos al menos tres polipéptidos (preferentemente de origen no viral) lo que da lugar a un

compromiso óptimo entre la manejabilidad del constructo y sus constituyentes (en particular, el ensamblaje de los poligenes) y la eficiencia de la expresión (en particular en el caso de que el polinucleótido de la presente invención se encuentre en un vector adecuado).

5 Por tanto, una realización adicional de la presente invención es un vector que contiene el polinucleótido descrito anteriormente. El vector puede seleccionarse del grupo formado por plásmidos, vectores de expresión y vectores de transferencia. De forma más preferente, el vector de la presente invención es útil para la transferencia de genes eucariotas, para la transferencia génica transitoria o viral mediada por vectores.

10 Son vectores especialmente preferentes los vectores de expresión eucariotas, tales como virus seleccionados del grupo formado por adenovirus, virus asociados a adenovirus (AAV), parvovirus autónomos, virus del herpes simple (HSV), retrovirus, rhadinovirus, virus Epstein-Barr, lentivirus, virus Semliki Forest (virus del bosque Semliki) y baculovirus. Los vectores más preferentes de la presente invención son los vectores de expresión baculovirus. Los baculovirus preferentes de la presente invención son las realizaciones en los que los genes *v-cath* y *chiA* se hallan
15 funcionalmente alterados, dado que ello conlleva un mantenimiento mejorado de los compartimentos celulares durante la infección y expresión proteica. El gen *v-cath* codifica la proteasa viral V-CATH la cual es activada en el momento de la muerte celular mediante un proceso dependiente de un gen yuxtapuesto en el ADN viral, *chia* que codifica una quitinasa. Preferentemente, se alteran ambos genes para eliminar la actividad V-CATH y ganar la opción de utilizar la cromatografía de afinidad a quitina sin interferencia por el producto del gen *chiA*. La calidad de los
20 productos de expresión generados por un sistema baculovirus que carece de la actividad funcional de los genes *v-cath* y *chiA* mejora significativamente debido a la reducción de la actividad proteolítica y lisis celular dependiente de virus.

Preferentemente, los vectores de acuerdo con la presente invención comprende un sitio para SSRs, preferentemente
25 LoxP para la recombinación específica de sitio cre-lox. De forma más preferente, el sitio cre-lox se ubica en uno o ambos genes de baculovirus *v-cath* y *chiA* de manera que alteran su función.

El vector de la presente invención contiene preferentemente uno o más genes marcadores para la selección de huéspedes transfectados con éxito con el vector ensamblado correctamente. Son ejemplos de marcadores adecuados
30 luciferasa, β -Gal, CAT, genes que codifican proteínas fluorescentes como GFP, BFP, YFP, CFP y variantes de las mismas, y el gen *lacZ α* . El gen o genes marcadores pueden ser variantes equivalentes funcionalmente, mutantes, fragmentos o análogos de los ejemplos mencionados o cualquier otro marcador adecuado conocido por los expertos. Las variantes, mutantes o análogos muestran preferentemente una homología de al menos el 75%, más preferentemente el 85%, y especialmente preferente el 90%, en particular al menos el 95% a nivel de aminoácidos en comparación con
35 el marcador del que deriva dicha variante, mutante o análogo.

En otra realización preferente, el vector de la presente invención comprende un elemento transposón, preferentemente el sitio de unión Tn7. Más preferentemente, dicho elemento transposón, por ejemplo el sitio de unión Tn7, se
40 ubica dentro de un gen marcador de manera que puede evaluarse una integración exitosa por transposición analizando el fenotipo proporcionado por el gen marcador funcional.

Los vectores de transferencia preferentes de la presente invención se basan en pFBDM o pUCDM tal como se da a conocer en la patente WO 2005/085456 A1 (PCT/EP2004/13381; ver identificador de secuencia nº 1 y 2 así como las Figuras 1 y 2, respectivamente dadas a conocer en la misma). Vectores de transferencia de la presente invención
45 adicionalmente preferentes se basan en derivados de los vectores anteriores pFBDM y pUCDM, respectivamente.

Son ejemplos de derivados particularmente preferentes de pFBDM y pUCDM los vectores de transferencia pSPL (Figura 3), pFL (Figura 4), pKL (Figura 5) y pKDM (Figura 6). Al igual que pUCDM, pSPL contiene un origen de replicación condicional (R6K γ) y pFL (como pFBDM) contiene un origen de replicación de elevado número de
50 copias (ColE1). pKDM y pKL poseen orígenes de replicación de bajo número de copias derivados de pBR322. En analogía con pFBDM, pFL pKL y pKDM contienen elementos transposón (Tn7R, Tn7L). Los vectores pSPL, pFL y pKL poseen un repetido LoxP invertido imperfecto que flanquea el casete de expresión dual (al igual que pUCDM). Todos los vectores contienen el módulo de multiplicación (M) descrito anteriormente para generar casetes1 multigen. pFL y pKL (y derivados) son vectores aceptores, pUCDM y pSPL (y derivados) son vectores donantes en plásmidos
55 de fusión mediados por Cre.

60

65

ES 2 324 495 T3

En la siguiente Tabla 1, se resumen las características importantes de los ejemplos preferentes anteriores de vectores de transferencia para generar los constructos de la presente invención.

TABLA 1

Características de los vectores de transferencia preferentes

Vector	Marcador de resistencia antibiótica	Replicón (origen)	Cepa huésped	Elementos de recombinación y multiplicación	Utilización
pFBDM	Ampicilina Gentamicina	ColE1	TOP10*	Tn7L, Tn7R, Módulo multiplicación M	integración en sitio MultiBac ** Tn7
pFL	Ampicilina Gentamicina	ColE1	TOP10	Tn7L, Tn7R, LoxP Módulo multiplicación M	Aceptor para fusiones de plásmidos; integración en sitio MultiBac ** Tn7
pKDM	Kanamicina Gentamicina	pBR322	TOP10	Tn7L, Tn7R, Módulo multiplicación M	integración en sitio MultiBac ** Tn7
pKL	Kanamicina Gentamicina	pBR322	TOP10	Tn7L, Tn7R, LoxP Módulo multiplicación M	Aceptor para fusiones de plásmidos; integración en sitio MultiBac ** Tn7
pUCDM	Cloramfenicol	R6ky	BW23473	LoxP, Módulo multiplicación M	Donante para fusiones de plásmidos; integración en sitio MultiBac ** LoxP
pSPL	Espectinomicina	R6ky	BW23473	LoxP, Módulo multiplicación M	Donante para fusiones de plásmidos; integración en sitio MultiBac ** LoxP
* o cualquier otra cepa de clonación de laboratorio general (<i>recA</i> ⁻ , <i>endA</i> ⁻ , <i>pir</i> ⁻)					
* ver WO 2005/085456 A1					

Por tanto, los poligenes de la presente invención se insertan en un vector tal como pFBDM, pUCDM, pSPL, pFL, pKL o pKDM en los sitios de clonación múltiple (MCS1, MCS2), ya sea mediante escisión por enzimas de restricción y ligadura o a través de recombinación (por ejemplo, utilizando la enzima BD In-Fusion). Los vectores de transferencia de baculovirus pFBDM, pUCDM, pSPL, pFL, pKL y pKDM, comprenden ADN de baculovirus receptor modificado, diseñado para una producción de proteínas mejorada y para permitir un método simple y rápido para integrar genes a través de dos sitios de acceso (attTn7 y LoxP) en el ADN de baculovirus en células de *E. coli* adaptadas a este objetivo.

De acuerdo con una realización adicional, la presente invención da a conocer una célula huésped que contiene el polinucleótido y/o el vector de la invención.

Son ejemplos de células huésped preferentes las células de mamífero, tales con células humanas, de roedor, porcinas tales como CPL, FS-13, y PK-15, células bovinas tales como MDB y BT, células ovinas tales como FLL-YFT, células de *C. elegans*, células de levaduras tales como *S. cerevisiae*, *S. pombe*, *P. pastoris* y *C. albicans*, células de insecto tales como células derivadas de *S. frugiperda*, preferentemente Sf9, Sf21, Express Sf+ o células High Five, h5, células derivadas de *D. melanogaster* tales como células Schneider S2 y bacterias tales como *E. coli*, preferentemente las cepas Top10, Dh5 α , DH10 α , HB101, TG1, BW23473 y BW23474.

Las células humanas preferentes se seleccionan del grupo formado por HeLa, Huh7, HEK293, HepG2, KATO-III, IMR32, MT-2, células β pancreáticas, queratinocitos, fibroblastos de médula ósea, CHP212, células neurales primarias, W12, SK-N-MC, Saos-2, WI38, hepatocitos primarios, FLC3, 143TK-, DLD-1, células de vena umbilical, fibroblastos de pulmón embrionario, fibroblastos del prepucio primarios, células de osteosarcoma, células MRC5 y células MG63.

Las células huésped que comprenden un polinucleótido y/o vector de acuerdo con la invención pueden ser células aisladas o pueden encontrarse en tejidos u órganos.

Una realización adicional de la presente invención hace referencia a un animal transgénico no humano transformado con al menos una secuencia de polinucleótido y/o vector de la invención. Son animales transgénicos no humanos preferentes las especies de roedores, porcinas, bovinas, y *C. elegans*.

El animal transgénico no humano de la presente invención es particularmente útil para elucidar el papel de los complejos de multiproteínas o para el cribado de compuestos para sus actividades biológicas *in vivo*.

Una realización adicional de la presente invención es un método para la producción del polinucleótido tal como se ha definido anteriormente que comprende las etapas de:

(a) proporcionar, preferentemente ampliar, las regiones de codificación de los genes que forman parte de al menos dos poligenes;

(b) proporcionar dichas regiones de codificación con secuencias que codifican al menos un sitio de escisión y/o al menos un péptido autoescindible; y

(c) ensamblar los fragmentos resultantes de las etapas (a) y (b) de manera que de cada poligen se origina un ORF individual; y

(d) combinar los al menos dos poligenes en un único polinucleótido.

Otro aspecto de la invención es un método para la producción del vector de acuerdo con la presente invención que comprende las etapas de:

(a) generar al menos dos poligenes comprendiendo cada uno de ellos al menos tres genes dentro de un ORF individual tal como se ha definido anteriormente (preferentemente mediante el método anterior para la producción del polinucleótido de la presente invención); y

(b) clonar los poligenes en un vector plásmido o viral,

en el que al menos uno de los genes de cada poligen es de origen no viral y al menos dos de los polipéptidos codificados por los genes son capaces de al menos interactuar transitoriamente con otro polipéptido codificado por los genes.

Preferentemente, uno de los genes ensamblados en los poligenes es un gen que codifica una proteasa capaz de escindir los sitios de escisión por proteasas que conectan los polipéptidos codificados por los poligenes. Las proteasas preferentes se han definido anteriormente.

La construcción de los poligenes, así como la de los vectores de la presente invención, puede realizarse mediante diversas técnicas de biología molecular que son generalmente conocidas por los expertos en la técnica (ver, por ejemplo, Ausubel y otros (eds.) Protocolos actuales en biología molecular ("Current Protocols in Molecular Biology"), John Wiley & Sons, Hoboken, NJ, USA, 2003). La producción de un poligen puede llevarse a cabo por ejemplo mediante

amplificación PCR de las secuencias de nucleótidos que codifican los polipéptidos particulares (por ejemplo utilizando los moldes de ADNc correspondientes), preferentemente utilizando un cebador (ya sea 5' o 3') que proporciona la secuencia o secuencias que codifican al menos un lugar de escisión por proteasas y/o al menos un péptido autoescindible. Preferentemente, los cebadores contienen además una secuencia de reconocimiento de una enzima de restricción adecuada. Preferentemente, cada cebador contiene un sitio de restricción que es diferente del sitio de restricción del otro cebador de manera que es posible la ligadura direccional del producto de amplificación resultante con otro producto de amplificación y/o un vector linealizado que contenga los mismos sitios de restricción. De acuerdo con otra realización preferente para la producción de los constructos mediante ligadura direccional, los cebadores pueden contener una secuencia de reconocimiento de un enzima de restricción que produce una proyección que no es auto-ligable tal como RsrII o BstEII. En caso de que los cebadores no contengan en sí mismos un sitio de restricción, también es posible proporcionar los productos de amplificación con adaptadores que comprenden el sitio o sitios de restricción deseados. Obviamente, también es posible proporcionar las regiones de codificación de los polipéptidos deseados utilizando cualquier fuente (más allá de la amplificación). Por ejemplo, las secuencias necesarias pueden estar ya presentes como tales o en los vectores correspondientes a partir de los cuales pueden cortarse las secuencias mediante las enzimas de restricción adecuadas. Los constructos que no contienen sitios de restricción adecuados etc., puede proporcionarse con las secuencias adecuadas (sitio o sitios de restricción, sitio o sitios de escisión por proteínas/péptido o péptidos autoescindibles; ligadores etc.) mediante la ligadura de los adaptadores adecuados que contengan los elementos necesarios. A continuación, los productos de amplificación o cualquier otro constructo adecuado (que contenga sitios de restricción) se cortan con las enzimas de restricción adecuadas. Obviamente, la secuencia de cualquier cebador utilizado se selecciona preferentemente de manera que tras el ensamblaje final de los poligenes, preferentemente en un vector adecuado, se produzca un ORF individual para cada poligen.

De acuerdo con una realización preferente de la invención, a continuación, los productos de amplificación u otras secuencias adecuadas pueden ligarse secuencialmente o simultáneamente en un vector adecuado, por ejemplo en MCS1 o MCS2 del sistema de vector MultiBac al que se ha hecho referencia anteriormente. Por ejemplo, puede introducirse un poligen en pFDBM y otro poligen puede ligarse en pUCDM (tal como se describe en WO 2005/085456 A1 (PCT/EP2004/013381)). A continuación, se utilizan los constructos resultantes para la producción de los bácmidos correspondientes mediante recombinación específica de sitio *cre-lox* (derivado pUCDM) y transposición específica Tn7 (derivado pFDBM) dando lugar a un vector de expresión de baculovirus listo para infectar las células de insecto correspondientes para expresar y purificar el complejo de multiproteínas de interés. Además de producir los poligenes e introducirlos en un vector adecuado mediante restricción/ligadura, también es posible ensamblar dichos constructos mediante recombinación homóloga utilizando las recombinasas adecuadas. Son ejemplos adecuados de técnicas clonadas basadas en recombinasas el sistema In-Fusion® disponible de BD Biosciences Clontech, Heidelberg, Alemania (ver Clontechniques, Octubre 2002, p. 10) y el sistema de recombinación Red®/ET® disponible de Gene Bridges GmbH, Dresden, Alemania (ver la patente WO-A-99/29837; <http://www.genebridges.com>).

El sistema In-Fusion® requiere regiones homólogas de 15 pares de bases (bp) en una molécula de ADN para ser fusionadas en un constructor lineal (tal como un vector adecuado) que poseen las regiones homólogas correspondientes. Por consiguiente, un vector que contiene un poligen de la presente invención puede ensamblarse secuencialmente o simultáneamente mediante la provisión de las regiones codificadoras que los constituyen (conjuntamente con secuencias de ligadura que codifican un sitio de escisión por proteínas y/o un péptido autoescindible) con las secuencias homólogas adecuadas de 15 bp. Si se desea, las secuencias homólogas de 15 bp pueden seleccionarse de manera que las regiones de codificación constituyentes se ensamblen en un orden deseado. Obviamente, pueden introducirse regiones de homología adecuadas mediante amplificación PCR de los fragmentos correspondientes utilizando cebadores que contengan las secuencias deseadas.

El sistema de recombinación Red®/ET® es diferente del sistema In-Fusion® en que se requieren secuencias de homología de 40 a 60 pares de bases, pero el constructor en el cual debe insertar un fragmento no necesita ser lineal. La recombinación se realiza *in vivo* en un huésped, preferentemente una cepa de *E. coli* que expresa el sistema de recombinasa dual "ET" RecE/RecT o Redα/Redβ. Así, los fragmentos que forman los poligenes pueden transformarse directamente conjuntamente con un vector adecuado en el huésped adecuado, preferentemente células de *E. coli*. De este modo, cada poligen puede ensamblarse en el vector adecuado ya sea secuencialmente fragmento por fragmento (preferentemente comprendiendo la región de codificación de cada polipéptido miembro del complejo de multiproteínas + al menos la secuencia de un sitio de escisión por proteasas/secuencia de un péptido autoescindible) o por transformación simultánea de todos los fragmentos.

Los constructos de polinucleótido especialmente preferentes de la presente invención contienen poligenes de al menos una longitud similar, dado que los inventores han hallado que la expresión de las poliproteínas correspondientes a dichos constructos de lugar a niveles de expresión comparables. De acuerdo con la presente invención, "poligenes de longitud similar" significa que las longitudes de las secuencias de nucleótidos de los poligenes difieren entre sí no más de un 50%, más preferentemente no más de un 30%, en particular no más de un 20% o incluso menos.

Además, la presente invención da a conocer un método para la producción de complejos de multiproteínas *en Vitro* que comprende las etapas de:

(a) cultivar la célula huésped de acuerdo con la presente invención en un medio adecuado bajo condiciones que permiten la expresión de los poligenes; y

(b) recuperar los productos de expresión codificados por los poligenes a partir del medio y/o las células huésped.

La presente invención también hace referencia a un método para la producción de complejos de multiproteínas *in vivo* que comprende las etapas de:

(a) generar al menos dos poligenes comprendiendo cada uno de ellos al menos tres genes dentro de un único ORF tal como se ha definido anteriormente; y

(b) transformar los poligenes en un animal no humano de manera que los poligenes se expresan en dicho animal no humano.

Preferentemente, la transformación del animal no humano con los poligenes de acuerdo con la etapa (b) se realiza mediante un vector, en particular un vector viral, más preferentemente un vector baculovirus. Los baculovirus son vehículos especialmente útiles para introducir poligenes en especies de mamíferos. El método *in vivo* anterior se realiza preferentemente en mamíferos no humanos, *C. elegans* o insectos. Los ejemplos particularmente preferentes de especies animales adecuadas se han definido previamente en la presente invención.

Las realizaciones de la presente invención también son útiles para la preparación de vacunas dirigidas contra ensamblajes de multisubunidades de proteínas. Los complejos de múltiples subunidades muestran habitualmente diferentes epitopos en comparación con las proteínas individuales que forman parte de los complejos. Por tanto, los complejos de multiproteínas producidos de acuerdo con la presente invención muestran los epitopos naturales relevantes de forma más adecuada, proporcionando de este modo mejores dianas antigénicas para la producción de anticuerpos.

Recientemente se han producido partículas similares a viriones (VLPs) formadas por cuatro proteínas a partir del coronavirus del síndrome respiratorio agudo severo (SARS) utilizando un vector de expresión baculovirus recombinante (consultar Mortola y otros (2004) FEBS Lett. 576, 174-178). La expresión efectiva de dichas partículas infecciosas para la preparación de vacunas puede facilitarse en gran medida utilizando el sistema de expresión de poligenes de acuerdo con la presente invención. Particularmente, mediante las herramientas de expresión de la presente invención se consigue un elevado rendimiento de expresión de ensamblajes de multisubunidades que contienen substancialmente más polipéptidos que en el ejemplo del SARS-VLPs.

Por lo tanto, es posible diseñar un método para la producción de una vacuna que comprende las etapas de:

(a) administrar al menos un polinucleótido y/o vector de la presente invención a un mamífero, de manera que el poligen de la invención es expresado en el mamífero;

(b) opcionalmente administrar un adyuvante al mamífero; y

(c) opcionalmente aislar los anticuerpos y/o células del bazo que producen anticuerpos específicos para al menos uno de los polipéptidos codificados por los poligenes.

La presente invención da a conocer una estrategia simple y adecuada para la producción recombinante de ensamblajes de multiproteínas. En estos ensamblajes de multisubunidades pueden evaluarse las interacciones de complejos proteicos o las modificaciones de las proteínas que forman parte del ensamblaje de multisubunidades. En los ensamblajes de multisubunidades producidos de acuerdo con la presente invención también puede analizarse su interacción con compuestos candidatos (moléculas orgánicas pequeñas, ácidos nucleicos, péptidos, polipéptidos, etc.) que pueden ejercer una actividad biológicamente significativa con valor médico.

Por lo tanto, la presente invención también se dirige a un método para analizar interacciones entre complejos proteicos o modificaciones de proteínas.

De acuerdo con una realización preferente, la presente invención da a conocer un método para el cribado de interacciones entre complejos proteicos o modificaciones de complejos de multiproteínas *in vitro* que comprende las etapas de:

(a) proporcionar una célula huésped de acuerdo con la presente invención que contiene al menos dos poligenes;

(b) mantener la célula huésped bajo condiciones que permitan la expresión de los poligenes; y

(c) detectar interacciones entre o modificaciones de los polipéptidos codificados por los poligenes

Otra realización preferente de la presente invención es un método para el cribado *in vitro* de compuestos candidatos capaces de (i) interaccionar con un complejo de multiproteínas y/o (ii) modificar proteínas dentro de un complejo de multiproteínas y/o (iii) inhibir interacciones dentro o entre complejos de multiproteínas y/o inhibir modificaciones de proteínas dentro de un complejo de multiproteínas, que comprende las etapas de:

(a) proporcionar una célula huésped de la presente invención que contiene al menos dos poligenes;

(b) mantener las células huésped bajo condiciones que permitan la expresión de los poligenes;

(c) poner en contacto el compuesto candidato con la célula huésped; y

5 (d) detectar interacciones de los productos de expresión con el compuesto candidato y/o interacciones entre los productos de expresión y/o modificaciones de los productos de expresión y/o inhibición de interacciones entre los productos de expresión.

10 Los polinucleótidos y/o vectores de la presente invención también son adecuados para el cribado de interacciones proteína-proteína, proteína-complejo de (multi)proteínas o complejo de multiproteínas-complejo de multiproteínas o modificaciones (fosforilación, glicosilación etc.) de complejos de multiproteínas *in vivo*.

15 Por lo tanto, es posible diseñar un método *in vivo* para el cribado de compuestos candidatos capaces de (i) interactuar con un complejo de multiproteínas y/o (ii) modificar proteínas dentro de un complejo de multiproteínas y/o (iii) inhibir interacciones dentro o entre complejos de multiproteínas y/o inhibir modificaciones de (a) proteínas dentro de un complejo de multiproteínas, que comprende las etapas de:

20 (a) proporcionar un animal no humano que comprende al menos un polinucleótido y/o vector de la presente invención que contiene al menos dos poligenes tal como se han definido anteriormente, de manera que los poligenes se expresan en el animal;

(b) administrar un compuesto candidato al animal; y

25 (c) detectar interacciones de los productos de expresión con el compuesto candidato y/o interacciones entre los productos de expresión y/o modificaciones de los productos de expresión y/o inhibición de interacciones entre los productos de expresión.

30 Las herramientas para la expresión de multiproteínas de la presente invención también tienen uso médico. En particular, pueden administrarse a animales o humanos complejos de multiproteínas bioactivas, así como combinaciones ventajosas médicamente de proteínas, por ejemplo mezclas de anticuerpos, opcionalmente en combinación con interleuquinas y/o adyuvantes, mediante los polinucleótidos y/o vectores génicos de la presente invención.

35 Por consiguiente, es posible utilizar el polinucleótido y/o vector y/o célula huésped descritos previamente para la preparación de un medicamento que comprende un vehículo de transferencia de poligenes para terapia génica.

Se están realizando enormes esfuerzos para desarrollar sistemas de administración de genes con fines terapéuticos. En el pasado, la terapia génica ha sido el centro de un intenso entusiasmo pero también de críticas. Hasta la fecha, se han realizado grandes progresos en la evaluación de la terapia génica en ensayos clínicos sobre cómo conseguir estrategias *in vivo* y *ex vivo* seguras y aplicables clínicamente para enfermedades humanas (ver Worgall S. (2004) Peediatr. Nephrol.). En resumen, en la actualidad la terapia génica constituye un método muy prometedor para la corrección de enfermedades genéticas y adquiridas que supongan la expresión permanente o transitoria de un producto genético terapéutico (Worgall S., ver anterior). Los vectores recombinantes basados en virus, en particular aquellos que no son competentes para replicarse en huéspedes mamíferos (por ejemplo los vectores baculovirus) han surgido recientemente como herramientas potentes para la administración de genes a células mamíferas y se han aplicado con éxito a una amplia gama de líneas celulares de mamífero incluyendo células humanas, de primates, de roedores, bovinas, porcinas y ovinas (revisado en Kost y Condreay (2002) Trends Biotech. 20, 173-180). Para obtener efectos de transferencia/terapia génica complejos, tanto *ex vivo* como *in vivo*, ha surgido una demanda en aumento de vectores virales policistrónicos para conseguir resultados más potentes utilizando una terapia génica combinada en lugar de una terapia génica individual (de Felipe (2002), Curr. Gene Ther. 2, 355-378; Planelles (2003) Math. Mol. Biol. 50 229, 273-284). También se ha demostrado la necesidad de incorporar genes accesorios en un virus portador que quiere administrarse *in vivo*, por ejemplo para bloquear la inactivación por el sistema de complemento, utilizando un baculovirus pseudotipado con una proteína de la envoltura gp64 de baculovirus que porta una fusión de dominio con la proteína factor de aceleración de la degradación humana (Hueser y otros (2001) Nat. Biotech. 19, 451-455), ejemplificando la necesidad de proporcionar modificaciones recombinantes a nivel de la producción de virus además de en los múltiples genes que deben transferirse con objetivos terapéuticos.

60 Por consiguiente, los baculovirus recombinante de la presente invención son preferentes para la preparación de medicamentos terapéuticos genéticos. Más preferentemente, el vector utilizado para el medicamento de la presente invención es un baculovirus que comprende al menos dos poligenes tal como se han definido anteriormente que codifican:

(i) uno o más polipéptidos terapéuticos y

(ii) una o más proteínas de baculovirus

65 En una realización preferente, la proteína o proteínas según (ii) son proteínas de baculovirus humanizadas expresadas a partir de baculovirus psudotipados, preferentemente una proteína de la envoltura gp64 de baculovirus humanizada, por ejemplo fusionada con una proteína humana como por ejemplo el factor acelerador de la desintegración.

Adicionalmente, son aplicaciones posibles los métodos de terapia génica *in vivo* que comprenden las etapas de:

- (a) proporcionar un vehículo de transferencia génica que comprende un polinucleótido de acuerdo con la invención;
- (b) administrar el vehículo de transferencia de poligenes a un paciente que padece una enfermedad genética.

Otra posibilidad de aplicación consiste en un método de terapia génica *ex vivo* que comprende las etapas de:

- (a) obtener células de un paciente que padece una enfermedad genética;
- (b) transformar las células obtenidas con un vehículo de transferencia de poligenes que comprende un polinucleótido de acuerdo con la presente invención; y
- (c) administrar las células transformadas al paciente.

Los complejos de multiproteínas producidos de acuerdo con la presente invención pueden utilizarse de forma ventajosa en estudios biofísicos, en particular estudios estructurales utilizando técnicas cristalográficas, de microscopía electrónica y/o de resonancia nuclear magnética, estudios de química proteica, en particular de interacciones proteína-proteína, y para el desarrollo de fármacos.

Así, es posible utilizar el polinucleótido y/o vector y/o célula huésped de la presente invención para la cristalización de complejos de multiproteínas.

Una realización adicional de la presente invención es un kit para la preparación de complejos de múltiples proteínas que comprende:

- (a) cebadores para amplificación PCR de las secuencias de codificación que constituyen los poligenes;
- (b) un vector plásmido o viral; y
- (c) opcionalmente células huésped adecuadas para la propagación del plásmido o vector

Los cebadores se diseñan convenientemente para adecuarse a las necesidades de la producción de un ORF individual para cada poligen y pueden contener sitios de restricción para la ligadura (secuencial o simultánea) en el vector plásmido o viral y/o los cebadores pueden contener secuencias para el ensamblado de los poligenes y/o la inserción en el vector plásmido o viral mediante recombinación homóloga (por ejemplo utilizando el sistema In-Fusion® o el sistema Red®/ET® tal como se ha descrito anteriormente).

Las Figuras muestran:

La Figura 1, muestra la secuencia de nucleótidos (identificador de secuencia nº 1) y la secuencia deducida de aminoácidos (identificador de secuencia nº 2) de un producto que codifica el núcleo de la proteína de unión a la caja TATA humana (hTBP) (hTBPC, fragmento c-terminal de la proteína completa truncada en la posición 159). Se indican las posiciones de los sitios de restricción RsrII (presentes en las secuencias de los cebadores).

La Figura 2, muestra las fotografías del análisis electroforético en gel de agarosa de la ligadura *in vitro* de segmentos del gen hTBPC y de la subclonación de la muestra. El gen hTPBC amplificado mediante PCR se digirió con RsrII y se purificó (carril 1). La incubación con ligasa da lugar a concatámeros en escalera que contienen 1, 2, 3 o más genes unidos cada uno de ellos en un ORF (carril 2, carril 3 es una escalera de 1 kb del marcador de ADN MBI). La subclonación de la mezcla de los constructos de expresión así originados conteniendo cada uno de ellos un poligen con un número diferente de genes hTBPC unidos que pueden liberarse mediante una digestión de restricción utilizando RsrII (carriles 4-7). Digestión fuera del poligen insertado muestras 1 (carril 8), 2 (carril 9), 3 (carril 10) y 5 (carril 11) genes hTBPC que dan lugar a un ORF individual en cada caso.

La Figura 3, muestra una representación esquemática del vector de transferencia básico pSPL que es la base de los constructos de vector de transferencia preferentes de la presente invención.

La Figura 4, muestra una representación esquemática del vector de transferencia básico pFL que es la base de los constructos de vector de transferencia preferentes de la presente invención.

La Figura 5, muestra una representación esquemática del vector de transferencia básico pKL que es la base de los constructos de vector de transferencia preferentes de la presente invención.

La Figura 6, muestra una representación esquemática del vector de transferencia básico pKDM que es la base de los constructos de vector de transferencia preferentes de la presente invención.

La Figura 7, muestra una representación esquemática del constructo de vector de transferencia pFBDO[hTBPC]3.

La Figura 8, muestra la secuencia de nucleótidos de pFBDO[hTBPC]3 (identificador de secuencia nº 3).

La Figura 9, muestra una representación esquemática del constructo de vector de transferencia pUCDMCSTAF1TBPcTAF2.

La Figura 10, muestra la secuencia de nucleótidos de pUCDMCSTAF1TBPcTAF2 (identificador de secuencia nº 4).

La Figura 11, muestra una representación esquemática del constructo de vector de transferencia pFBDO[HisTEVTAF6TAF9]his.

La Figura 12, muestra la secuencia de nucleótidos de pFBDO[HisTEVTAF6TAF9]his (identificador de secuencia nº 5).

La presente invención se ilustra adicionalmente mediante los siguientes ejemplos no limitantes.

Ejemplos

Ejemplo 1

Producción de poligenes y ligadura en vectores de expresión

El principio de la generación de poligenes se muestra en este ejemplo utilizando el núcleo de la proteína de unión a la caja TATA humana (hTBP) (hTBPC, fragmento c-terminal de la proteína completa truncada en la posición 159). Se amplificó el gen que codifica hTBPC mediante reacción en cadena de la polimerasa (PCR) utilizando un cebador sentido que se hibrida al extremo 5' del gen que contiene una proyección que posee un sitio de restricción RsrII y que codifica además un separador de aminoácidos y un sitio de escisión del virus del grabado del tabaco (TEV). El cebador antisentido se hibrida al extremo 3' del gen y contiene un sitio de restricción RsrII. RsrII es un enzima de restricción que produce una proyección asimétrica de 3 nucleótidos que no se autoligan, por tanto, el producto de restricción es asimétrico y la ligadora produce un producto direccional. El producto de la PCR se digirió con RsrII y se purificó. En la Figura 1 se muestran la secuencia de ADN (identificador de secuencia nº 1) y la secuencia de aminoácidos deducida (identificador de secuencia nº 2) del producto de la PCR.

La ligadura dio lugar a concatámeros de la hTBPC tal como se muestra en la Figura 2. La subclonación de la mezcla de reacción de la ligadura *in vitro* en un vector adecuado dio lugar a constructos de expresión que contenían poligenes que codificaban 1, 2, 3 y 5 proteínas hTBP en una única poliproteína separadas por sitios de escisión para la proteasa TEV. En las Figuras 3 y 4 se muestran, respectivamente, una representación esquemática y la secuencia de nucleótidos (identificador de secuencia nº 3) de uno de los vectores de expresión resultantes (pFBDO[hTBPC]3).

Ejemplo 2

Generación de vectores de transferencia de baculovirus que contienen poligenes que codifican subunidades de un factor de transcripción general humano

Se generó un poligen que codifica una poliproteína que comprende los factores asociados a TBP hTAF1 y hTAF2 además de hTBPC insertados en un vector de transferencia pUDCM (ver la patente WO 2005/085456 A1 (PCT/EP 2004/013381)) para su expresión por baculovirus, con los genes separados por secuencias que codifican un separador de aminoácidos y un sitio para la proteasa TEV. En la Figura 9 se muestra una representación esquemática del constructo resultante pUCDMCSTAF1TBPcTAF2. La secuencia de nucleótidos del constructo se muestra en la Figura 10 (identificador de secuencia nº 4). Se generó un constructo adicional que contiene un poligen que codifica una poliproteína que comprende una proteasa TEV y los factores asociados a TBP hTAF6 y hTAF9 insertados en el vector de transferencia pFBDM (ver la patente WO 2005/085456 A1 (PCT/EP2004/013381)) para su expresión por baculovirus, con los genes separados por un separador de aminoácidos y un sitio para la proteasa TEV. En la Figura 11 se muestra una representación esquemática del constructo resultante pFDDO[HisTEVTAF6TAF9]his. La secuencia de nucleótidos de este constructo se muestra en la Figura 12 (identificador de secuencia nº 5).

Ejemplo 3

Preparación de constructos de bácmidos, infección de células de insecto y expresión de proteínas

Para la construcción de constructos de bácmidos que comprenden los dos poligenes anteriores, se introdujeron los constructos pUCDMCSTAF1TBPcTAF2 (derivado de pUDCM) y pFDDO[HisTEVTAF6TAF9]his (derivado de pFBDM) en células DH10MultiBac^{Cre} tal como se describe en los Ejemplos 5 (para pUCDMCSTAF1TBPcTAF2; recombinación específica de sitio Cre-lox) y 6 (para pFDDO[HisTEVTAF6TAF9]his; transposición Tn7) de la patente WO 2005/085456 A1 (PCT/EP2004/013381). Si se desea también puede realizarse la transposición/recombinación específica de sitio cre-lox en una sola etapa en células DH10MultiBac^{Cre} tal como se describe en la patente WO 2005/085456 A1 (PCT/EP2004/013381). La preparación de bácmidos, la infección de células de insecto y la expresión de proteínas se realizó de acuerdo con protocolos establecidos (ver, por ejemplo, O'Reilly y otros (1994) Vectores de

ES 2 324 495 T3

expresión baculovirus. Un manual de laboratorio. ("Baculovirus expression vectors. A laboratory manual") Oxford University Press, New York - Oxford; Manual de sistemas de expresión en baculovirus "Bac-to Bac™" ("Bac-to Bac™ Baculovirus Expression Systems Manual") Invitrogen, Life Technologies. Inc., 2000).

5 El siguiente listado de secuencias es parte de la presente descripción en la que las secuencias representan:

El identificador de secuencia nº 1 es la secuencia de nucleótidos del producto de PCR que codifica el núcleo de la proteína de unión a la caja TATA humana (hTBP) (hTBPC, fragmento c-terminal de la proteína completa truncada en la posición 159) mostrada en la Figura 1.

10 El identificador de secuencia nº 2 es la secuencia de aminoácidos del núcleo de la proteína de unión a la caja TATA humana (hTBPC) mostrada en la Figura 1.

15 El identificador de secuencia nº 3 es la secuencia de nucleótidos de pFBDO[hTBPC]3 mostrado en la Figura 8.

El identificador de secuencia nº 4 es la secuencia de nucleótidos de pUCDMCSTAF1TBPcTAF2 mostrado en la Figura 10.

20 El identificador de secuencia nº 5 es la secuencia de nucleótidos de pFBDO[HisTEVTAF6TAF9]his mostrado en la Figura 12.

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Polinucleótido que codifica, cada uno de ellos dentro de un marco de lectura abierto individual (ORF), al menos dos poligenes que codifican cada uno de ellos al menos tres polipéptidos activos biológicamente en el que al menos dos de los polipéptidos codificados por los genes que forman parte de los poligenes son de origen no viral, al menos dos de los polipéptidos codificados por los genes que forman parte de los poligenes son capaces de interactuar al menos transitoriamente con al menos otro de los polipéptidos codificados por los genes de dichos poligenes, y en el que los genes que forman parte de cada poligen están conectados entre sí por una secuencia que codifica al menos un sitio de escisión por proteasas y/o una secuencia que codifica al menos un péptido autoescindible.

2. Polinucleótido, según la reivindicación 1, en el que cada poligen está unido operativamente a una secuencia promotor capaz de controlar la expresión del poligen.

3. Polinucleótido, según las reivindicaciones 1 o 2, en el que cada ORF está flanqueada por una secuencia de finalización.

4. Polinucleótido, según cualquiera de las reivindicaciones precedentes, en el que el sitio de escisión por proteasas se selecciona del grupo formado por los sitios de escisión para las proteasas NIa de potivirus, proteasas HC de potivirus, proteasas P1 (P35) de potivirus, proteasas NIa de biovirus, proteasas codificadas por ARN-2 de biovirus, proteasas L de aftovirus, proteasas 2A de enterovirus, proteasas 2A de rinovirus, proteasas 3C de picoma, proteasas 24K de comovirus, proteasas 24K de nepovirus, proteasa similar a 3C del RTSV (virus esférico del tungro del arroz), proteasa similar a 3C de PYVF (virus de la mancha amarilla de la chirivía), trombina, factor Xa y enteroquinasa.

5. Polinucleótido, según la reivindicación 4, en el que el sitio de escisión por proteasas es un sitio de escisión de proteasa TEV (virus del grabado del tabaco).

6. Polinucleótido, según cualquiera de las reivindicaciones 1 a 4, en el que el péptido autoescindible se selecciona del grupo formado por péptidos 2A de aftovirus, y péptidos 2A de cardiovirus.

7. Polinucleótido, según cualquiera de las reivindicaciones precedentes, que contiene además el gen o genes de la proteasa o proteasas capaces de escindir el sitio o sitios de escisión por proteasas.

8. Polinucleótido, según cualquiera de las reivindicaciones precedentes, que contiene además al menos un sitio para su integración en un vector o una célula huésped.

9. Polinucleótido, según cualquiera de las reivindicaciones precedentes, que comprende una disposición funcional según la fórmula I

X-T1-MCS1-P1-[A-B]-P2-MCS2-T2-Y (I)

que comprende:

(a) al menos dos casetes de expresión T1-MCS1-P1 y P2-MCS2-T2 en una disposición cabeza a cabeza, cabeza a cola o cola a cola, comprendiendo cada uno de ellos un sitio de clonación múltiple MCS1 o MCS2, flanqueado por un promotor P1 y una secuencia de finalización T1 para MCS1 y flanqueado por un promotor P2 y una secuencia de finalización T2 para MCS2,

(b) al menos un módulo de multiplicación M entre los promotores P1 y P2 que comprende al menos dos sitios de restricción A y B,

(c) al menos dos sitios de restricción X e Y flanqueando cada uno de ellos los casetes de expresión,

en el que

(d) los sitios de restricción A y X así como B y Y son compatibles, pero

(e) los productos de unión AY y BX no son escindibles enzimáticamente por los enzimas de restricción a, b, x o y específicos de los sitios de restricción A, B, X e Y, respectivamente, y

(f) los sitios de restricción A y B así como los sitios de restricción X e Y son incompatibles, y

en el que cada poligen está insertado en uno de los casetes de expresión.

10. Polinucleótido, según la reivindicación 10, en el que

(a) los promotores P1 y P2 se seleccionan del grupo formado por polh y p10;

(b) las secuencias de finalización se seleccionan del grupo formado por SV40 y HSVtk;

(c) los sitios de restricción A y B en el módulo de multiplicación M se seleccionan del grupo formado por los sitios de restricción BstZ171I, SpeI, ClaI y NruI;

(d) los sitios de restricción X e Y se seleccionan del grupo formado por los sitios de restricción PmeI y AvrII; y

(e) los sitios para la integración en virus se seleccionan del grupo formado por un sitio para recombinasas específicas de sitio (SSRs) y un sitio para un elemento transposón.

11. Vector que comprende el polinucleótido, según cualquiera de las reivindicaciones precedentes.

12. Vector, según la reivindicación 11, que es un vector de expresión eucariota, o un vector de transferencia o un vector baculovirus.

13. Vector, según la reivindicación 12, que es un vector baculovirus, en el que los genes de baculovirus v-cath y chiA están funcionalmente alterados.

14. Célula huésped seleccionada del grupo formado por bacterias, células de levadura, células de insecto, células de nematodo y células de mamífero, que comprende el polinucleótido según cualquiera de las reivindicaciones 1 a 10 y/o el vector según cualquiera de las reivindicaciones 11 a 13.

15. Animal transgénico no humano transformado con el polinucleótido según cualquiera de las reivindicaciones 1 a 10 y/o el vector según cualquiera de las reivindicaciones 11 a 13.

16. Método para la producción de complejos de multiproteínas *in vitro* que comprende las etapas de:

(a) cultivar la célula huésped de la reivindicación 14 en un medio adecuado bajo condiciones que permitan la expresión de los poligenes; y

(b) recuperar los productos de expresión codificados por los poligenes a partir del medio y/o de las células huésped.

Fig. 1

Código universal

Número total de aminoácidos: 819, peso molecular=91858

Máximo de inicios de ORF en posición AA 541 (puede ser pos. ADN 1623) para 286 AA (858 bases), peso molecular= 32543

AAseq:

		RsrII
1		CGAATTCCTCGAGCGGT
1		R
61	CCGGAGGTAACGGATCCGAAAACCTGTATTTTCAGGGTTCGGGATTGTACCGCAGCTGC	
20	S G G N G S E N L Y F Q G S G I V P Q L	
121	AAAATATTGTATCCACAGTGAATCTTGGTTGTAACTTGACCTAAAGACCATTGCACTTC	
40	Q N I V S T V N L G C K L D L K T I A L	
181	GTGCCCGAAACGCCGAATATAATCCCAAGCGGTTTGCTGCGGTAATCATGAGGATAAGAG	
60	R A R N A E Y N P K R F A A V I M R I R	
241	AGCCACGAACCACGGCACTGATTTTCAGTTCTGGGAAAATGGTGTGCACAGGAGCCAAGA	
80	E P R T T A L I F S S G K M V C T G A K	
301	GTGAAGAACAGTCCAGACTGGCAGCAAGAAAATATGCTAGAGTTGTACAGAAGTTGGGTT	
100	S E E Q S R L A A R K Y A R V V Q K L G	
361	TTCCAGCTAAGTTCTTGGACTTCAAGATTCAGAACATGGTGGGGAGCTGTGATGTGAAGT	
120	F P A K F L D F K I Q N M V G S C D V K	
421	TTCCTATAAGGTTAGAAGGCCTTGTGCTCACCACCAACAATTTAGTAGTTATGAGCCAG	
140	F P I R L E G L V L T H Q Q F S S Y E P	
481	AGTTATTTCTGGTTTAATCTACAGAATGATCAAACCCAGAATTGTTCTCCTTATTTTG	
160	E L F P G L I Y R M I K P R I V L L I F	
541	TTTCTGGAAAAGTTGTATTAACAGGTGCTAAAGTCAGAGCAGAAAATTATGAAGCATTG	
180	V S G K V V L T G A K V R A E I Y E A F	
		RsrII
601	AAAACATCTACCCTATTCTAAAGGGATTTCAGGAAGACGACGCGGTCCGGC	
200	E N I Y P I L K G F R K T T R	

Fig. 2

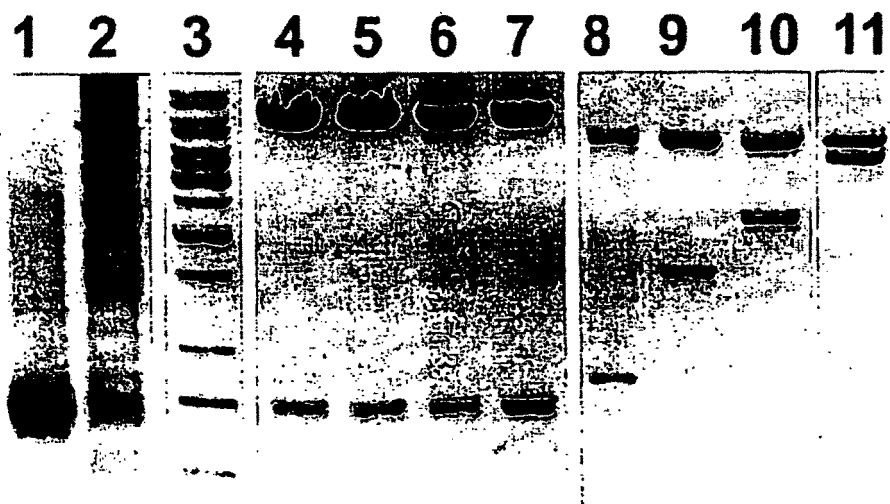


Fig. 3

pSPL

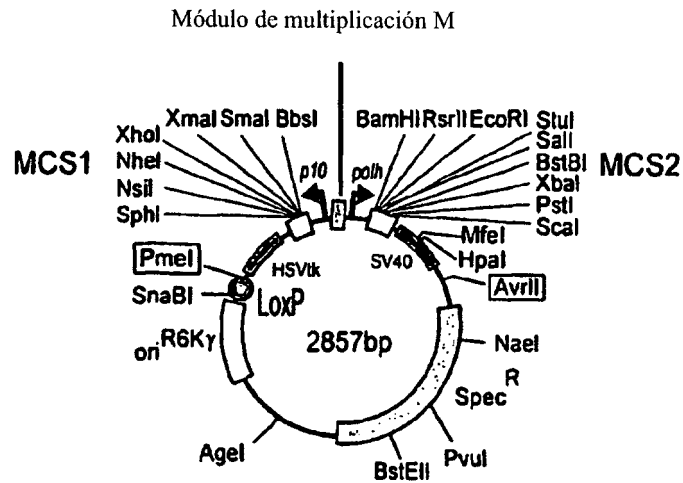


Fig. 4

pFL

Módulo de multiplicación M

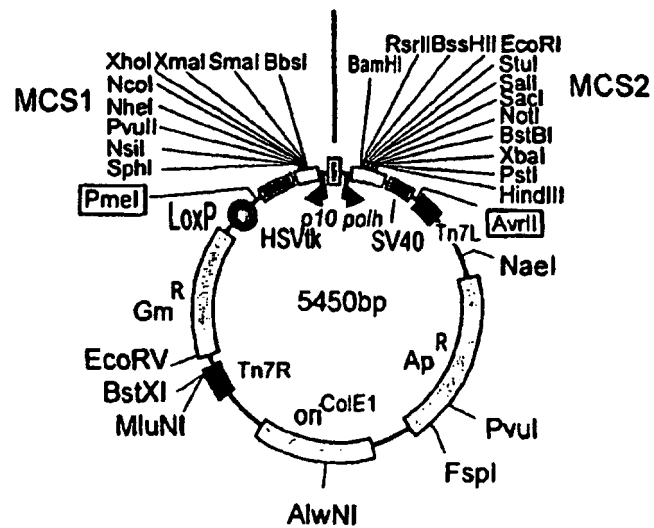


Fig. 5

pKL

Módulo de multiplicación M

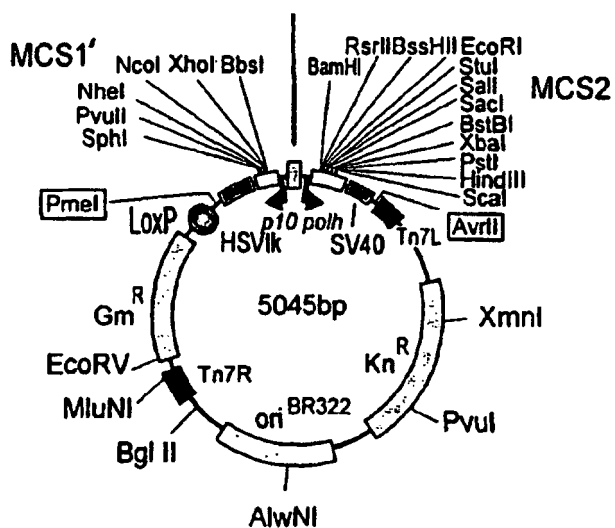


Fig. 6

pKDM

Módulo de multiplicación M

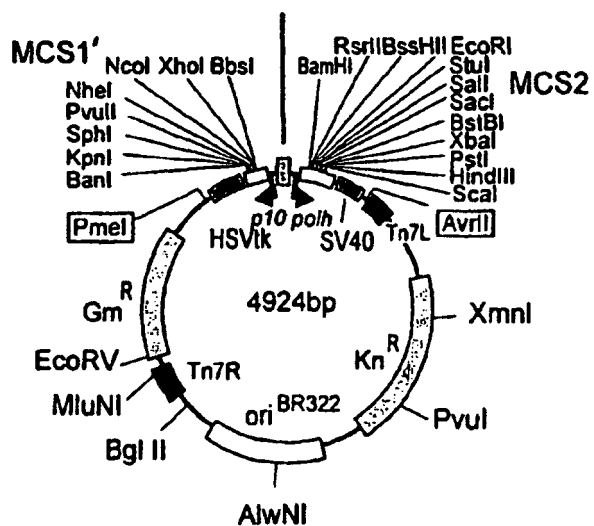


Fig. 7

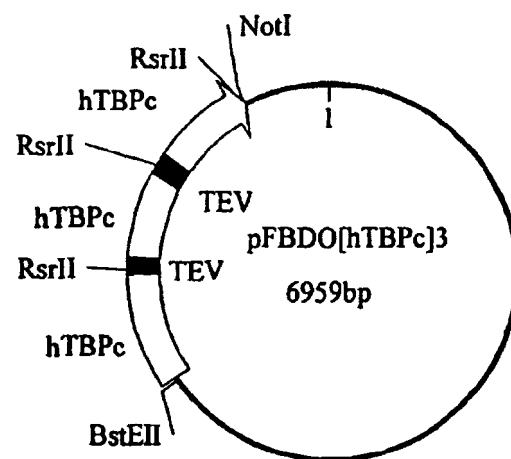


Fig. 8

SEQ pFBDO[hTBPc]3: 6959 bp;
 Composición 1827 A; 1609 C; 1674 G; 1849 T; 0 OTROS
 Porcentaje 26.3% A; 23.1% C; 24.1% G; 26.6% T; 0.0% OTROS
 Peso molecular (kDa): ssDNA: 2149.69 dsDNA: 4290.03
 PALABRA CLAVE CIRCULAR

ORIGEN

```

1      TTCTCTGTCA CAGAATGAAA ATTTTCTGT CATCTCTTCG TTATTAATGT TTGTAATTGA
61     CTGAATATCA ACGCTTATTT GCAGCCTGAA TGGCGAATGG GACGCGCCCT GTAGCGGCGC
121    ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT
181    AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG GCTTTCGCCG
241    TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA
301    CCCCCAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT
361    TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG
421    AACAACTCTC AACCCATATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT TGCCGATTTT
481    GGCCTATTGG TTAATAAATG AGCTGATTTA AAAAAAATTT AACCGGAATT TTAACAAAA
541    ATTAACGTTT ACAATTTTCA GTGGCATTGT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG
601    TTTATTTTTC TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT
661    GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCCGTG TCGCCCTTAT
721    TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT
781    AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTAC ATCGAACTGG ATCTCAACAG
841    CGGTAAGATC CTTGAGAGTT TTCGCCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA
901    AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCT
961    CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT
1021   TACGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
1081   TCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA
1141   CAACATGGGG GATCATGTAA CTCGCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT
1201   ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TCGCAAAC
1261   ATTAAGTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC
1321   GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
1381   TAAATCTGGA GCCCGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG
1441   TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG
1501   AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTTGGTAA TGTCAGACCA
1561   AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTAATTTA AAAGGATCTA
1621   GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCTGTTCA
1681   CGTAATCTGC TGCTTGCAAA CAAAAAACCC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA
1741   TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA
1801   TACTGTCCCT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC
1861   TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
1921   TCCTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC
1981   GGGGGGTTCT TGCACACAGC CCAGCTTGGG GCGAACGACC TACACCGAAC TGAGATACCT
2041   ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC
2101   GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG
2161   GTATCTTTAT AGTCCTGTCT GGTTCGCTCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
2221   CTCGTGAGGG GGGCGGAGCC TATGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTTCT
2281   GGCCFTTTGC TGGCCTTTTG CTCACATGTT CTTTCTGCTG TTATCCCTG ATTCTGTGGA
2341   TAACCGTATT ACCGCCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
2401   CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCTGATG CGGTATTTTC TCCTTACGCA
2461   TCTGTGCGGT ATTTACACCC GCAGACCAGC CGCGTAACCT GGCAAAATCG GTTACGGTTG
2521   AGTAATAAAT GGATGCCCTG CGTAAGCGGG TGTGGGCGGA CAATAAAGTC TTAACCTGAA
2581   CAAAATAGAT CTAAACTATG ACAATAAAGT CTTAACTAG ACAGAATAGT TGTAAACTGA
2641   AATCAGTCCA GTTATGCTGT GAAAAAGCAT ACTGGACTTT TGTATGGCT AAAGCAAAC
2701   CTTCATTTTC TGAAGTGCAA ATTGCCGCTC GTATTAAAGA GGGGCGTGGC CAAGGGCATG
2761   GTAAAGACTA TATTGCGGCG GTTGTGACAA TTTACCGAAC AACTCCGCGG CCGGGAAGCC
2821   GATCTCGGCT TGAACGAATT GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTGCATCAC

```

Fig. 8 (continuación)

```

2881 TTCTTCCCGT ATGCCCAACT TTGTATAGAG AGCCACTGCG GGATCGTCAC CGTAATCTGC
2941 TTGCACGTAG ATCACATAAG CACCAAGCGC GTTGGCCTCA TGCTTGAGGA GATTGATGAG
3001 CGCGGTGGCA ATGCCCTGCC TCCGGTGCTC GCCGAGACT GCGAGATCAT AGATATAGAT
3061 CTCACTACGC GGCTGCTCAA ACCTGGGCAG AACGTAAGCC GCGAGAGCGC CAACAACCGC
3121 TTCTTGGTCG AAGGCAGCAA GCGCGATGAA TGTCTTACTA CGGAGCAAGT TCCCGAGGTA
3181 ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGCTACGTCT CCGAACTCAC GACCGAAAAG
3241 ATCAAGAGCA GCCCGCATGG ATTTGACTTG GTCAGGGCCG AGCCTACATG TGCGAATGAT
3301 GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG CCCTGCTGCG TAACATCGTT
3361 CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA CAAGCCATGA
3421 AAACCGCCAC TGCGCCGTTA CCACCGTTGC GTTCGGTCAA GGTCTGGAC CAGTTGCGTG
3481 AGCGCATACG CTACTTGCAT TACAGTTTAC GAACCGAACA GGCTTATGTC AACTGGGTTT
3541 GTGCCTTTCA CCGTTTCCAC GGTGTGCGTC ACCCGCAAC CTTGGGCAGC AGCGAAGTCG
3601 AGGCATTTC TCTCTGGCTG GCGAAGGAGC GCAAGGTTTC GGTCTCCACG CATCTCAGG
3661 CATTGGCGGC CTTGCTGTTT TTCTACGGCA ASGTGCTGTG CACGGATCTG CCCTGGCTTC
3721 AGGAGATCGG TAGACCTCGG CCGTCGCGCG GCTTGCCGGT GGTGCTGACC CCGGATGAAG
3781 TGGTTCGCAT CCTCGGTTTT CTGGAAGCGG AGCATCGTTT GTTCGCCCAG GACTCTAGCT
3841 ATAGTTCTAG TGGTTGGCCT ACAGCTTTGT TTAACAAAG CTGTACCCGT AGTGGCTATG
3901 GCAGGGCTTG CCGCCCCGAC GTTGGCTGCG AGCCCTGGGC CTTACCCGA ACTTGGGGT
3961 TGGGGTGGGG AAAAGGAAGA AACCGGGCG TATTGGTCCC AATGGGGTCT CGGTGGGGTA
4021 TCGACAGAGT GCCAGCCCTG GGACCGAACC CCGCGTTTAT GAACAAACGA CCCAACACCC
4081 GTGCGTTTTA TTCTGTCTTT TTATTGCCGT CATAGCGCGG GTTCCTTCCG GTATTGTCTC
4141 CTTCCGTGTT TCAGTTAGCC TCCCCATCT CCCGGTACCG CATGCTATGC ATCAGCTGCT
4201 AGCACCATGG CTCGAGATCC CGGGTGATCA AGTCTTCGTC GAGTGATTGT AAATAAATG
4261 TAATTTACAG TATAGTATTT TAATTAATAT ACAAATGATT TGATAATAAT TCTTATTTAA
4321 CTATAATATA TTGTGTTGGG TTGAATTAAA GGTCCGTATA CTAGTATCGA TTCGCGACCT
4381 ACTCCGGAAT ATTAATAGAT CATGGAGATA ATTAAATGA TAACCATCTC GCAAATAAAT
4441 AAGTATTTTA CTGTTTTCTG AACAGTTTGT TAATAAAAAA ACCTATAAAT ATTCCGGATT
4501 ATTCATACCG TCCCACCATC GGGCGCGGAT CCTCGAGATG GGTAAACATG ACAAGCGACG
4561 ATGGA AAAAG AATTT CATAG CCGTCTCAGC AGCCAACCGC TTTAAGAAAA TCTCATCCTC
4621 CGGGGCAGCT AGCTGGAGCC ACCCGCAGTT CGAAAAAGGC GCCGACGACG ACGACGACAA
4681 GGGCTCCCAT ATGTCTGGGA TTGTACCGCA GCTGCAAAAT ATTGTATCCA CAGTGAATCT
4741 TGGTTGTAAA CTTGACCTAA AGACCATTGC ACTTCGTGCC CGAAACGCCG AATATAATCC
4801 CAAGCGGTTT GCTGCGGTAA TCATGAGGAT AAGAGAGCCA CGAACCACGG CACTGATTTT
4861 CAGTTCGGG AAAATGGTGT GCACAGGAGC CAAGAGTGAA GRACAGTCCA GACTGGCAGC
4921 AAGAAAATAT GCTAGAGTTG TACAGAAGTT GGGTTTTCCA GCTAAGTTCT TGGACTTCAA
4981 GATTCAGAAC ATGGTGGGGA GCTGTGATGT GAAGTTTCTT ATAAGGTTAG AAGGCCTTGT
5041 GCTACCCAC CAACAATTTA GTAGTTATGA GCCAGAGTTA TTTCTGGTT TAATCTACAG
5101 AATGATCAAA CCCAGAATTG TTCTCCTTAT TTTTGTCTT GGAAGAGTTG TATTAACAGG
5161 TGCTAAAGTC AGAGCAGAAA TTTATGAAGC ATTTGAAAC ATCTACCCTA TTCTAAAGGG
5221 ATTCAGGAAG ACGACGCGGT CCGGAGGTAA CGGATCCGAA AACCTGTATT TTCAGGGTTC
5281 TGGGATTGTA CCGCAGCTGC AAAATATTGT ATCCACAGTG AATCTTGGTT GTAACCTTGA
5341 CCTAAAGACC ATTGCACTT CCGCCGAAA CGCCGAATAT AATCCCAAGC GGTGCTGTC
5401 GGTAATCATG AGGATAAGAG AGCCACGAAC CACGGCACTG ATTTTCAGTT CTGGGAAAAAT
5461 GGTGTGCACA GGAGCCAAGA GTGAAGAACA GTCCAGACTG GCAGCAAGAA AATATGCTAG
5521 AGTTGTACAG AAGTTGGGTT TTCCAGCTAA GTTCTTGAC TTCAAGATTC AGAACATGGT
5581 GGGGAGCTGT GATGTGAAGT TTCCTATAAG GTTAGAAGGC CTTGTGCTCA CCCACCAACA
5641 ATTTAGTAGT TATGAGCCAG AGTTATTTCC TGGTTTAATC TACAGAATGA TCAAAACCCAG
5701 AATTGTTCTC CTTATTTTGT TTTCTGGAAA AGTTGTATTA ACAGGTGCTA AAGTCAGAGC
5761 AGAAATTTAT GAAGCATTTG AAAACATCTA CCCTATTCTA AAGGGATTCA GGAAGACGAC
5821 CGGTTCCGGA GGTAACGGAT CCGAAAACCT GTATTTTCAG GGTCTGGGA TTGTACCGCA
5881 GCTGCAAAAT ATTGTATCCA CAGTGAATCT TGGTTGTAAA CTTGACCTAA AGACCATTGC
5941 ACTTCGTGCC CGAAACGCCG AATATAATCC CAAGCGGTTT GCTGCGGTAA TCATGAGGAT
6001 AAGAGAGCCA CGAACCACGG CACTGATTTT CAGTTCGGG AAAATGGTGT GCACAGGAGC
6061 CAAGAGTGAA GAACAGTCCA GACTGGCAGC AAGAAAATAT GCTAGAGTTG TACAGAAGTT
6121 GGGTTTTCCA GCTAAGTTCT TGGACTTCAA GATTCAAGAC ATGGTGGGGA GCTGTGATGT

```


Fig. 8 (continuación)

```

6181 GAAGTTTCCT ATAAGGTTAG AAGGCCTTGT GCTCACCAC CAACAATTTA GTAGTTATGA
6241 GCCAGAGTTA TTTCCTGGTT TAATCTACAG AATGATCAAA CCCAGAATTG TTCTCCTTAT
6301 TTTTGTTTCT GGAAAAGTTG TATTAACAGG TGCTAAAGTC AGAGCAGAAA TTTATGAAGC
6361 ATTTGAAAAC ATCTACCCTA TTCTAAAGGG ATTCAGGAAG ACGACGCGGT CCGGCCACCA
6421 TCATCACCAC CATTGATAAG CTAGCGGCCG CTTTCGAATC TAGAGCCTGC AGTCTCGACA
6481 AGCTTGTCGA GAAGTACTAG AGGATCATAA TCAGCCATAC CACATTTGTA GAGGTTTAC
6541 TTGCTTTAAA AAACCTCCCA CACCTCCCC TGAACCTGAA ACATAAAATG AATGCAATTG
6601 TTGTTGTTAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCACAA
6661 ATTTACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTTGTCC AAACATCA
6721 ATGTATCTTA TCATGTCTGG ATCTGATCAC TGCTTGAGCC TAGGAGATCC GAACCAGATA
6781 AGTGAAATCT AGTTCCAAAC TATTTTGTCA TTTTAAATT TCGTATTAGC TTACGACGCT
6841 ACACCCAGTT CCCATCTATT TTGTCACTCT TCCCTAAATA ATCCTTAAAA ACTCCATTTC
6901 CACCCCTCCC AGTTCCCAAC GCCAACTCCA TGTGACAAAC CGTCATCTTC GGCTACTTT

```

Fig. 9

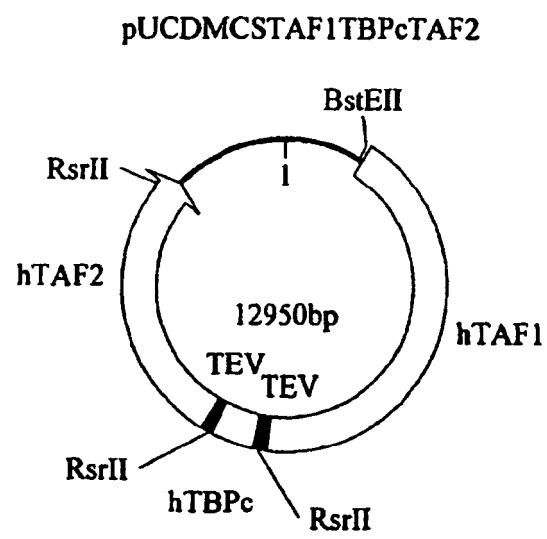


Fig. 10

SEQ pUCDMCSTAF1TBPCtAF2: 12950 bp;
 Composición 3820 A; 2716 C; 3059 G; 3355 T; 0 OTROS
 Porcentaje: 29.5% A; 21.0% C; 23.6% G; 25.9% T; 0.0% OTROS
 Peso molecular (kDa): ssDNA: 4006.92 dsDNA: 7982.98
 PALABRA CLAVE CIRCULAR

ORIGEN

```

1      AATTCTGTCA GCCGTTAAGT GTTCTGTGT CACTGAAAT TGCTTTGAGA GGCTCTAAGG
61     GCTTCTCAGT GCGTTACATC CTTGGCTTGT TGTCCACAAC CGTTAAACCT TAAAAGCTTT
121    AAAAGCCTTA TATATTCTTT TTTTCTTTAT AAAACTTAAA ACCTTAGAGG CTATTTAAGT
181    TGCTGATTTA TATTAATTTT ATTGTTCAAA CATGAGAGCT TAGTACGTGA AACATGAGAG
241    CTTAGTACGT TAGCCATGAG AGCTTAGTAC GTTAGCCATG AGGGTTTGTG TCGTTAAACA
301    TGAGAGCTTA GTACGTTAAA CATGAGAGCT TAGTACGTGA AACATGAGAG CTTAGTACGT
361    ACTATCAACA GGTTGAACTG CTGATCAACA GATCCTCTAC GCGGCCGCGG TACCATAACT
421    TCGTATAGCA TACATTATAC GAAGTTATCT GGAGTACCCG TAGTGGCTAT GGCAGGGCTT
481    GCGGCCCGCA CGTTGGCTGC GAGCCCTGGG CCTTCACCCG AACTTGGGGG TTGGGGTGGG
541    GAAAAGGAAG AAACGCGGGC GTATTGGTCC CAATGGGGTC TCGGTGGGGT ATCGACAGAG
601    TGCCAGCCCT GGGACCGAAC CCCGCGTTTA TGAACAAACG ACCCAACACC CGTGCCTTTT
661    ATTCTGTCTT TTTATTGCCG TCATAGCGCG GGTTCCTTCC GGTATTGTCT CTTCCGTGT
721    TTCAGTTAGC CTCCCCATC TCCCGGTACC GCATGCTATG CATCAGCTGC TAGCACCATG
781    GCTCGAGATC CCGGGTGATC AAGTCTTCGT CGAGTGATTG TAAATAAAAT GTAATTTACA
841    GTATAGTATT TTAATTAATA TACAAATGAT TTGATAATAA TTCTTATTTA ACTATAATAT
901    ATTGTGTTGG GTTGAATTAA AGGTCCGTAT ACTAGTATCG ATTCGCGACC TACTCCGGAA
961    TATTAATAGA TCATGGAGAT AATTAAAATG ATAACCATCT CGCAAATAAA TAAGTATTTT
1021   ACTGTTTTTC TAACAGTTTT GTAATAAAAA AACCTATAAA TATTCCGGAT TATTCATACC
1081   GTCCACCAT CGGGCGCGGA TCCTCGAGAT GGGTAACCAT GACAAGCGAC GATGGA AAAA
1141   GAATTTTATA GCCGTCTCAG CAGCCAACCG CTTAAGAAA ATCTCATCCT CCGGGGCAGC
1201   TAGCTGGAGC CACCCGAGT TCGAAAAAGG CGCCGACGAC GACGACGACA AGGGCTCCCA
1261   TAGGGGACCC GGCTGCGATT TGCTGCTGCG GACAGCAGCT ACCATCACTG CTGCGCCATG
1321   CATGTCAGAC ACGGACAGCG ACGAAGATTG CGCTGGAGGC GGCCCATTTT CTTTAGCGGG
1381   TTTCTTTTTC GGCAACATCA ATGGAGCCGG GCAGCTGGAG GGGGAAAGCG TCTTGATGA
1441   TGAATGTAAG AAGCACTTGG CAGGCTTGGG GGCTTTGGGG CTGGGCAGCC TGATCACTGA
1501   ACTCACGGCA AATGAAGAAT TGACCGGGAC TGACGGTGCC TTGGTAAATG ATGAAGGGTG
1561   GGTTAGGAGT ACAGAAGATG CTGTGGACTA TTCAGACATC AATGAGGTGG CAGAAGATGA
1621   AAGCCGAAGA TACCAGCAGA CGATGGGGAG CTTGCAGCCC CTTTGCCACT CAGATTATGA
1681   TGAAGATGAC TATGATGCTG ATTGTGAAGA CATTGATTGC AAGTTGATGC CTCCTCCACC
1741   TCCACCCCGG GGACCAATGA AGAAGGATAA GGACCAGGAT TCTATTACTG GTGAGAAAGT
1801   GGACTTCAGT AGTTCCTCTG ACTCAGAATC TGAGATGGGA CCTCAGGAAG CAACACAGGC
1861   AGAATCTGAA GATGGAAAGC TGACCCCTTC ATTGGCTGGG ATTATGCAGC ATGATGCCAC
1921   CAAGCTGTTG CCAAGTGTC AAGAACTTTT TCCAGAATTT CGACCTGGAA AGGTGTTACG
1981   TTTTCTACGT CTTTTTGGAC CAGGGAAGAA TGTCCCATCT GTTTGGCGGA GTGCTCGGAG
2041   AAAGAGGAAG AAGAAGCACC GTGAGCTGAT ACAGGAAGAG CAGATCCAGG AGGTGGAGTG
2101   CTCAGTAGAA TCAGAAGTCA GCCAGAAGTC TTTGTGGAAC TACGACTACG CTCCACCACC
2161   ACCTCCAGAG CAGTGTCTCT CTGATGATGA AATCACGATG ATGGCTCCTG TGGAGTCCAA
2221   ATTTTCCCAA TCAACTGGAG ATATAGATAA AGTGACAGAT ACCAAACCAA GAGTGGCTGA
2281   GTGGCGTTAT GGGCCTGCCC GACTGTGGTA TGATATGCTG GGTGTCCCTG AAGATGGCAG
2341   TGGGTTTGAC TATGGCTTCA AACTGAGAAA GACAGAACAT GAACCTGTGA TAAAATCTAG
2401   AATGATAGAG GAATTTAGGA AACTTGAGGA AAACAATGGC ACTGATCTTC TGGCTGATGA
2461   AAACCTCCTG ATGGTGACAC AGCTGCATTG GGAGGATGAT ATCATCTGGG ATGGGGAGGA
2521   TGTCAAACAC AAAGGGACAA AACCTCAGCG TGCAAGCCTG GCAGGCTGGC TTCCTTCTAG
2581   CATGACTAGG AATGCGATGG CTTACAATGT TCAGCAAGGT TTTGCAGCCA CTCTTGATGA
2641   TGACAAACCT TGGTACTCCA TTTTCCCAT TGACAATGAG GATCTGGTAT ATGGACGCTG
2701   GGAGGACAAAT ATCATTTGGG ATGCTCAGGC CATGCCCGG CTGTTGGAAC CTCCTGTTTT
2761   GACACTTGAT CCCAATGATG AGAACCTCAT TTTGGAAATT CCTGATGAGA AGGAAGAGGC
2821   CACCTCTAAC TCCCCCTCCA AGGAGAGTAA GAAGGAATCA TCTCTGAAGA AGAGTCGAAT

```

Fig. 10 (continuación)

```

2881 TCTCTTAGGG AAAACAGGAG TCATCAAGGA GGAACCACAG CAGAACATGT CTCAGCCAGA
2941 AGTGAAAGAT CCATGGAATC TCTCCAATGA TGAGTATTAT TATCCCAAGC AACAGGGTCT
3001 TCGAGGCACC TTTGGAGGGA ATATTATCCA GCATTCAATT CCTGCTGTGG AATTACGGCA
3061 GCCCTTCTTT CCCACCCACA TGGGGCCCAT CAAACTCCGG CAGTTCCATC GCCACCTCT
3121 GAAAAAGTAC TCATTGGTG CACTTTCTCA GCCAGGTCCC CACTCAGTCC AACCTTTGCT
3181 AAAGCACATC AAAAAAAGG CCAAGATGAG AGAACAAGAG AGGCAAGCTT CAGGTGGTGG
3241 AGAGATGTTT TTTATGCGCA CACCTCAGGA CCTCACAGGC AAAGATGGTG ATCTTATTCT
3301 TGCAGAAATAT AGTGAGGAAA ATGGACCCTT AATGATGCAG GTTGGCATGG CAACCAAGAT
3361 AAAGAACTAT TATAAACGGA AACCTGGAAG AGATCCTGGA GCACCAGATT GTAAATATGG
3421 GGAAACTGTT TACTGCCATA CATCTCCTTT CCTGGGTTCT CTCCATCCTG GCCAATTGCT
3481 GCAAGCATTT GAGAACCAACC TTTTTCGTGC TCCAATTTAT CTTATAAGA TGCCAGAAAC
3541 TGATTTCCTG ATCATTTCGA CAAGACAGGG TTTACTATAT CGGGAATTAG TGGATATTTT
3601 TGTGGTTGGC CAGCAGTGTC CCTTGTTTGA AGTTCCTGGG CCTAACTCCA AAAGGGCCAA
3661 TACGCATATT CGAGACTTTC TACAGGTTTT TATTACCGC CTTTCTGGA AAAGTAAAGA
3721 TCGGCCACGG AGGATACGAA TGGAGATAT AAAAAAGCC TTTCTTCCC ATTCAGAAAG
3781 CAGCATCCGG AAGAGGCTAA AGCTCTGCGC TGACTTCAA CGCACAGGGA TGGACTCAA
3841 CTGGTGGGTG CTTAAGTCTG ATTTTCGTTT ACCAACGGAA GAAGAGATCA GAGCTATGGT
3901 GTCACCAGAG CAGTGCTGTG CTTATTATAG CATGATAGCT GCAGAGCAAC GACTGAAGGA
3961 TGCTGGCTAT GGTGAGAAAT CCTTTTTCG TCCAGAAGAA GAAATGAGG AAGATTTCCA
4021 GATGAAGATT GATGATGAAG TTCGCACTGC CCCTTGGAAC ACCACAAGGG CCTTCATTGC
4081 TGCCATGAAG GGCAAGTGTG TGCTAGAGGT GACTGGGGTG GCAGATCCCA CGGGGTGTGG
4141 TGAAGGATTC TCCTATGTGA AGATTCCAAA CAAACCAACA CAGCAGAAGG ATGATAAAGA
4201 ACCGCAGCTA GTGAAGAAGA CAGTGACAGG AACAGATGCA GACCTTCGTC GCCTTCCCT
4261 GAAAAATGCC AAGCAACTTC TACGTAAATT TGGTGTGCCT GAGGAAGAGA TTAAAAAGTT
4321 GTCCCGCTGG GAAGTGATTG ATGTGGTGGC CACAATGTCA ACAGAACAGG CTCGTCTCGG
4381 AGAGGGGCCC ATGAGTAAAT TTGCCGTGG ATCAAGGTTT TCTGTGGCTG AGCATCAAGA
4441 GCGTTACAAA GAGGAATGTC AGCGCATCTT TGACCTACAG AACAGGTTT TGTCATCAAC
4501 TGAAGTCTTA TCAACTGACA CAGACAGCAG CTCAGCTGAA GATAGTGACT TTGAAGAAAT
4561 GGGAAAGAAC ATTGAGAACA TGTTCGAGAA CAAGAAAACC AGCTCTCAGC TTTCACGTGA
4621 ACGGGAGGAA CAGGAGCGGA AGGAACTACA GCGAATGCTA CTGGCAGCAG GCTCAGCAGC
4681 ATCCGGAAC AATCACAGAG ATGATGACAC AGCTTCCGTG ACTAGCCTTA ACTCTCTGC
4741 CACTGGACGC TGTCTCAAGA TTTATCGCAC GTTTCGAGAT GAAGAGGGGA AAGAGTATGT
4801 TCGCTGTGAG ACAGTCCGAA AACCAGCTGT CATTGATGCC TATGTGCGCA TACGGACTAC
4861 AAAAGATGAG GAATTCAITC GAAAAATTGC CCTTTTGTAT GAACAACATC GGGAGAGAT
4921 GCGAAAAGAA CGGCGGAGGA TTCAAGAGCA ACTGAGGCGG CTTAAGAGGA ACCAGGAAAA
4981 GGAGAAAGCTT AAGGGTCTC CTGAGAAGAA GCCCAAGAAA ATGAAGGAGC GTCTGACCT
5041 AAAACTGAAA TGTGGGGCAT GTGGTGCCAT TGGACACATG AGGACTAACA AATTCTGCCC
5101 CCTCTATTAT CAAACAAATG CGCCACCTTC CAACCCTGTT GCCATGACAG AAGAACAGGA
5161 GGAGGAGTTG GAAAAGACAG TCATTCATAA TGATAATGAA GAACTTATCA AGGTTGAAGG
5221 GACCAAAATT GTCTTGGGGA AACAGCTAAT TGAGAGTGCG GATGAGGTTT GCAGAAAATC
5281 TCTGTTCTC AAGTTTCCTA AACAGCAGCT TCCTCCAAAG AAGAAACGGC GAGTTGGAAC
5341 CACTGTTTAC TGTGACTATT TGAATAGACC TCATAAGTCC ATCCACCGGC GCCGCACAGA
5401 CCCTATGGTG ACGCTGTCGT CCATCTTGGA GTCTATCATC AATGACATGA GAGATCTTCC
5461 AAATACATAC CCTTTCACAC CTCCAGTCAA TGCAAGGTT GTAAAGGACT ACTACAAAT
5521 CATCACTCGG CCAATGGACC TACAAACACT CCGCGAAAAC GTGCGTAAAC GCCTCTACCC
5581 ATCTCGGGAA GAGTTCAGAG AGCATCTGGA GCTAATTGTG AAAAAATAGT CAACCTACAA
5641 TGGGCCAAAA CACTCATTGA CTCAGATCTC TCAATCCATG CTGGATCTCT GTGATGAAAA
5701 ACTCAAAGAG AAAGAAGACA AATTAGCTCG CTTAGAGAAA GCTATCAACC CCTTGCTGGA
5761 TGATGATGAC CAAGTGGCGT TTTCTTTCAT TCTGGACAAC ATTGTCACCC AGAAAATGAT
5821 GGCAGTTCCA GATTCTTGGC CATTTTCATC CCCAGTTAAT AAGAAAATTG TTCCAGATTA
5881 TTACAAAGTG ATTGTCAATC CAATGGATTT AGAGACCATA CGTAAGAACA TCTCCAAGCA
5941 CAAGTATCAG AGTCGGGAGA GCTTCTGGA TGATGTAAAC CTTATTCTGG CCAACAGTGT
6001 TAAGTATAAT GGACCTGAGA GTCAGTATAC TAAGACTGCC CAGGAGATTG TGAACGCTCG
6061 TTACCAGACA TTGACTGAGT ATGATGAACA TTTGACTCAA CTTGAGAAGG ATATTGTGAC

```

Fig. 10 (continuación)

6121	TGCTAAAGAA	GCAGCTTTGG	AGGAAGCAGA	ATTAGAAAGC	CTGGACCCAA	TGACCCAGG
6181	GCCCTACACG	CCTCAGCCTC	CTGATTTGTA	TGATACCAAC	ACATCCCTCA	GTATGTCTCG
6241	AGATGCCTCT	GTATTTCAAG	ATGAGAGCAA	TATGTCTGTC	TTGGATATTC	CCAGTGCCAC
6301	TCCAGAAAAG	CAGGTAACAC	AGGAAGGTGA	AGATGGAGAT	GGTGATCTTG	CAGATGAAGA
6361	GGAAGGAACT	GTACAACAGC	CTCAAGCCAG	TGTCCTGTAT	GAGGATTTGC	TTATGTCTGA
6421	AGGAGAAGAT	GATGAGGAAG	ATGCTGGGAG	TGATGAAGAA	GGAGACAATC	CTTTCTCTGC
6481	TATCCAGCTG	AGTGAAAGTG	GAAGTGACTC	TGATGTGGGA	TCTGGTGGAA	TAAGACCCAA
6541	ACAACCCCGC	ATGCTTCAGG	AGAACACAAG	GATGGACATG	GAAAATGAAG	AAAGCATGAT
6601	GTCTTATGAG	GGAGACGGTG	GGGAGGCTTC	CCATGGTTTG	GAGGATAGCA	ACATCAGTTA
6661	TGGGAGCTAT	GAGGAGCCTG	ATCCCAAGTC	GAACACCCAA	GACACAAGCT	TCAGCAGCAT
6721	CGGTGGGTAT	GAGGTATCAG	AGGAGGAAGA	AGATGAGGAG	GAGGAAGAGC	AGCGCTCTGG
6781	GCCGAGCGTA	CTAAGCCAGG	TCCACCTGTC	AGAGGACGAG	GAGGACAGTG	AGGATTTCCA
6841	CTCCATTGCT	GGGGACAGTG	ACTTGGACTC	TGATGAACGG	TCCGGAGGTA	ACGGATCCGA
6901	AAACCTGTAT	TTTCAGGGTT	CTGGGATGTT	ACCGCAGCTG	CAAAATATTG	TATCCAGAT
6961	GAATCTTGGT	TGTAACCTTG	ACCTAAAGAC	CATTGCACCT	CGTGCCCGAA	AAAGCCAGTA
7021	TAATCCCAAG	CGGTTTGCTG	CGGTAATCAT	GAGGATAAGA	GAGCCACGAA	CCACGGCACT
7081	GATTTTCAGT	TCTGGGAAAA	TGGTGTGCAC	AGGAGCCAAG	AGTGAAGAAC	AGTCCAGACT
7141	GGCAGCAAGA	AAATATGCTA	GAGTTGTACA	GAAGTTGGGT	TTTCCAGCTA	AGTTCCTGGA
7201	CTTCAAGATT	CAGAACATGG	TGGGGAGCTG	TGATGTGAAG	TTTCCTATAA	GGTTAGAAGG
7261	CCTTGTGCTC	ACCCACCAAC	AATTTAGTAG	TTATGAGCCA	GAGTTATTTT	CTGGTTTAA
7321	CTACAGAATG	ATCAAACCCA	GAATTGTTCT	CCTTATTTTT	GTTTCTGGAA	AAGTTGTATT
7381	AACAGGTGCT	AAAGTCAGAG	CAGAAATTTA	TGAAGCATT	GAAAACATCT	ACCCATTTCT
7441	AAAGGGATT	AGGAAGACGA	CGCGGTCCGG	AGGTAACGGA	TCCGAAAACC	TGTATTTTCA
7501	GGGTGACTAC	AAAGACGATG	ACGATAAAAA	CAGGAAGAAA	GGAGACAAGG	GCTTTGAAAG
7561	CCCAAGGCCA	TATAAATTAA	CCCATCAGGT	CGTCTGCATC	AACAACATAA	ATTTCCAGAG
7621	AAAATCTGTT	GTGGGATTTG	TGGAACGTAC	TATATTTCCC	ACAGTTGCAA	ACTTGAATAG
7681	AATCAAGTTG	AACAGCAAAC	AGTGTAGAAT	ATACCGAGTA	AGGATCAATG	ATTTAGAGGC
7741	TGCTTTTATT	TATAATGACC	CAACCTTGG	AGTTTGTGAC	AGTGAATCAA	AACAGAGAAA
7801	CCTCAATTAT	TTTTCCAATG	CTTATGCAGC	TGCAGTTAGT	GCTGTGGACC	CTGATGCAGG
7861	AAATGGAGAA	CTTTGCATTA	AGGTTCCATC	AGAGCTATGG	AAACACGTTG	ATGAGTTAAA
7921	GGTCCTGAAG	ATACACATCA	ATTTTCTTTT	GGATCAGCCC	AAAGGAGGTC	TTCATTTTGT
7981	GGTACCCAGT	GTAGAGGGAA	GTATGGCAGA	GAGAGGTGCT	CATGTTTTCT	CTTGTGGGTA
8041	TCAAAATTCT	ACAAGATTTT	GGTCCCTTG	TGTTGATTCA	TACTCTGAAT	TGTGTACATG
8101	GAAATTAGAA	TTTACAGTAG	ATGCTGCAAT	GGTTGCTGTT	TCTAATGGCG	ATTTGGTGG
8161	GACAGTGAT	ACTCATGATA	TGAGGAAGAA	AACCTTCCAT	TATATGCTTA	CCATTCCCTAC
8221	AGCAGCGTCA	AATATCTCCT	TGGCCATTGG	ACCATTGGA	ATACTGGTAG	ATCCATACAT
8281	GATGAGGTT	ACTCATTTT	GTTTGCCCCA	ACTTCTTCCA	TTGCTGAAAC	ATACCACATC
8341	ATACCTTCAT	GAAGTCTTTG	AATTTTATGA	AGAAATTCTT	ACATGTCGTT	ACCCATACTC
8401	CTGTTTTAAG	ACTGTCTTCA	TTGATGAGGC	TTATGTTGAA	GTGGCTGCTT	ATGCTTCCAT
8461	GAGCATTTT	AGCACAAATC	TTTTACACAG	TGCCATGATT	ATAGATGAGA	CACCTTTGAC
8521	TAGAAGGTGT	TTAGCCCAAT	CCTTGGCCCA	GCAGTTTTTT	GGTTGTTTCA	TATCTAGAAT
8581	GTCTTGGTCT	GATGAATGGG	TGCTGAAGGG	AATTTAGGC	TATATCTATG	GACTTTGGAT
8641	GAAAAAACT	TTTGGTGTGA	ATGAGTACCG	CCATTGGATT	AAAGAGGAGC	TAGACAAAAT
8701	AGTGGCATAT	GAACAAAAA	CTGGTGGGGT	TTTACTACAT	CCCATATTTG	GTGGAGGAAA
8761	AGAGAAGGAT	AATCCGGCTT	CCCATCTACA	CTTTTCAATA	AAGCATCCAC	ATACACTGTC
8821	CTGGGAATAC	TACACTATGT	TTCAAGTGTA	AGCCACCTT	GTGATGAGAT	TGATTGAAAA
8881	TAGGATCAGT	ATGGAATTTA	TGCTACAAGT	TTTCAATAAA	CTGCTAAGTC	TGGCTAGTAC
8941	TGCTTCATCT	CAGAAGTTCC	AGTCACATAT	GTGGAGTCAG	ATGTTGGTTT	CCCATCTGG
9001	GTTTTTGA	TCCATTTCAA	ATGTCTCTGG	CAAAGATATT	CAGCCGTTAA	TAAAGCAGTG
9061	GGTAGATCAG	AGTGGAGTGG	TAAAATTTTA	TGGAAGTTTT	GCATTTAATA	GAAAACGAAA
9121	TGCTCTGGAA	CTGGAATATA	AACAGGACTA	TACATCTCCT	GGAACCTCAG	AATACGTGGG
9181	ACCACTTAAA	GTGACAGTGC	AGGAGTTAGA	TGGATCCTTC	AATCATACAC	TGCAAAATTGA
9241	AGAAAACAGC	CTTAAACATG	ATATACCCTG	CCATTCCAAA	AGTAGAAGGA	ATAAAAAGAA
9301	AAAAATCCCA	CTGATGAATG	GAGAAGAAGT	TGACATGGAT	CTTTCTGCAA	TGGATGCTGA
9361	TTCCCTTTG	CTGTGGATAA	GGATAGACCC	AGATATGTCA	GTATTGAGGA	AGGTAGAATT

Fig. 10 (continuación)

```

9421 TGAGCAAGCT GATTTTATGT GGCAGTATCA GCTCCGCTAT GAGAGAGATG TTGTTGCACA
9481 GCAGGAATCC ATTTTGGCTT TGGAAAAATT CCCTACTCCA GCATCTCGGC TTGCACTCAC
9541 TGATATATTA GAACAAGAGC AGTGTTCCTA CAGAGTAAGA ATGTCAGCTT GCTTCTGTCT
9601 TGCAAAGATT GCAAATTCCA TGGTGAGCAC ATGGACAGGA CCACCAGCCA TGAAGTCACT
9661 CTTACTAGG ATGTTTGTGTT GTAAAAGTTG TCCAAACATT GTGAAAACAA ACAACTTTAT
9721 GAGCTTTCAA AGTTATTTTC TACAGAAGAC TATGCCAGTT GCAATGGCTT TATTAAGAGA
9781 TGTTCATAAT CTTTGTCTTA AAGAAGTCTT AACGTTTATT TTAGACTTAA TCAAGTACAA
9841 TGACAAAGG AAAAATAAGT TTTCAGATAA CTATTATCGT GCAGAAATGA TTGATGCCCT
9901 GGCCAACTCT GTTACACCTG CAGTCAGTGT GAATAATGAA GTTAGAACTT TGGATAACTT
9961 AAATCCTGAT GTGCGACTCA TTCTTGAAGA AATCACCAGA TTTTGAATA TGGAAAAACT
10021 TCTTCCGAT TACAGGCATA CCATCAGTGT CAGTGTGTTG AGAGCATAC TGGTAACTCA
10081 GAAGAACGGA CATGTGCCAA GTGATCCAGC TCTTTTAA TCTTATGCTG AATATGGCCA
10141 CTTTGTGGAC ATTAGGATAG CAGCTTTGGA AGCAGTTGTT GATTATACTA AAGTGGACAG
10201 AAGTTATGAA GAACTGCAAT GGCTACTTAA TATGATTCAG AATGACCCTG TACCCTATGT
10261 AAGGCATAAG ATTCTCAACA TGTTGACTAA GAACCCACCA TTTACTAAGA ACATGGAGTC
10321 TCCCTTATGC AATGAAGCCC TGGTAGATCA ACTTTGAAA CTTATGAATT CTGGTACTTC
10381 ACATGACTGG AGGTTACGGT GTGGTGCTGT GGACTTGATC TTCACACTTT TTGGCCTCAG
10441 TAGACCTTCC TGTTTACCCT TGCCAGAGCT TGGGTGTTT CTTAATCTAA AGGAGAAAAA
10501 AGCTGTCTTG AATCCTACCA TAATTCCAGA GTCAGTAGCA GGCAACCAAG AAGCTGCAAA
10561 TAATCCAAGC AGTCACCCAC AGCTAGTTGG ATTTCAGAAC CCTTTTCCA GTTCTCAAGA
10621 TGAGGAGGAG ATTGATATGG ATACTGTTCA TGATAGCCAG GCCTTCATTT CCCATCATTT
10681 AAACATGCTT GAAAGGCCGT CAACTCCAGG GCTCTCGAAG TATCGGCCAG CTAGCTCCCG
10741 ATCTGCTTTA ATACCCAGC ACTCAGCAGG CTGCGACAGC ACACCCACCA CAAAACCCCA
10801 GTGGAGTTG GAACTTGCA GGAAGGGAAC AGGTAAAGAA CAAGCACCTT TGGAGATGAG
10861 TATGCATCCG GCGGCAAGCG CTCCACTCTC AGTCTTACT AAGGAATCTA CAGCCTCCAA
10921 ACACAGTGAC CACCATCACC ACCATCACCA TGAGCACAAG AAAAGAAGA AGAAGCATAA
10981 ACATAAGCAC AAACACAAGC ATAAGCATGA CAGTAAAGAA AAGGACAAGG AGCCTTTCAC
11041 TTTCTCCAGC CCTGCCAGTG GCAGGTCTAT TCGTTCTCCT TCCCTTTCAG ACCGGTCCGG
11101 CCACCATCAT CACCACCATT GATAAGCTAG CGGCCGCTTT CGAATCTAGA GCCTGCAGTC
11161 TCGACAAGCT TGTGAGAAG TACTAGAGGA TCATAATCAG CCATACCACA TTTGTAGAGG
11221 TTTTACTTGC TTTAAAAAAC CTCCACACC TCCCTTGAA CCTGAAACAT AAAATGAATG
11281 CAATTGTTGT TGTTAACTTG TTTATTGCAG CTTATAATGG TTACAAATAA AGCAATAGCA
11341 TCACAAATTT CACAAATAAA GCATTTTTTT CACTGCATTC TAGTTGTGGT TTGTCCAAAC
11401 TCATCAATGT ATCTTATCAT GTCTGGATCT GATCACTGCT TGAGCCTAGA AGATCCGGCT
11461 GCTAACAAAG CCCGAAAGGA AGCTGAGTTG GCTGTGCCA CCGCTGAGCA ATAACTAGCA
11521 TAACCCCTTG GGGCCTCTAA ACGGGTCTTG AGGGGTTTT TTGCTGAAAG AGGAACTATA
11581 TCCGATCTG AACAGGAGG ACAGCTGATA GAAACAGAAG CCACTGGAGC ACCTCAAAA
11641 CACCATCATA CACTAAATCA GTAAGTTGGC AGCATCACCC GACGCACTTT GCGCCGAATA
11701 AATACCTGTG ACGGAAGATC ACTTCGCAGA ATAAATAAAT CCTGGTGTCC CTGTTGATAC
11761 CGGGAAGCCC TGGGCCAACT TTTGGCGAAA ATGAGACGTT GATCGGCACG TAAGAGGTTT
11821 CAACTTTCAC CATAATGAAA TAAGATCACT ACCGGGCGTA TTTTGTGAGT TATCGAGATT
11881 TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAATCACT GGATATACCA CCGTTGATAT
11941 ATCCCAATGG CATCGTAAAG AACATTTTGA GGCATTTTCA TCAGTTGCTC AATGTACCTA
12001 TAACCAGACC GTTCAGCTGG ATATTACGGC CTTTTTAAAG ACCGTAAAGA AAAATAAGCA
12061 CAAGTTTTAT CCGGCCTTTA TTCACATTCT TGCCCGCTG ATGAATGCTC ATCCGGAATT
12121 CCGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT AGTGTTCACC CTTGTTACAC
12181 CGTTTTCCAT GAGCAAACCTG AAACGTTTTT ATCGCTCTGG AGTGAATACC ACGACGATTT
12241 CCGGCAGTTT CTACACATAT ATTGCGAAGA TGTGGCGTGT TACGGTGAAA ACCTGGCCTA
12301 TTTCCCTAAA GGGTTTATTG AGAATATGTT TTTCTCTCA GCCAATCCCT GGGTGAGTTT
12361 CACCAAGTTT GATTTAAACG TGGCCAATAT GGACAACTTC TTCGCCCCG TTTTACCAT
12421 GGGCAAAATAT TATACGCAAG GCGCAAGGT GCTGATGCCG CTGGCGATT CAGTTTCATCA
12481 TGCCGTCTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAT GAATTACAAC AGTACTGCGA
12541 TGAGTGGCAG GCGGGGCGT AATTTTTTTA AGGCAGTTAT TGGTGCCCTT AAACGCTGG
12601 TGCTACGCCT GAATAAGTGA TAATAAGCGG ATGAATGGCA GAAATTCGAA AGCAAATTCG
12661 ACCGGTCTGT CGGTTTCAGG CAGGGTCTGT AAATAGCCG TATGTCTAT TGCTGGTTTA

```

Fig. 10 (continuación)

```

12721 CCGGTTTATT GACTACCGGA AGCAGTGTGA CCGTGTGCTT CTCAAATGCC TGAGGCCAGT
12781 TTGCTCAGGC TCTCCCCGTG GAGGTAATAA TTGACGATAT GATCATTAT TCTGCCTCCC
12841 AGAGCCTGAC ATTCATCCGG GGTCAGCACC GTTCTGCGG ACTGGCTTTC TACGTGTTC
12901 GCTTCCTTTA GCAGCCCTTG CGCCCTGAGT GCTTGCGGCA GCGTGAAGCT
    
```

Fig. 11

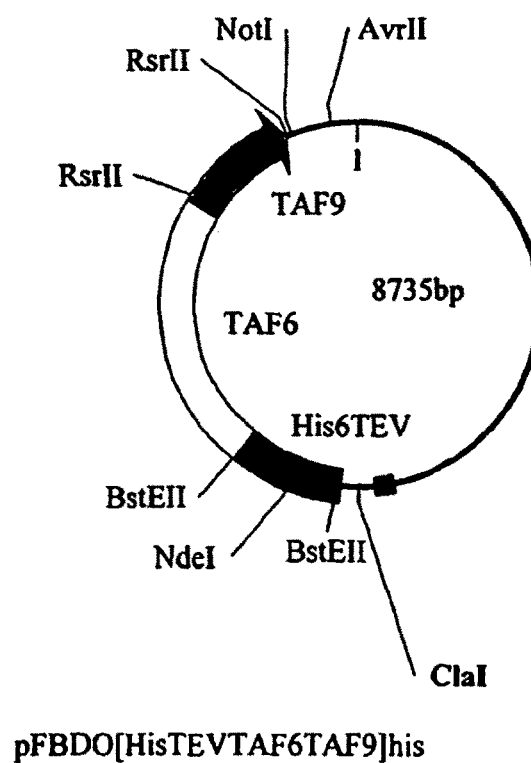


Fig. 12

SEQ pFBDO[HisTEV8032]his: 8735 bp;
 Composición 2196 A; 2272 C; 2120 G; 2147 T; 0 OTROS
 Porcentaje: 25.1% A; 26.0% C; 24.3% G; 24.6% T; 0.0% OTROS
 Peso molecular (kDa): ssDNA: 2694.12 dsDNA: 5385.15
 PLABRA CLAVE CIRCULAR

ORIGEN

```

1      TTCTCTGTCA CAGAATGAAA ATTTTCTGT CATCTCTTCG TTATTAATGT TTGTAATTGA
61     CTGAATATCA ACGCTTATTT GCAGCTGAA TGGCGAATGG GACGCGCCCT GTAGCGGCGC
121    ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT
181    AGCGCCCGCT CCTTTCGCTT TCCTCCCTTC CTTTCTCGCC ACGTTCGCCG GCTTTCGCCG
241    TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA
301    CCCCCAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT
361    TTTTCGCCCT TTGACGTTGG AGTCCACGTT CTTAATAGT GGACTCTTGT TCCAAACTGG
421    AACAACACTC AACCCTATCT CGGTCTATTC TTTTGATTTA TAAGGGATTT TGCCGATTTT
481    GGCCTATTGG TTAATAAATG AGCTGATTTA AAAAAATTT AACGCGAATT TTAACAAAAT
541    ATTAACGTTT ACAATTTTCA GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG
601    TTTATTTTTC TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGTATAAT
661    GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCGGTG TCGCCCTTAT
721    TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGCTCAC CCAGAAACGC TGGTGAAAGT
781    AAAAGATGCT GAAGATCAGT TGGGTGCACG AGTGGGTAC ATCGAACTGG ATCTCAACAG
841    CGGTAAGATC CTTGAGAGTT TTCGCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA
901    AGTTCCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCT
961    CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT
1021   TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACCATGA GTGATAACAC
1081   TCGGCGCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA
1141   CAACATGGGG GATCATGTAA CTCGCCCTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT
1201   ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TCGCGAACT
1261   ATTAACGGC GAACACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC
1321   GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA
1381   TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG
1441   TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG
1501   AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTACAGCCA
1561   AGTTTACTCA TATATACTTT AGATTGATTT AAAACTTCAT TTTTAATTTA AAAGGATCTA
1621   GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCGTTCCA
1681   CGTAATCTGC TGCTTGCAAA CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCCGGA
1741   TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA
1801   TACTGTCTTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC
1861   TACATACCTC GCTCTGTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCTGT
1921   TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CCGGCTGAAC
1981   GGGGGGTTTCG TGCACACAGC CCAGCTTGGG GCGAACGACC TACACCGAAC TGAGATACCT
2041   ACAGCGTGAG CATTGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC
2101   GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCAGGGG GAAACGCCTG
2161   GTATCTTTAT AGTCCTGTCT GGTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG
2221   CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCGGCTTTT TACGGTTCCT
2281   GGCCTTTTGC TGGCCTTTTG CTCACATGTT CTTTCCTGCG TTATCCCTG ATTCTGTGGA
2341   TAACCGTATT ACCGCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
2401   CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCTGATG CGGTATTTTC TCCTTACGCA
2461   TCTGTGCGGT ATTTACACAC GCAGACCAGC CGCGTAACCT GGCAAAATCG GTTACGGTTG
2521   AGTAATAAAT GGATGCCCTG CGTAAGCGGG TGTGGGCGGA CAATAAAGTC TTAAGTGA
2581   CAAAATAGAT CTAAGTATG CTAAGTATG CTTAACTAG ACAGAATAGT TGTAACTGA
2641   AATCAGTCCA GTTATGCTGT AAAAAAGCAT ACTGGACTTT TGTATGGCT AAAGCAAAC
2701   CTTCAATTTT TGAAGTGCAA ATTGCCGCTC GTATTAAAGA GGGGCGTGGC CAAGGGCATG
2761   GTAAAGACTA TATTCGCGGC GTTGTGACAA TTTACCGAAC AACTCCGCGG CCGGGAAGCC

```


Fig. 12 (continuación)

2821	GATCTCGGCT	TGAACGAATT	GTTAGGTGGC	GGTACTTGGG	TCGATATCAA	AGTGCATCAC
2881	TTCTTCCCGT	ATGCCCAACT	TTGTATAGAG	AGCCACTGCG	GGATCGTCAC	CGTAATCTGC
2941	TTGCACGTAG	ATCACATAAG	CACCAAGCGC	GTTGGCCTCA	TGCTTGAGGA	GATTGATGAG
3001	CGCGGTGGCA	ATGCCCTGCC	TCCGGTGCTC	GCCGGAGACT	GCGAGATCAT	AGATATAGAT
3061	CTCACTACGC	GGCTGCTCAA	ACCTGGGCAG	AACGTAAGCC	GCGAGAGCGC	CAACAACCGC
3121	TTCTTGGTCG	AAGGCAGCAA	GCGCGATGAA	TGTCTTACTA	CGGAGCAAGT	TCCCGAGGTA
3181	ATCGGAGTCC	GGCTGATGTT	GGGAGTAGGT	GGCTACGTCT	CCGAACTCAC	GACCGAAAAG
3241	ATCAAGAGCA	GCCCCGATGG	ATTTGACTTG	GTCAGGGCCG	AGCCTACATG	TGCGAATGAT
3301	GCCCATACTT	GAGCCACCTA	ACTTTGTTTT	AGGGCGACTG	CCCTGCTGCG	TAACATCGTT
3361	CTTGCTGCTT	GGATGCCCGA	GGCATAGACT	GTACAAAAAA	ACAGTCATAA	CAAGCCATGA
3421	AAACGCGCCAC	TGCGCCGTTA	CCACCGCTGC	GTTCCGTCOA	GGTTCTGGAC	CAGTTGCGTG
3481	AGCGCATACG	CTACTTGCAAT	TACAGTTTAC	GAACCGAACA	GGCTTATGTC	AAGTGGGTTT
3541	GTGCCTTCAT	CCGTTTCCAC	GGTGTGCGTC	ACCCGGCAAC	CTTGGGCAGC	AGCGAAGTCG
3601	AGGCATTTCT	GTCCTGGCTG	GCGAACGAGC	GCAAGGTTTC	GGTCTCCACG	CATCGTCAGG
3661	CATGGCGGGC	CTTGCTGTTC	TTCTACGGCA	AGGTGCTGTG	CACGGATCTG	CCCTGGCTTC
3721	AGGAGATCGG	TAGACCTCGG	CCGTGCGGGC	GCTTGCCGGT	GGTGTGACCC	CCGGATGAAG
3781	TGGTTCGCAT	CCTCGGTTTT	CTGGAAGGCG	AGCATCGTTT	GTTGCGCCAG	GACTCTAGCT
3841	ATAGTTCTAG	TGGTTGGCCT	ACAGCTTTGT	TTAAACAAAG	CTGTACCCGT	AGTGGCTATG
3901	GCAGGGCTTG	CCGCCCCGAC	GTTGGCTGCG	AGCCCTGGGC	CTTCACCCGA	ACTTGGGGGT
3961	TGGGGTGGGG	AAAAGGAAGA	AACGCGGGCG	TATTGGTCCC	AATGGGGTCT	CGGTGGGGTA
4021	TCGACAGAGT	GCCAGCCCTG	GGACCGAACC	CCGCGTTTAT	GAACAAACGA	CCCAACACCC
4081	GTGCGTTTTA	TTCTGTCTTT	TTATTGCCCT	CATAGCGCGG	GTTCCCTCCG	GTATTGTCTC
4141	CTTCCGTGTT	TCAGTTAGCC	TCCCCATCT	CCCGGTACCG	CATGCTATGC	ATCAGCTGCT
4201	AGCAACATGG	CTGAGATCC	CGGGTGATCA	AGTCTTCGTC	GAGTGATTGT	AAATAAAATG
4261	TAATTTACAG	TATAGTATTT	TAATTAATAT	ACAAATGATT	TGATAATAAT	TCTTATTTAA
4321	CTATAATATA	TTGTGTTGGG	TTGAATTAAA	GGTCCGTATA	CTAGTATCGA	TTCGCGACCT
4381	ACTCCGGAAT	ATTAATAGAT	CATGGAGATA	ATTAATAATGA	TAACCATCTC	GCAAATAAAT
4441	AAGTATTTTA	CTGTTTTTCG	AACAGTTTTG	TAATAAAAAA	ACCTATAAAT	ATTCCGGAAT
4501	ATTCATACCG	TCCCACCATC	GGGCGCGGAT	CCTCGAGATG	GGTAACCATC	ATCATCATCA
4561	TCACGGAGAA	AGCTTGTTTA	AGGGACCACG	TGATTACAAC	CCGATATCGA	GCACCATTGG
4621	TCATTTGACG	AATGAATCTG	ATGGGCACAC	AACATCGTTG	TATGGTATTG	GATTGGTCC
4681	CTTCATCATT	ACAAACAAGC	ACTTGTTTAG	AAGAAATAAT	GGAACACTGT	TGGTCCAATC
4741	ACTACATGGT	GTATTCAAGG	TCAAGAACAC	CACGACTTTG	CAACAACACC	TCATTGATGG
4801	GAGGGACATG	ATAATTATTC	GCATGCCTAA	GGATTTCCTA	CCATTTCCTC	AAAAGCTGAA
4861	ATTTAGAGAG	CCACAAAGGG	AAGAGCGCAT	ATGTCTTGTC	ACAACCAACT	TCCAACTAA
4921	GAGCATGTCT	AGCATGGTGT	CAGACACTAG	TTGCACATTC	CCTTCATCTG	ATGGCATATT
4981	CTGGAAGCAT	TGGATTCAAA	CCAAGGATGG	GCAGTGTGGC	AGTCCATTAG	TATCAACTAG
5041	AGATGGGTTT	ATTGTTGGTA	TACACTCAGC	ATCGAATTTT	ACCAACACAA	ACAATTATTT
5101	CACAAGCGTG	CCGAAAAACT	TCATGGAATT	GTTGACAAAT	CAGGAGGCGC	AGCAGTGGGT
5161	TAGTGGTTGG	CGATTAAATG	CTGACTCAGT	ATTGTGGGGG	GGCCATAAAG	TTTTCATGAG
5221	CAAACTGAA	GAGCCTTTTC	AGCCAGTTAA	GGAAGCGACT	CAACTCATGA	ATGAATTGGT
5281	GTAATCGCAA	GGTGGTGGTG	AAAACCTGTA	CTTCCAGGGT	AACCACGCTG	AGGAGAAGAA
5341	GCTGAAGCTT	AGCAACACTG	TGCTGCCCTC	GGAGTCCATG	AAGGTGGTGG	CTGAATCCAT
5401	GGGCATCGCC	CAGATTCAGG	AGGAGACCTG	CCAGTGCTTA	ACGGATGAGG	TCAGCTACCG
5461	CATCAAAGAG	ATCGCACAGG	ATGCCTTGAA	GTTTATGCAC	ATGGGGAAGC	GGCAGAAGCT
5521	CACCACCAGT	GACATTGACT	ACGCCTTGAA	GCTAAAGAAT	GTCGAGCCAC	TCTATGGCTT
5581	CCACGCCAG	GAGTTTATT	CTTTCCGCTT	CGCCTCTGGT	GGGGGCGGGG	AGCTTTACTT
5641	CTATGAGGAG	AAGGAGGTTG	ATCTGAGCGA	CATCATCAAT	ACCCCTCTGC	CCCGGGTGCC
5701	CTGGGACGTC	TGCCTCAAAG	CTCATTGGCT	GAGCATCGAG	GGCTGCCAGC	CAGCTATCCC
5761	CGGGAACCCG	CCCCAGCTC	CCAAAGAGCA	ACAGAAGGCT	GAAGCCACAG	AACCCCTGAA
5821	GTCAGCCAAG	CCAGGCCAGG	AGGAAGACGG	ACCCCTGAAG	GGCAAGGTC	AAGGGGCCAC
5881	CACAGCCGAC	GGCAAAGGGA	AAGAGAAGAA	GGCGCCGCC	TTGCTGGAGG	GGGCCCCCTT
5941	GCGACTGAAG	CCCCGGAGCA	TCCACGAGTT	GTCTGTGGAG	CAGCAGCTCT	ACTACAAGGA
6001	GATCACCGAG	GCCTGCGTGG	GCTCCTGCGA	GGCCAAGAGG	GCGGAAGCCC	TGCAAAGCAT
6061	TGCCACGGAC	CCTGGACTGT	ATCAGATGCT	GCCACGGTTC	AGTACCTTTA	TCTCGGAGGG

Fig. 12 (continuación)

6121	GGTCCGTGTG	AACGTGGTTC	AGAACAACCT	GGCCCTACTC	ATCTACCTGA	TGCGTATGGT
6181	GAAAGCGCTG	ATGGACAACC	CCACGCTCTA	TCTAGAAAAA	TACGTCCATG	AGCTGATTCC
6241	AGCTGTGATG	ACCTGCATCG	TGAGCAGACA	GTTGTGCCTG	CGACCAGATG	TGGACAAATCA
6301	CTGGGCACTC	CGAGACTTTG	CTGCCCCGCT	GGTGGCCCCAG	ATCTGCAAGC	ATTTTAGCAC
6361	AACCACTAAC	AACATCCAGT	CCCGGATCAC	CAAGACCTTC	ACCAAGAGCT	GGGTGGACGA
6421	GAAGACGCCC	TGGACGACTC	GTTATGGCTC	CATCGCAGGC	TTGGCTGAGC	TGGGACACGA
6481	TGTTATCAAG	ACTCTGATTC	TGCCCCGGCT	GCAGCAGGAA	GGGGAGCGGA	TCCGCAGTGT
6541	GCTGGACGGC	CCTGTGCTGA	GCAACATTGA	CCGGATTGGA	GCAGACCATG	TGCAGAGCCT
6601	CCTGCTGAAA	CACTGTGCTC	CTGTTCTGGC	AAAGCTGCGC	CCACCGCCTG	ACAATCAGGA
6661	CGCCTATCGG	GCAGAATTCC	GGTCCCTTGG	GCCCCCTCTC	TGCTCCCAGG	TGGTCAAGGC
6721	TCGGGCCCCAG	GCTGCTCTGC	AGGCTCAGCA	GGTCAACAGG	ACCACTCTGA	CCATCAGCA
6781	GCCCCGGCCC	ACGCTGACCC	TCTCGCAGGC	CCCACAGCCT	GGCCCTCGCA	CCCCTGGCTT
6841	GTGAAGGTT	CCTGGCTCCA	TGCACCTTCC	TGTCCAGACA	CTGGTGTCTG	CACGAGCGGC
6901	TGCCCCACCA	CAGCCTTCCC	CTCCTCCAAC	CAAGTTTATT	GTAATGTCTG	CGTCTCCAG
6961	CGCCCCATCC	ACCCAGCAGG	TCCTGTCCCT	CAGCACCTCG	GCCCCCGGCT	CAGGTTCCAC
7021	CACCACTTCG	CCCGTCACCA	CCACCGTCCC	CAGCGTGCAG	CCCATCGTCA	AGTTGGTCTC
7081	CACCGCCACC	ACCGCACCCC	CCAGCACTGC	TCCCTCTGGT	CCTGGGAGTG	TCCAGAAGTA
7141	CATCGTGGTC	TCACTTCCCC	CAACAGGGGA	GGGCAAAGGA	GGCCCCACCT	CCCATCCTTC
7201	TCCAGTTCCT	CCCCCGGCAT	CGTCCCCGTC	CCCACTCAGC	GGCAGTGCCC	TTTGTGGGGG
7261	GAAGCAGGAG	GCTGGGGACA	GTCCCCCTCC	AGCTCCAGGG	ACTCCAAAAG	CCAATGGCTC
7321	CCAGCCCAAC	TCCGGCTCCC	CTCAGCCTGC	TCCGCGGTCC	GGTGGTGGTG	GTGAAAACCT
7381	GTATTTTCAG	GGCGAGTCTG	GCAAGACGGC	TTCTCCCAAG	AGCATGCCGA	AAGATGCACA
7441	GATGATGGCA	CAATCCTGA	AGGATATGGG	GATTACAGAA	TATGAGCCAA	GAGTTATAAA
7501	TCAGATGTTG	GAGTTTGCCT	TCCGATATGT	GACCACAATT	CTAGATGATG	CAAAAATTTA
7561	TTCAAGCCAT	GCTAAGAAAG	CTACTGTTGA	TGCAGATGAT	GTGCGATTGG	CAATCCAGTG
7621	CCGCGCTGAT	CAGTCTTTTA	CCTCTCCTCC	CCCAAGAGAT	TTTTTATTAG	ATATTGCAAG
7681	GCAAAGAAAT	CAAACCCCTT	TGCCATTGAT	CAAGCCATAT	TCAGGTCCAA	GGTTGCCACC
7741	TGATAGATAC	TGCTTAACAG	CTCCAAACTA	TAGGCTGAAA	TCTTTACAGA	AAAAGGCATC
7801	AACTTCTGCG	GGAAGAATAA	CAGTCCCAGG	GTAAAGTGTT	GGTTCAGTTA	CTAGCAGACC
7861	AAGTACTCCC	ACACTAGGCA	CACCAACCCC	ACAGACCATG	TCTGTTTCAA	CTAAAGTAGG
7921	GACTCCCATG	TCCCTCACAG	GTCAAAGGTT	TACAGTACAG	ATGCCTACTT	CTCAGTCTCC
7981	AGCTGTAAAA	GCTTCAATTC	CTGCAACCTC	AGCAGTTCAG	AATGTTCTGA	TTAATCCATC
8041	ATTAATCGGG	TCCAAAAACA	TTCTTATTAC	CACTAATATG	ATGTCATCAC	AAAATACTGC
8101	CAATGAATCA	TCAAATGCAT	TGAAAAGAAA	ACGTGAAGAT	GATGATGATG	ACGATGATGA
8161	TGATGATGAC	TATGATAATC	TGCGGTCCGG	CCACCATCAT	CACCACCATT	GATAAGCTAG
8221	CGGCCGCTTT	CGAATCTAGA	GCCTGCAGTC	TCGACAAGCT	TGTCGAGAAG	TACTAGAGGA
8281	TCATAATCAG	CCATACCACA	TTGTAGAGG	TTTTACTTGC	TTTAAAAAAC	CTCCACACC
8341	TCCCCCTGAA	CCTGAAACAT	AAAATGAATG	CAATTGTTGT	TGTTAACTTG	TTTATTGCAG
8401	CTTATAATGG	TTACAAATAA	AGCAATAGCA	TCACAAATT	CACAAATAAA	GCATTTTTTT
8461	CACTGCATTC	TAGTTGTGGT	TTGTCCAAAC	TCATCAATGT	ATCTTATCAT	GTCTGGATCT
8521	GATCACTGCT	TGAGCCTAGG	AGATCCGAAC	CAGATAAGTG	AAATCTAGTT	CCAAACTATT
8581	TTGTCATTTT	TAATTTTCGT	ATTAGCTTAC	GACGCTACAC	CCAGTCCCA	TCTATTTTGT
8641	CACTCTTCCC	TAAATAATCC	TTAAAACTC	CATTTCCACC	CCTCCCAGTT	CCCAACGCCA
8701	ACTCCATGTG	ACAAACCGTC	ATCTTCGGCT	ACTTT		

ES 2 324 495 T3

LISTA DE SECUENCIAS

<110> Eidgenössische Technische Hochschule Zürich
 Berger, Imre
 Fitzgerald, Daniel J.
 Richmond, Timothy J.
 <120> Expresión recombinante de complejos de multiproteínas utilizando poligenes
 <130> T-05-019
 <150> EP 05024271.8
 <151> 2005-11-08
 <160> 5
 <170> PatentIn version 3.1
 <210> 1
 <211> 607
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Producto PCR que codifica hTBPC
 <400> 1

```

    cgaattcctc gagcgggtccg gaggtaacgg atccgaaaac ctgtattttc aggggttctgg      60
    gattgtaccg cagctgcaaa atattgtatc cacagtgaat cttgggttgta aacttgacct      120
    aaagaccatt gcacttcgtg cccgaaacgc cgaatataat cccaagcggg ttgctgcggg      180
    aatcatgagg ataagagagc cacgaaccac ggcactgatt ttcagttctg ggaaaatggg      240
    gtgcacagga gccaagagtg aagaacagtc cagactggca gcaagaaaat atgctagagt      300
    tgtacagaag ttgggttttc cagctaagtt cttggacttc aagattcaga acatggtggg      360
    gagctgtgat gtgaagtttc ctataagggt agaaggcctt gtgctcacc accaacaatt      420
    tagtagttat gagccagagt tatttcctgg tttaatctac agaatgatca aaccagaaat      480
    tgttctcctt atttttggtt ctggaaaagt tgtattaaca ggtgctaaag tcagagcaga      540

    aatttatgaa gcatttgaaa acatctaccc tattctaaag ggattcagga agacgacgcg      600
    gtccggc                                           607
  
```

<210> 2
 <211> 196
 <212> PRT
 <213> *Homo sapiens*

ES 2 324 495 T3

<400> 2

5 Arg Ser Gly Gly Asn Gly Ser Glu Asn Leu Tyr Phe Gln Gly Ser Gly
 1 5 10 15
 10 Ile Val Pro Gln Leu Gln Asn Ile Val Ser Thr Val Asn Leu Gly Cys
 20 25 30
 15 Lys Leu Asp Leu Lys Thr Ile Ala Leu Arg Ala Arg Asn Ala Glu Tyr
 35 40 45
 20 Asn Pro Lys Arg Phe Ala Ala Val Ile Met Arg Ile Arg Glu Pro Arg
 50 55 60
 25 Thr Thr Ala Leu Ile Phe Ser Ser Gly Lys Met Val Cys Thr Gly Ala
 65 70 75 80
 30 Lys Ser Glu Glu Gln Ser Arg Leu Ala Ala Arg Lys Tyr Ala Arg Val
 85 90 95
 35 Val Gln Lys Leu Gly Phe Pro Ala Lys Phe Leu Asp Phe Lys Ile Gln
 100 105 110
 40 Asn Met Val Gly Ser Cys Asp Val Lys Phe Pro Ile Arg Leu Glu Gly
 115 120 125
 45 Leu Val Leu Thr His Gln Gln Phe Ser Ser Tyr Glu Pro Glu Leu Phe
 130 135 140
 50 Pro Gly Leu Ile Tyr Arg Met Ile Lys Pro Arg Ile Val Leu Leu Ile
 145 150 155 160
 55 Phe Val Ser Gly Lys Val Val Leu Thr Gly Ala Lys Val Arg Ala Glu
 165 170 175
 60 Ile Tyr Glu Ala Phe Glu Asn Ile Tyr Pro Ile Leu Lys Gly Phe Arg
 180 185 190
 65 Lys Thr Thr Arg
 195

<210> 3

60 <211> 6959

<212> ADN

<213> Secuencia artificial

<220>

65 <223> Plásmido pFBDO[hTBPC]3

ES 2 324 495 T3

<400> 3

	ttctctgtca cagaatgaaa atttttctgt catctcttcg ttattaatgt ttgtaattga	60
5	ctgaatatca acgcttattt gcagcctgaa tggcgaatgg gacgcgccct gtagcggcgc	120
	attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg ccagcgccct	180
	agcgcgccgt cctttcgctt tcttcccttc ctttctcgcc acgttcgccg gctttccccg	240
10	tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac ggcacctcga	300
	ccccaaaaa cttgattagg gtgatggttc acgtagtggg ccatcgccct gatagacggg	360
15	ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaactgg	420
	aacaacactc aaccctatct cggctctattc ttttgattta taagggattt tgccgatttc	480
	ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat	540
20	attaacgttt acaatttcag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg	600
	tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat	660
	gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat	720
25	tccctttttt gcggcatttt gccttctgt ttttgctcac ccagaaacgc tggtgaaagt	780
	aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag	840
30	cggtaaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcactttttaa	900
	agttctgcta tgtggcgcgg tatttatcccg tattgacgcc gggcaagagc aactcgggtcg	960
	ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag aaaagcatct	1020
35	tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga gtgataacac	1080
	tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca	1140
	caacatgggg gatcatgtaa ctgccttga tcggtgggaa ccggagctga atgaagccat	1200
40	accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact	1260
	attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc	1320
	ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggg ttattgctga	1380
45	taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg	1440
	taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg	1500
50	aaatagacag atcgtgaga taggtgcctc actgattaag cattggtaac tgtcagacca	1560

55

60

65

ES 2 324 495 T3

	agtttactca tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta	1620
	ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca	1680
5	cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga	1740
	tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa	1800
	ractgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc	1860
10	tacatacctc gctctgctaa tcctgttacc agtggtgct gccagtggcg ataagtcgtg	1920
	tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggc cgggctgaac	1980
15	ggggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct	2040
	acagcgtgag cattgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggatatcc	2100
	ggtaagcggc agggctcgaa caggagagcg cagagggag cttccagggg gaaacgcctg	2160
20	gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg	2220
	ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct	2280
	ggccttttgc tggccttttg ctcacatggt ctttctgctg ttatcccctg attctgtgga	2340
25	taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg	2400
	cagcagtcga gtgagcgagg aagcgggaaga gcgcctgatg cggatatttc tccttacgca	2460
30	tctgtgcggt atttcacacc gcagaccagc cgcgtaacct ggcaaaatcg gttacggttg	2520
	agtaataaat ggatgccctg cgtaagcggg tgtgggcgga caataaagtc ttaaactgaa	2580
	caaaatagat ctaaaactatg acaataaagt cttaaactag acagaatagt tgtaaactga	2640
35	aatcagtcga gttatgctgt gaaaaagcat actggacttt tgttatggct aaagcaaaact	2700
	cttcattttc tgaagtgcaa attgcccgtc gtattaaaga ggggcgtggc caagggcatg	2760
	gtaaagacta tattcgcggc gttgtgacaa tttaccgaac aactccgcgg ccgggaagcc	2820
40	gatctcggct tgaacgaatt gttaggtggc ggtacttggg tcgatatcaa agtgcacac	2880
	ttcttcccgt atgcccact ttgtatagag agccactgcg ggatcgtcac cgtaatctgc	2940
	ttgcacgtag atcacataag caccaagcgc gttggcctca tgcttgagga gattgatgag	3000
45	cgcggtggca atgcccgtcc tccggtgctc gccggagact gcgagatcat agatatagat	3060
	ctcactacgc ggctgctcaa acctgggcag aacgtaagcc gcgagagcgc caacaaccgc	3120
50	ttcttggtcg aaggcagcaa gcgcgatgaa tgtcttacta cggagcaagt tcccagggtg	3180
	atcggagtcc ggctgatggt gggagtaggt ggctacgtct ccgaactcac gaccgaaaag	3240
	atcaagagca gcccgcattg atttgacttg gtcagggccg agcctacatg tgcgaatgat	3300
55	gcccatactt gagccacctt actttgtttt agggcgactg ccctgctgcg taacatcggt	3360
	cttgctgctt ggatgcccga ggcatagact gtacaaaaaa acagtcataa caagccatga	3420
	aaaccgccac tgcgccgtta ccaccgtgc gttcgggtcaa ggttctggac cagtgcgtg	3480
60	agcgcatacg ctacttgcac tacagtttac gaaccgaaca ggcttatgtc aactgggttc	3540
	gtgccttcat ccgtttccac ggtgtgcgtc acccggcaac cttgggcagc agcgaagtcg	3600

65

ES 2 324 495 T3

	aggcatttct gtcctggctg gcgaacgagc gcaagggttc ggtctccacg catcgtcagg	3660
	cattggcggc cttgctgttc ttctacggca aggtgctgtg cacggatctg ccctggcttc	3720
5	aggagatcgg tagacctcgg ccgtcgcggc gcttgccggt ggtgctgacc ccgatgaag	3780
	tggttcgcat cctcggtttt ctggaaggcg agcatcgttt gttcgcccag gactctagct	3840
	atagttctag tggttggcct acagctttgt ttaaacaag ctgtaccctg agtggctatg	3900
10	gcagggcttg ccgccccgac gttggctgcg agccctgggc cttcaccga acttgggggt	3960
	tggggtgggg aaaaggaaga aacgcgggcg tattggtccc aatggggtct cggtggggta	4020
15	tcgacagagt gccagccctg ggaccgaacc ccgcgtttat gaacaaacga cccaacaccc	4080
	gtgcgtttta ttctgtcttt ttattgccgt catagcgcgg gttccttccg gtattgtctc	4140
	cttcctgtgt tcagtttagcc tccccatct cccgggtaccg catgctatgc atcagctgct	4200
20	agcaccatgg ctcgagatcc cgggtgatca agtcttcgtc gagtgattgt aaataaaatg	4260
	taattttacag tatagtatct taattaatat acaaagtatt tgataataat tcttatttaa	4320
	ctataatata ttgtgttggg ttgaattaaa ggtccgtata ctagtatcga ttcgcgacct	4380
25	actccggaat attaatagat catggagata attaaaatga taaccatctc gcaaataaat	4440
	aagtatttta ctgttttcgt aacagttttg taataaaaaa acctataaat attccggatt	4500
	attcataccg tcccaccatc gggcgcggat cctcagatg ggtaaccatg acaagcgacg	4560
30	atggaaaaag aatttcatag ccgtctcagc agccaaccgc ttttaagaaa tctcatcctc	4620
	cggggcagct agctggagcc acccgcagtt cgaaaaaggc gccgacgacg acgacgacaa	4680
35	gggtctccat atgtctggga ttgtaccgca gctgcaaaat attgtatcca cagtgaatct	4740
	tggttgtaaa cttgacctaa agaccattgc acttcgtgcc cgaaacgccg aatataatcc	4800
	caagcggttt gctgcggtaa tcatgaggat aagagagcca cgaaccacgg cactgatttt	4860
40	cagttctggg aaaatggtgt gcacaggagc caagagtga gaacagtcca gactggcagc	4920
	aagaaaatat gctagagttg tacagaagtt ggggtttcca gctaagttct tggacttcaa	4980
	gattcagaac atggtgggga gctgtgatgt gaagtttcct ataaggttag aaggccttgt	5040
45	gctcaccac caacaattta gtagttatga gccagagtta tttcctgggt taatctacag	5100
	aatgatcaaa ccagaattg ttctccttat ttttgtttct ggaaaagttg tattaacagg	5160
	tgctaaagtc agagcagaaa tttatgaagc atttgaaaac atctacccta ttctaaaggg	5220
50	attcaggaag acgacgcggt ccggaggtaa cggatccgaa aacctgtatt ttcagggttc	5280
	tgggattgta ccgcagctgc aaaatattgt atccacagtg aatcttggtt gtaaaactga	5340
55	cctaaagacc attgcacttc gtgcccgaaa cgccgaatat aatcccaagc ggtttgctgc	5400
	ggtaatcatg aggataagag agccacgaac cacggcactg attttcagtt ctgggaaaat	5460
	ggtgtgcaca ggagccaaga gtgaagaaca gtccagactg gcagcaagaa aatatgctag	5520
60	agttgtacag aagttgggtt ttccagctaa gttcttggtt ttcaagattc agaacatggt	5580
	ggggagctgt gatgtgaagt ttcctataag gttagaaggc cttgtgctca cccaccaaca	5640

ES 2 324 495 T3

	atttagtagt tatgagccag agttatttcc tggtttaatc tacagaatga tcaaaccag	5700
	aattgttctc cttatttttg tttctggaaa agttgtatta acaggtgcta aagtcagagc	5760
5	agaaatttat gaagcatttg aaaacatcta ccctattcta aagggattca ggaagacgac	5820
	gcggtccgga ggtaacggat ccgaaaacct gtattttcag ggttctggga ttgtaccgca	5880
	gctgcaaaat attgtatcca cagtgaatct tggttgtaaa cttgacctaa agaccattgc	5940
10	acttcgtgcc cgaaacgccg aatataatcc caagcgggtt gctgcggtaa tcatgaggat	6000
	aagagagcca cgaaccacgg cactgatttt cagttctggg aaaatgggtg gcacaggagc	6060
15	caagagtga gaacagtcca gactggcagc aagaaaatat gctagagttg tacagaagtt	6120
	gggttttcca gctaagtctt tggacttcaa gattcagaac atgggtgggga gctgtgatgt	6180
	gaagtttctc ataaggttag aaggccttgt gctcaccac caacaattta gtagttatga	6240
20	gccagagtta tttcctgggt taatctacag aatgatcaaa cccagaattg ttctccttat	6300
	ttttgttctt ggaaaagttg tattaacagg tgctaaagtc agagcagaaa tttatgaagc	6360
	atttgaaaac atctacccta ttctaaaggg attcaggaag acgacgcggt ccggccacca	6420
25	tcatcaccac cattgataag ctagcggccg ctttcgaatc tagagcctgc agtctcgaca	6480
	agcttgctga gaagtactag aggatcataa tcagccatac cacatttgta gaggttttac	6540
	ttgttttaaa aaacctccca cacctccccc tgaacctgaa acataaaatg aatgcaattg	6600
30	ttgttggtta cttgtttatt gcagcttata atggttacia ataaagcaat agcatcacia	6660
	atttcacaaa taaagcattt ttttactgct attctagttg tggtttgctc aaactcatca	6720
35	atgtatctta tcatgtctgg atctgatcac tgcttgagcc taggagatcc gaaccagata	6780
	agtgaatct agttccaaac tattttgtca tttttaattt tcgtattagc ttacgacgct	6840
	acaccagtt cccatctatt ttgtcactct tccctaaata atccttaaaa actccatttc	6900
40	caccctccc agttcccaac gccaaactcca tgtgacaaac cgtcatcttc ggctacttt	6959

<210> 4

45 <211> 12950

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Plásmido pUCDMSTAF1TBPCTAF2

<400> 4

55	aattctgtca gccgttaagt gttcctgtgt cactgaaaat tgctttgaga ggctctaagg	60
	gcttctcagt gcgttacatc cctggcttgt tgtccacaac cgttaaacct taaaagcttt	120
	aaaagcctta tatattcttt tttttcttat aaaacttaaa accttagagg ctatttaagt	180
60	tgctgattta tattaatttt attgttcaaa catgagagct tagtacgtga aacatgagag	240

65

ES 2 324 495 T3

	cttagtacgt tagccatgag agcttagtac gtttagccatg agggtttagt tcgttaaaca	300
	tgagagctta gtacgttaaa catgagagct tagtacgtga aacatgagag ctttagtacgt	360
5	actatcaaca ggttgaactg ctgatcaaca gatcctctac gcggccgcgg taccataact	420
	tcgtatagca tacattatac gaagttaact ggagtacccg tagtggctat ggcagggctt	480
	gcccggcga cggtggctgc gagccctggg ccttcacccg aacttggggg ttgggggtggg	540
10	gaaaaggaag aaacgcgggc gtattggtcc caatggggtc tcggtggggg atcgacagag	600
	tgccagccct gggaccgaac cccgcgttta tgaacaaacg acccaacacc cgtgcgtttt	660
15	attctgtctt tttattgccg tcatagcgcg ggttccttcc ggtattgtct ccttcctgtg	720
	ttcagttagc ctccccatc tcccgggtacc gcatgctatg catcagctgc tagcaccatg	780
	gctcgagatc ccgggtgatc aagtcttcgt cgagtgattg taaataaaat gtaatttaca	840
20	gtatagtatt ttaattaata tacaaatgat ttgataataa ttcttattta actataatat	900
	attgtgttgg gttgaattaa aggtccgtat actagtatcg attcgcgacc tactccggaa	960
	tattaataga tcatggagat aattaaaatg ataaccatct cgcaaataaa taagtatttt	1020
25	actgttttcg taacagtttt gtaataaaaa aacctataaa tattccggat tattcatacc	1080
	gtcccaccat cgggcgcgga tcctcgagat gggtaaccat gacaagcgac gatggaaaaa	1140
30	gaatttcata gccgtctcag cagccaaccg ctttaagaaa atctcatcct ccggggcagc	1200
	tagctggagc caccgcgagt tcgaaaaagg cgccgacgac gacgacgaca agggctccca	1260
	tatgggaccc ggctgcgatt tgctgctgcg gacagcagct accatcactg ctgccgccat	1320
35	catgtcgac acggacagcg acgaagattc cgctggaggc ggcccathtt ctttagcggg	1380
	tttccttttc ggcaacatca atggagccgg gcagctggag ggggaaagcg tcttgatga	1440
	tgaatgtaag aagcacttgg caggcttggg ggctttgggg ctgggcagcc tgatcactga	1500
40	actcacggca aatgaagaat tgaccgggac tgacggtgcc ttggtaaatg atgaagggtg	1560
	ggttaggagt acagaagatg ctgtggacta ttcagacatc aatgagggtg cagaagatga	1620
	aagccgaaga taccagcaga cgatggggag cttgcagccc ctttgccact cagattatga	1680
45	tgaagatgac tatgatgctg attgtgaaga cattgattgc aagttagtgc ctctccacc	1740
	tccacccccg ggaccaatga agaaggataa ggaccaggat tctattactg gtgagaaagt	1800
50	ggacttcagt agttcctctg actcagaatc tgagatggga cctcaggaag caacacaggc	1860
	agaatctgaa gatggaaagc tgacccttcc attggctggg attatgcagc atgatgccac	1920
	caagctgttg ccaagtgtca cagaactttt tccagaattt cgacctggaa aggtgttacg	1980
55	ttttctacgt ctttttggac cagggaaagaa tgtcccatct gtttggcggg gtgctcggag	2040
	aaagaggaag aagaagcacc gtgagctgat acaggaagag cagatccagg aggtggagtg	2100
	ctcagtagaa tcagaagtca gccagaagtc tttgtggaac tacgactacg ctccaccacc	2160
60	acctccagag cagtgtctct ctgatgatga aatcacgatg atggctcctg tggagtccaa	2220
	attttcccaa tcaactggag atatagataa agtgacagat accaaaccaa gagtggctga	2280

65

ES 2 324 495 T3

	gtggcgttat gggcctgccc gactgtggtg tgatatgctg ggtgtccctg aagatggcag	2340
	tgggtttgac tatggcttca aactgagaaa gacagaacat gaacctgtga taaaatctag	2400
5	aatgatagag gaatttagga aacttgagga aaacaatggc actgatcttc tggctgatga	2460
	aaacttcctg atggtgacac agctgcattg ggaggatgat atcatctggg atggggagga	2520
	tgtaaacac aaagggacaa aacctcagcg tgcaagcctg gcaggctggc ttccttctag	2580
10	catgactagg aatgcgatgg cttacaatgt tcagcaaggc tttgcagcca ctcttgatga	2640
	tgacaaacct tggtagtcca tttttcccat tgacaatgag gatctggtat atggacgctg	2700
15	ggaggacaat atcatttggg atgctcaggg catgccccgg ctgttggaac ctctgtttt	2760
	gacacttgat cccaatgatg agaacctcat tttggaaatt cctgatgaga aggaagaggc	2820
	cacctctaac tccccctcca aggagagtaa gaaggaaatca tctctgaaga agagtcgaat	2880
20	tctcttaggg aaaacaggag tcatcaagga ggaaccacag cagaacatgt ctacgccaga	2940
	agtgaagat ccatggaatc tctccaatga tgagtattat tatccaagc aacagggtct	3000
	tcgaggcacc tttggagggg atattatcca gcattcaatt cctgctgtgg aattacggca	3060
25	gcccttcttt cccaccaca tggggcccat caaactccgg cagttccatc gccacctct	3120
	gaaaaagtac tcatttggtg cactttctca gccaggctcc cactcagtc aacctttgct	3180
	aaagcacatc aaaaaaagg ccaagatgag agaacaagag aggcaagctt cagggtggtg	3240
30	agagatgttt tttatgcgca cacctcagga cctcacaggc aaagatggtg atcttattct	3300
	tgagaatat agtgaggaaa atggaccctt aatgatgcag gttggcatgg caaccaagat	3360
35	aaagaactat tataaacgga aacctggaaa agatcctgga gcaccagatt gtaaatatgg	3420
	ggaaactggt tactgccata catctccttt cctgggttct ctccatcctg gccaatgct	3480
	gcaagcattt gagaacaacc ttttctgtgc tccaatttat cttcataaga tgccagaaac	3540
40	tgatttcttg atcattcgga caagacaggg ttactatatt cgggaattag tggatatttt	3600
	tgtggttggc cagcagtgct ccttgtttga agttcctggg cctaactcca aaagggccaa	3660
	tacgcatatt cgagactttc tacagggttt tatttaccgc ctttctgga aaagtaaaga	3720
45	tcggccacgg aggatacgaa tggaagatat aaaaaagcc tttccttccc attcagaaag	3780
	cagcatccgg aagaggctaa agctctgctg tgacttcaaa cgcacaggga tggactcaaa	3840
50	ctggtgggtg cttaagtctg attttctgtt accaacggaa gaagagatca gagctatggt	3900
	gtcaccagag cagtgtgtg cttattatag catgatagct gcagagcaac gactgaagga	3960
	tgctggctat ggtgagaaat ctttttttgc tccagaagaa gaaaatgagg aagatttcca	4020
55	gatgaagatt gatgatgaag ttcgcactgc cccttggaa accacaaggg ccttcattgc	4080
	tgccatgaag ggcaagtgtc tgctagaggt gactggggtg gcagatcca cggggtgtgg	4140
	tgaaggattc tcctatgtga agattccaaa caaaccaaca cagcagaagg atgataaaga	4200
60	accgcagcca gtgaagaaga cagtgcaggg aacagatgca gaccttctgc gcctttccct	4260
	gaaaaatgcc aagcaacttc tacgtaaatt tgggtgtgct gaggaagaga taaaaagtt	4320

65

ES 2 324 495 T3

	gtcccgcctgg	gaagtgattg	atgtggtgcg	cacaatgtca	acagaacagg	ctcgttcttg	4380
	agaggggccc	atgagtaaat	ttgcccgctg	atcaagggtt	tctgtggctg	agcatcaaga	4440
5	gcgttacaaa	gaggaatgtc	agcgcattct	tgacctacag	aacaagggtc	tgcatcaaac	4500
	tgaagtctta	tcaactgaca	cagacagcag	ctcagctgaa	gatagtgact	ttgaagaaat	4560
	gggaaagaac	attgagaaca	tgttgcagaa	caagaaaacc	agctctcagc	tttcacgtga	4620
10	acgggaggaa	caggagcgga	aggaactaca	gcgaatgcta	ctggcagcag	gctcagcagc	4680
	atccggaaac	aatcacagag	atgatgacac	agcttccgtg	actagcctta	actcttctgc	4740
	cactggacgc	tgtctcaaga	tttatcgcac	gtttcgagat	gaagagggga	aagagtatgt	4800
15	tcgctgtgag	acagtccgaa	aaccagctgt	cattgatgcc	tatgtgcgca	tacggactac	4860
	aaaagatgag	gaattcattc	gaaaatttgc	cctttttgat	gaacaacatc	gggaagagat	4920
20	gcgaaaagaa	cggcggagga	ttcaagagca	actgaggcgg	cttaagagga	accaggaaaa	4980
	ggagaagctt	aagggtcctc	ctgagaagaa	gcccagaaga	atgaaggagc	gtcctgacct	5040
	aaaactgaaa	tgtggggcat	gtggtgccat	tggaacacat	aggactaaca	aattctgccc	5100
25	cctctattat	caaacaaatg	cgccaccttc	caaccctgtt	gccatgacag	aagaacagga	5160
	ggaggagttg	gaaaagacag	tcattcataa	tgataatgaa	gaacttatca	aggttgaagg	5220
	gaccaaattt	gtcttgggga	aacagctaata	tgagagtgcg	gatgagggtc	gcagaaaatc	5280
30	tctggttctc	aagtttccta	aacagcagct	tcctccaaag	aagaaacggc	gagttggaac	5340
	cactgttcac	tgtgactatt	tgaatagacc	tcataagtcc	atccaccggc	gccgcacaga	5400
35	ccctatgggt	acgctgtcgt	ccatcttgga	gtctatcatc	aatgacatga	gagatcttcc	5460
	aaatacatat	cctttccaca	ctccagtcaa	tgcaaagggt	gtaaaggact	actacaaaat	5520
	catcactcgg	ccaatggacc	tacaaacact	ccgcgaaaac	gtgcgtaaac	gcctctaccc	5580
40	atctcgggaa	gagttcagag	agcatctgga	gctaattgtg	aaaaatagtg	caacctacaa	5640
	tgggccaaaa	cactcattga	ctcagatctc	tcaatccatg	ctggatctct	gtgatgaaaa	5700
	actcaaagag	aaagaagaca	aattagctcg	cttagagaaa	gctatcaacc	ccttgctgga	5760
45	tgatgatgac	caagtggcgt	tttctttcat	tctggacaac	attgtcacc	agaaaatgat	5820
	ggcagttcca	gattcttggc	catttcatca	cccagttaat	aagaaatttg	ttccagatta	5880
	ttacaaagtg	attgtcaatc	caatggattt	agagaccata	cgtaagaaca	tctccaagca	5940
50	caagtatcag	agtcgggaga	gctttctgga	tgatgtaaac	cttattcttg	ccaacagtgt	6000
	taagtataat	ggacctgaga	gtcagtatac	taagactgcc	caggagattg	tgaacgtctg	6060
55	ttaccagaca	ttgactgagt	atgatgaaca	tttgactcaa	cttgagaagg	atatttgtac	6120
	tgctaaagaa	gcagcttttg	aggaagcaga	attagaaagc	ctggacccaa	tgaccccagg	6180
	gccctacacg	cctcagcctc	ctgatttgta	tgataccaac	acatccctca	gtatgtctcg	6240
60	agatgcctct	gtatttcaag	atgagagcaa	tatgtctgtc	ttggatattc	ccagtgccac	6300
	tccagaaaag	caggtaacac	aggaagggtg	agatggagat	ggtgatcttg	cagatgaaga	6360

65

ES 2 324 495 T3

ggaaggaact gtacaacagc ctcaagccag tgtcctgtat gaggatttgc ttatgtctga 6420
 aggagaagat gatgaggaag atgctgggag tgatgaagaa ggagacaatc ctttctctgc 6480
 5 tatccagctg agtgaaagtg gaagtgactc tgatgtggga tctggtgga taagacccaa 6540
 acaaccccg c atgcttcagg agaacacaag gatggacatg gaaaatgaag aaagcatgat 6600
 10 gtccatagag ggagacggtg gggaggcttc ccatggtttg gaggatagca acatcagtta 6660
 tgggagctat gaggagcctg atcccaagtc gaacacccaa gacacaagct tcagcagcat 6720
 cgggtgggtat gaggtatcag aggaggaaga agatgaggag gaggaagagc agcgtctctg 6780
 15 gccgagcgta ctaagccagg tccacctgct agaggacgag gaggacagtg aggatttcca 6840
 ctccattgct ggggacagtg acttggaactc tgatgaacgg tccggaggta acggatccga 6900
 aaacctgtat tttcaggggt ctgggattgt accgcagctg caaaatattg tatccacagt 6960
 20 gaatcttggg tgtaaacttg acctaaagac cattgcactt cgtgcccga acgccgaata 7020
 taatcccaag cggtttgctg cggtaatcat gaggataaga gagccacgaa ccacggcact 7080
 gattttcagt tctgggaaaa tgggtgtcac aggagccaag agtgaagaac agtccagact 7140
 25 ggcagcaaga aaatatgcta gagttgtaca gaagtgggt tttccagcta agttcttgga 7200
 cttcaagatt cagaacatgg tggggagctg tgatgtgaag tttcctataa ggtagaagg 7260
 cttgtgtctc accaccaac aatttagtag ttatgagcca gagttatttc ctggtttaat 7320
 30 ctacagaatg atcaaacca gaattgttct cttattttt gtttctggaa agtgtgtatt 7380
 aacagggtgct aaagtcagag cagaaattta tgaagcattt gaaaacatct accctattct 7440
 35 aaagggattc aggaagacga cgcggtccgg aggtaacgga tccgaaaacc tgtattttca 7500
 ggggtgactac aaagacgatg acgataaaaa caggaagaaa ggagacaagg gctttgaaag 7560
 cccaaggcca tataaattaa cccatcagggt cgtctgcac aacaacataa atttccagag 7620
 40 aaaatctgtt gtgggatttg tggaactgac tatatttccc acagttgcaa acttgaatag 7680
 aatcaagttg aacagcaaac agttagaat ataccgagta aggatcaatg atttagaggc 7740
 tgcttttatt tataatgacc caaccttgga agtttgtcac agtgaatcaa aacagagaaa 7800
 45 cctcaattat ttttccaatg cttatgcagc tgcagttagt gctgtggacc ctgatgcagg 7860
 aaatggagaa ctttgcatta aggttccatc agagctatgg aaacacgttg atgagttaa 7920
 50 ggtcctgaag atacacatca attttcttt ggatcagccc aaaggagggtc ttcattttgt 7980
 ggtacccagt gtagaggga gtagggcaga gagagggtg catgttttct cttgtgggta 8040
 tcaaaattct acaagatttt ggttcccttg tgttgattca tactctgaat tgtgtacatg 8100
 55 gaaattagaa tttacagtag atgctgcaat ggttgctgtt tctaattggc atttggtgga 8160
 gacagtgtat actcatgata tgaggaagaa aactttccat tatatgctta ccattcctac 8220
 agcagcgtca aatatctcct tggccattgg accatttgaa atactggtag atccatacat 8280
 60 gcatgagggt actcattttt gtttgcacca acttcttcca ttgctgaaac ataccacatc 8340
 atacctcat gaagtcttg aattttatga agaaattctt acatgtcgtt acccatactc 8400

65

ES 2 324 495 T3

	ctgttttaag actgtcttca ttgatgaggc ttatgttgaa gtggctgctt atgcttccat	8460
	gagcattttt agcacaaatc ttttacacag tgccatgatt atagatgaga cacctttgac	8520
5	tagaaggtgt tttagcccaat ccttggccca gcagtttttt ggttgtttca tatctagaat	8580
	gtcttggtct gatgaatggg tgctgaaggg aatttcaggc tatatctatg gactttggat	8640
	gaaaaaaact tttggtgtta atgagtaccg ccattggatt aaagaggagc tagacaaaat	8700
10	agtggcatat gaactaaaaa ctggtggggg tttactacat cccatatttg gtggaggaaa	8760
	agagaaggat aatccggctt cccatctaca cttttcaata aagcatccac atacactgtc	8820
15	ctgggaatac tacactatgt ttcagtgtaa agcccacctt gtgatgagat tgattgaaaa	8880
	taggatcagt atggaattta tgctacaagt tttcaataaa ctgctaagtc tggctagtac	8940
	tgcttcatct cagaagtccc agtcacatat gtggagtcag atgttggttt ccacatctgg	9000
20	gtttttgaaa tccatttcaa atgtctctgg caaagatatt cagccgttaa taaagcagtg	9060
	ggtagatcag agtggagtgg taaaatttta tgggaagtttt gcatttaata gaaaacgaaa	9120
	tgtcttgga ctggaaataa aacaggacta tacatctcct ggaactcaga aatacgtggg	9180
25	accacttaaa gtgacagtgc aggagttaga tggatccttc aatcatacac tgcaaatga	9240
	agaaaacagc cttaaacatg atataccctg ccattccaaa agtagaagga ataaaaagaa	9300
	aaaaatccca ctgatgaatg gagaagaagt tgacatggat ctttctgcaa tggatgctga	9360
30	ttcccccttg ctgtggataa ggatagaccc agatatgtca gtattgagga aggtagaatt	9420
	tgagcaagct gattttatgt ggcagtatca gctccgctat gagagagatg ttgttgaca	9480
35	gcaggaatcc attttggtt tggaaaaatt ccctactcca gcattctggc ttgcactcac	9540
	tgatatatta gaacaagagc agtgtttcta cagagtaaga atgtcagctt gcttctgtct	9600
	tgcaaaagatt gcaaattcca tggtgagcac atggacagga ccaccagcca tgaagtca	9660
40	cttacttagg atgttttggt gtaaaagtgt tccaaacatt gtgaaaacaa acaactttat	9720
	gagctttcaa agttattttc tacagaagac tatgccagtt gcaatggctt tattaagaga	9780
	tgttcataat ctttgccta aagaagtctt aacgtttatt ttagacttaa tcaagtacaa	9840
45	tgacaacagg aaaaataagt tttcagataa ctattatcgt gcagaaatga ttgatgccct	9900
	ggccaactct gttacacctg cagtcagtggt gaataatgaa gttagaactt tggataactt	9960
	aatcctgat gtgcgactca ttcttgaaga aatcaccaga tttttgaata tggaaaaact	10020
50	tcttccgagt tacaggcata ccatcactgt cagttgtttg agagccatac gggactttca	10080
	gaagaacgga catgtgccaa gtgatccagc tctttttaaa tcttatgctg aatatggcca	10140
	ctttgtggac attaggatag cagctttgga agcagttggt gattatacta aagtggacag	10200
55	aagttatgaa gaactgcaat ggctacttaa tatgattcag aatgaccctg taccctatgt	10260
	aaggcataag atttcaaca tgttgactaa gaaccacca tttactaaga acatggagtc	10320
60	tcccttatgc aatgaagccc tggtagatca actttggaaa cttatgaatt ctggtacttc	10380
	acatgactgg aggttacggg gtggtgctgt ggacttgtag ttcacacttt ttggcctcag	10440

65

ES 2 324 495 T3

	tagaccttcc tgtttaccct tgccagagct tgggttggtt cttaatctaa aggagaaaaa	10500
	agctgtcttg aatcctacca taattccaga gtcagtagca ggcaaccaag aagctgcaaa	10560
5	taatccaagc agtcaccac agctagttag atttcagaac cctttttcca gttctcaaga	10620
	tgaggaggag attgatatgg atactgttca tgatagccag gccttcattt cccatcattt	10680
	aaacatgctt gaaaggccgt caactccagg gctctcgaag tatcggccag cttagctcccg	10740
10	atctgcttta ataccccagc actcagcagg ctgcgacagc acaccacca caaaaccca	10800
	gtggagtttg gaacttgac ggaagggaac aggtaaagaa caagcacctt tggagatgag	10860
	tatgcatccg gcggcaagcg ctccactctc agtctttact aaggaatcta cagcctccaa	10920
15	acacagtgc caccatcacc accatcacca tgagcacaag aaaaagaaga agaagcataa	10980
	acataagcac aaacacaagc ataagcatga cagtaaagaa aaggacaagg agcctttcac	11040
20	tttctccagc cctgccagtg gcagggtctat tctgttctct tccctttcag accgggtccgg	11100
	ccaccatcat caccaccatt gataagctag cggccgcttt cgaatctaga gcctgcagtc	11160
	tcgacaagct tgtcgagaag tactagagga tcataatcag ccataccaca tttgtagagg	11220
25	ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg	11280
	caattgttgt tgttaacttg tttattgcag cttataatgg ttacaaataa agcaatagca	11340
	tcacaaatth cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac	11400
30	tcatcaatgt atcttatcat gtctggatct gatcactgct tgagcctaga agatccggct	11460
	gctaacaaag cccgaaagga agctgagttg gctgctgcca ccgctgagca ataactagca	11520
	taaccccttg gggcctctaa acgggtcttg aggggttttt tgctgaaagg aggaactata	11580
35	tccggatctg aacaggaggg acagctgata gaaacagaag ccactggagc acctcaaaaa	11640
	caccatcata cactaaatca gtaagttagc agcatcaccc gacgcacttt gcgccgaata	11700
40	aatacctgtg acggaagatc acttcgcaga ataaataaat cctggtgtcc ctgttgatac	11760
	cgggaagccc tgggccaact tttggcgaat atgagacgtt gatcggcacg taagaggttc	11820
	caactttcac cataatgaaa taagatcact accgggctga ttttttgagt tatcgagatt	11880
45	ttcaggagct aaggaagcta aaatggagaa aaaaatcact ggatatacca ccgttgatat	11940
	atcccaatgg catcgtaaag aacattttga ggcatttcag tcagttgctc aatgtacctc	12000
	taaccagacc gttcagctgg atattacggc ctttttaaa accgtaaaga aaaataagca	12060
50	caagttttat ccggccttta ttcacattct tgcgcgctg atgaatgctc atccggaatt	12120
	ccgtatggca atgaaagacg gtgagctggt gatatgggat agtggttcacc cttgttacac	12180
	cgtttttcat gagcaaaact aaacgttttc atcgctctgg agtgaatacc acgacgattt	12240
55	ccggcagttt ctacacatat attcgcaaga tgtggcgtgt tacgggtgaaa acctggccta	12300
	tttccctaiaa gggttttatt agaatatgtt tttcgtctca gccaatccct gggtagattt	12360
60	caccagtttt gatttaaacg tggccaatat ggacaacttc ttcgcccccg ttttcaccat	12420
	gggcaaatat tatacgcaag gcgacaaggt gctgatgccg ctggcgattc aggttcatca	12480

65

ES 2 324 495 T3

5 **tgccgtctgt gatggcttcc atgtcggcag aatgcttaat gaattacaac agtactgcga 12540**
 tgagtggcag ggcggggcgt aattttttta aggcagttat tggtgccctt aaacgcctgg 12600
 10 **tgctacgcct gaataagtga taataagcgg atgaatggca gaaattcgaa agcaaattcg 12660**
 acccggtcgt cggttcaggg cagggtcgtt aaatagccgc ttatgtctat tgctggttta 12720
 ccggtttatt gactaccgga agcagtgtga ccgtgtgctt ctcaaatgcc tgaggccagt 12780
 15 **ttgctcaggc tctccccgtg gaggaataa ttgacgatat gatcatttat tctgcctccc 12840**
 agagcctgac attcatccgg ggtcagcacc gtttctgcgg actggctttc tacgtgttcc 12900
 gcttccttta gcagcccttg cggcctgagt gcttgcgga gcgtgaagct 12950

<210> 5

<211> 8735

20 <212> ADN

<213> Secuencia artificial

<220>

25 <223> Plásmido pFBOO[HiSTEVTAF6TAF9]his

<400> 5

30 **ttctctgtca cagaatgaaa atttttctgt catctcttcg ttattaatgt ttgtaattga 60**
 ctgaatatca acgcttattt gcagcctgaa tggcgaatgg gacgcgccct gtagcggcgc 120
 attaagcgcg gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg ccagcgcctt 180
 35 **agcgcgccgt cttttcgtt tcttcccttc ctttctcgcc acgttcgccg gctttccccg 240**
 tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac ggcacctcga 300
 ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccacgcgccct gatagacggt 360
 40 **ttttcgccct ttgacgttgg agtccacgtt ctttaatatg ggactcttgt tccaaactgg 420**
 aacaacactc aaccttatct cggcttattc ttttgattta taagggattt tgccgatttc 480
 ggcctatttg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat 540
 45 **attaacgttt acaatttcag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg 600**
 tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat 660
 gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat 720
 50 **tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt 780**
 aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag 840
 55 **cggttaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa 900**
 agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcgggtcg 960
 ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag aaaagcatct 1020
 60 **tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga gtgataacac 1080**

65

ES 2 324 495 T3

	tgccggccaac ttactttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca	1140
	caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga atgaagccat	1200
5	accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaaact	1260
	attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc	1320
	ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga	1380
10	taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg	1440
	taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg	1500
	aaatagacag atcgttgaga taggtgcctc actgattaag cattggtaac tgtcagacca	1560
15	agtttactca tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta	1620
	ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca	1680
	cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga	1740
20	tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataaccaaa	1800
	tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc	1860
	tacatacctc gctctgctaa tcctgttacc agtggtgctg gccagtggcg ataagtcgtg	1920
25	tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg cgggctgaac	1980
	gggggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct	2040
30	acagcgtgag cattgagaaa gcgccacgct tcccgaaggg agaaaaggcg acaggtatcc	2100
	ggtaagcggc agggctcgga caggagagcg cacgaggag cttccagggg gaaacgcctg	2160
	gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg	2220
35	ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct	2280
	ggccttttgc tggccttttg ctacatggt ctttcctgcg ttatcccctg attctgtgga	2340
	taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg	2400
40	cagcgagtca gtgagcgagg aagcggaaga gcgcctgatg cgggtattttc tccttacgca	2460
	tctgtcggtt atttcacacc gcagaccagc cgcgtaacct ggcaaaatcg gttacggttg	2520
	agtaataaat ggatgccctg cgtaagcggg tgtgggcgga caataaagtc ttaaactgaa	2580
45	caaaatagat ctaaactatg acaataaagt cttaaactag acagaatagt tgtaaactga	2640
	aatcagtcca gttatgctgt gaaaaagcat actggacttt tgttatggct aaagcaaact	2700
50	cttcattttc tgaagtgcaa attgcccgtc gtattaaaga ggggcgtggc caagggcatg	2760
	gtaaagacta tattcgcggc gttgtgacaa ttaccgaac aactccgcgg ccgggaagcc	2820
	gatctcggct tgaacgaatt gttaggtggc ggtacttggg tcgatatcaa agtgcacac	2880
55	ttcttcccgt atgcccaact ttgtatagag agccactgcg ggatcgtcac cgtaatctgc	2940
	ttgcacgtag atcacataag caccaagcgc gttggcctca tgcttgagga gattgatgag	3000
	cgcggtggca atgccctgcc tccggtgctc gccggagact gcgagatcat agatatagat	3060
60	ctcactacgc ggctgctcaa acctgggcag aacgtaagcc gcgagagcgc caacaaccgc	3120

ES 2 324 495 T3

	ttcttggtcg aaggcagcaa gcgcgatgaa tgtcttacta cggagcaagt tcccagggtg	3180
	atcggagtcc ggctgatgtt gggagtaggt ggctacgtct ccgaactcac gaccgaaaaag	3240
5	atcaagagca gcccgcatgg atttgacttg gtcagggccg agcctacatg tgcgaatgat	3300
	gcccatactt gagccaccta actttgtttt agggcgactg ccctgctgcg taacatcggt	3360
	cttgctgctt ggatgcccga ggcatagact gtacaaaaaa acagtcataa caagccatga	3420
10	aaaccgccac tgcgccgtta ccaccgctgc gttcgggtcaa ggttctggac cagttgcgtg	3480
	agcgcatacg ctacttgcac tacagtttac gaaccgaaca ggcttatgtc aactgggttc	3540
15	gtgccttcac ccgtttccac ggtgtgcgtc acccggcaac cttgggcagc agcgaagtcg	3600
	aggcatttct gtcctggctg gcgaacgagc gcaagggttc ggtctccacg catcgtcagg	3660
	cattggcggc cttgctgttc ttctacggca aggtgctgtg cacggatctg ccctggcttc	3720
20	aggagatcgg tagacctcgg ccgtcgcggc gcttgccggg ggtgctgacc ccgatgaag	3780
	tggttcgcac cctcggtttt ctggaaggcg agcatcgttt gttcgcccag gactctagct	3840
	atagttctag tggttggcct acagctttgt ttaaacaaga ctgtaccctg agtggtcatg	3900
25	gcagggcttg ccgccccgac gttggctgcg agccctgggc cttcacccga acttgggggt	3960
	tggggggggg aaaaggaaga aacgcgggcg tattgggtccc aatgggggtct cgggtggggtg	4020
	tcgacagagt gccagccctg ggaccgaacc ccgcgtttat gaacaaacga cccaacaccc	4080
30	gtgcgtttta ttctgtcttt ttattgccgt catagcgcgg gttccttccg gtattgtctc	4140
	cttccgtgtt tcagttagcc tccccatct cccggtagcg catgctatgc atcagctgct	4200
	agcaccatgg ctcgagatcc cgggtgatca agtcttcgtc gagtgattgt aaataaaatg	4260
35	taatttacag tatagtattt taattaatat acaaatgatt tgataataat tcttatttaa	4320
	ctataatata ttgtgttggg ttgaattaaa ggtccgtata ctagtatcga ttcgcgacct	4380
	actccggaat attaatagat catggagata attaaaatga taaccatctc gcaaataaat	4440
40	aagtatttta ctgttttcgt aacagttttg taataaaaaa acctataaat attccggatt	4500
	attcataccg tcccaccatc gggcgcggat cctcgagatg ggtaaccatc atcatcatca	4560
45	tcacggagaa agcttgttta agggaccacg tgattacaac ccgatatcga gcaccatttg	4620
	tcatttgacg aatgaatctg atgggcacac aacatcgttg tatggatttg gatttgggtc	4680
	cttcatcatt acaaacaagc acttgtttag aagaaataat ggaacactgt tgggtccaatc	4740
50	actacatggt gtattcaagg tcaagaacac cagactttg caacaacacc tcattgatgg	4800
	gagggacatg ataattattc gcatgcctaa ggatttccca ccatttcctc aaaagctgaa	4860
	atttagagag ccacaaaggg aagagcgcat atgtcttgtg acaaccaact tccaaactaa	4920
55	gagcatgtct agcatggtgt cagacactag ttgcacattc cttcatctg atggcatatt	4980
	ctggaagcat tggattcaaa ccaaggatgg gcagtgtggc agtccattag tatcaactag	5040
	agatgggttc attgttggtg tacactcagc atcgaatttc accaacacaa acaattattt	5100
60	cacaagcgtg ccgaaaaact tcatggaatt gttgacaaat caggaggcgc agcagtgggt	5160

ES 2 324 495 T3

	tagtggttgg cgattaaatg ctgactcagt attgtggggg ggccataaag ttttcatgag	5220
	caaacctgaa gagccttttc agccagttaa ggaagcgact caactcatga atgaattggt	5280
5	gtactcgcaa ggtggtggtg aaaacctgta cttccagggg aaccacgctg aggagaagaa	5340
	gctgaagctt agcaacactg tgctgccctc ggagtccatg aagggtggtg ctgaatccat	5400
	gggcatcgcc cagattcagg aggagacctg ccagctgcta acggatgagg tcagctaccg	5460
10	catcaaagag atcgcacagg atgccttgaa gttcatgcac atggggaagc ggcagaagct	5520
	caccaccagt gacattgact acgccttgaa gctaaagaat gtcgagccac tctatggctt	5580
	ccacgcccag gagttcattc ctttccgctt cgcctctggt gggggccggg agctttactt	5640
15	ctatgaggag aaggagggtg atctgagcga catcatcaat acccctctgc cccgggtgcc	5700
	cctggacgtc tgcttcaaag ctcatgggt gagcatcgag ggctgccagc cagctatccc	5760
	cgagaacccg cccccagctc ccaaagagca acagaaggct gaagccacag aaccctgaa	5820
20	gtcagccaag ccaggccagg aggaagacgg accctgaag ggcaaaggct aaggggccac	5880
	cacagccgac ggcaaaggga aagagaagaa ggcgcggccc ttgctggagg gggccccctt	5940
	gcgactgaag ccccgagca tccacgagtt gtctgtggag cagcagctct actacaagga	6000
	gatcacgag gcctgcgtgg gtcctgcga ggccaagagg gcggaagccc tgcaaagcat	6060
	tgccacggac cctggactgt atcagatgct gccacggttc agtaccttta tctcggaggg	6120
30	ggtccgtgtg aacgtgggtc agaacaacct ggccctactc atctacctga tgcgtatggt	6180
	gaaagcgctg atggacaacc ccacgtctta tctagaaaaa tacgtccatg agctgattcc	6240
	agctgtgatg acctgcatcg tgagcagaca gttgtgcctg cgaccagatg tggacaatca	6300
35	ctgggcactc cgagactttg ctgcccgcct ggtggcccag atctgcaagc attttagcac	6360
	aaccactaac aacatccagt cccggatcac caagaccttc accaagagct ggggtggacga	6420
	gaagacgccc tggacgactc gttatggctc catcgcaggc ttggctgagc tgggacacga	6480
40	tgttatcaag actctgattc tgccccggct gcagcaggaa ggggagcgga tccgcagtgt	6540
	gctggacggc cctgtgtgta gcaacattga ccggattgga gcagaccatg tgcagagcct	6600
	cctgtgtaaa cactgtgctc ctgttctggc aaagctgcgc ccaccgcctg acaatcagga	6660
45	cgcctatcgg gcagaattcg ggtcccttgg gcccctctc tgctcccagg tggtaaggc	6720
	tcggggccag gctgctctgc aggtcagca ggtcaacagg accactctga ccatcacgca	6780
	gccccggccc acgctgacct tctcgcaggc cccacagcct ggccctcgca cccctggctt	6840
50	gctgaagggt cctggctcca tcgcatttc tgtccagaca ctggtgtctg cagagcggc	6900
	tgccccacca cagccttccc ctctccaac caagtttatt gtaatgtcat cgtctccag	6960
55	cgccccatcc acccagcagg tcctgtccct cagcacctcg gccccggct caggttccac	7020
	caccacttcg cccgtcacca ccaccgtccc cagcgtgcag cccatcgtca agttggtctc	7080
	caccgccacc accgcacccc ccagcaetgc tccctctggg cctgggagtg tccagaagta	7140
60	catcgtggtc tcacttcccc caacagggga gggcaaagga ggccccacct cccatccttc	7200

ES 2 324 495 T3

	tccagttcct cccccggcat cgtccccgtc cccactcagc ggcagtgtccc tttgtggggg	7260
	gaagcaggag gctggggaca gtccccctcc agctccaggg actccaaaag ccaatggctc	7320
5	ccagcccaac tccggctccc ctccagctgc tccgcggtcc ggtgggtggtg gtgaaaacct	7380
	gtattttcag ggcgagtctg gcaagacggc ttctcccaag agcatgccga aagatgcaca	7440
	gatgatggca caaatcctga aggatatggg gattacagaa tatgagccaa gagttataaa	7500
10	tcagatgttg gagtttgcct tccgatatgt gaccacaatt ctagatgatg caaaaattta	7560
	ttcaagccat gctaagaaag ctactgttga tgcagatgat gtgcgattgg caatccagtg	7620
15	ccgcgctgat cagtctttta cctctcctcc cccaagagat tttttattag atattgcaag	7680
	gcaaagaaat caaacccctt tgccattgat caagccatat tcagggtccaa ggttgccacc	7740
	tgatagatac tgcttaacag ctccaaacta taggctgaaa tctttacaga aaaaggcatc	7800
20	aacttctgcg ggaagaataa cagtcccgcg gttaagtgtt gggtcagtta ctagcagacc	7860
	aagtactccc aactaggca caccaacccc acagaccatg tctgtttcaa ctaaagtagg	7920
	gactcccatg tccctcacag gtcaaagggt tacagtacag atgcctactt ctcagtctcc	7980
25	agctgtaaaa gcttcaattc ctgcaacctc agcagttcag aatgttctga ttaatccatc	8040
	attaatcggg tccaaaaaca ttcttattac cactaatatg atgtcatcac aaaatactgc	8100
	caatgaatca tcaaatgcat tgaaaagaaa acgtgaagat gatgatgatg acgatgatga	8160
30	tgatgatgac tatgataatc tgcggtccgg ccaccatcat caccaccatt gataagctag	8220
	cggccgcttt cgaatctaga gcctgcagtc tcgacaagct tgtcgagaag tactagagga	8280
35	tcataatcag ccataccaca tttgtagagg ttttacttgc tttaaaaaac ctcccacacc	8340
	tccccctgaa cctgaaacat aaaatgaatg caattgttgt tgtaaacttg tttattgcag	8400
	cttataatgg ttacaaataa agcaatagca tcacaaattt cacaaataaa gcattttttt	8460
40	cactgcattc tagttgtggg ttgtccaaac tcatcaatgt atcttatcat gtctggatct	8520
	gatcactgct tgagcctagg agatccgaac cagataagtg aaatctagtt ccaaactatt	8580
	ttgtcatttt taattttcgt attagcttac gacgctacac ccagttccca tctattttgt	8640
45	cactcttccc taaataatcc ttaaaaactc catttccacc cctcccagtt cccaacgcca	8700
	actecatgtg acaaaccgtc atcttcgggt actttt	8735

50

55

60

65