



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 295 433**

(51) Int. Cl.:

C12N 15/63 (2006.01)

C12N 15/67 (2006.01)

C12N 15/70 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **02788244 .8**

(86) Fecha de presentación : **23.12.2002**

(87) Número de publicación de la solicitud: **1461437**

(87) Fecha de publicación de la solicitud: **29.09.2004**

(54) Título: **Vectores de expresión y promotores para la expresión de genes heterólogos.**

(30) Prioridad: **21.12.2001 PT 10270401**

(45) Fecha de publicación de la mención BOPI:
16.04.2008

(45) Fecha de la publicación del folleto de la patente:
16.04.2008

(73) Titular/es: **BiotecnoI S.A.**
Lagoas Park Edificio 7 - 1 Piso Norte
2741-901 Porto Salvo, PT

(72) Inventor/es: **Kelly, Andrew, Graham;**
Filipe, Susana, Melquiades;
Gnadt, Nicole, Claudia;
Barbas, Ana, Lucia y
Pissarra, Pedro de Noronha

(74) Agente: **Elzaburu Márquez, Alberto**

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

Vectores de expresión y promotores para la expresión de genes heterólogos.

5 La presente invención se refiere a nuevos vectores y secuencias de ácido nucleico, que comprenden promotores que se inducen por nucleótidos (con o sin un grupo fosfato). La invención se refiere además a vectores de expresión que utilizan elementos del operón de la xantosina y a métodos para la producción de proteínas heterólogas.

10 La expresión de genes clonados introducidos en bacterias es aún el mecanismo más ampliamente utilizado para producir grandes cantidades de una proteína de interés para fines diagnósticos y terapéuticos. Con el fin de producir eficazmente proteínas en un anfitrión procariota, un promotor fuerte, regulado es un elemento esencial del sistema de expresión. Los promotores y los sistemas de expresión que se conocen en la técnica incluyen los promotores pL y pR del bacteriófago lambda, que pueden regularse por un represor sensible a la temperatura que reprime la transcripción a partir de ese promotor a temperaturas bajas y que permite la expresión de la proteína heteróloga mediante una inducción por la temperatura (EP-A-41767) o alternativamente mediante la utilización de un siste-
15 ma antirrepresor inducible que puede inducirse utilizando isopropil- β -D-tiogalactopiranosido (IPTG) a temperaturas más bajas, aún utilizando los promotores fuertes pL/pR (WO 98/48025). Sin embargo, las temperaturas altas dan lugar rutinariamente a la formación de la proteína heteróloga como agregados insolubles, o cuerpos de inclusión, lo que requiere posteriormente etapas de repliegamiento costosas para las que se utilizan compuestos químicos tóxi-
20 cos.

Otros promotores incluyen el operón lac y derivados, tales como el promotor trp-lac o el promotor tac (EP-A-67540) que se ha utilizado para producir niveles altos de proteínas en *E. coli*. Este promotor se induce en presencia de IPTG. Con el fin de someter al promotor a la represión, debe utilizarse junto con un plásmido que sobreproduzca el represor lac o en una cepa de *E. coli* que produzca la proteína represora lac. Los promotores basados en el promotor T7 (EEUU5869320) que utilizan la polimerasa de ARN T7, expresada de manera inducible, se utilizan rutinariamente en los laboratorios de investigación para sobreproducir proteínas, de nuevo utilizando IPTG. Sin embargo, la fuerza de dichos promotores puede dar lugar de nuevo a la producción de cuerpos de inclusión, o agregados, de las proteínas heterólogas.

30 *Escherichia coli*, una bacteria Gram-negativa, es capaz de crecer en xantosina como única fuente de carbono (Hammer-Jespersen *et al.*, 1980). Esto ocurre a través de la inducción y utilización de genes del operón de la xantosina, que comprende al menos tres genes, localizados a aproximadamente 52' en el cromosoma de *E. coli* (Figura 1) (Seeger *et al.*, 1995). Los genes del operón son xapA, xapB y xapR. XapA codifica la xantosina fosforilasa, que cataliza la degradación del enlace N-glicosídico de las moléculas de nucleósidos lo que resulta en una base libre y una pentosa-1-fosfato. Las bases libres pueden utilizarse posteriormente en las vías de recuperación de purina y pirimidina y también como una fuente de nitrógeno (Nygaard, 1983) y la molécula de pentosa puede utilizarse posteriormente como una fuente de carbono, lo que explica la capacidad de *E. coli* de crecer en medio mínimo con xantosina. La función de la proteína xapB no se ha resuelto totalmente pero parece ser, por su homología con otras proteínas y por su localiza-
40 ción celular, una proteína de transporte de membrana, implicada casi seguro en el transporte de nucleósidos a través de la membrana celular (Seeger, *supra*). XapR muestra una homología significativa con la familia LysR de proteínas reguladoras de la transcripción, proteínas de unión al ADN implicadas en la activación génica (Henikoff *et al.*, 1988) y parece que se expresa constitutivamente. Los genes xapA y xapB parece que se co-expresan y la expresión de xapA se encuentra sólo en presencia de xantosina y la proteína xapR (Seeger, *supra*). La proteína xapR, si es realmente un miembro de la familia LysR, se uniría a regiones específicas del promotor xapA, permitiendo la unión posterior de la Polimerasa de ARN de *E. coli* y la transcripción posterior de los genes xapA (y xapB). Dichas proteínas de unión al ADN parecen reconocer una secuencia consenso complementaria invertida ATATTGTTT (Bohannon 1989) y parece que dicha región existe en la región del promotor xapA (Figura 2) 120pb antes del extremo 5' del inicio de la transcripción (CCAATACAGTTTT) y la secuencia correspondiente a 40pb antes del extremo 5', superponiéndose a la región -35 (AAACTGTATTGG) (Seeger, *supra*). Las secuencias de los genes que comprenden el operón de la xantosina de *E. coli* se han depositado en numerosas bases de datos públicas en línea, por ejemplo, en la base de datos Genbank, accesible a través de la página principal del Nacional Center for Biotechnology Information (NCBI), Bethesda, Maryland, EEUU en www.ncbi.nlm.nih.gov, utilizando los números de registro AE000328, D90869, D90870 o X73828.

55 La presente invención se refiere a la expresión de proteínas heterólogas utilizando un vector de expresión y sistema de expresión nuevo, que evita los problemas de los cuerpos de inclusión insolubles asociado con la inducción por la temperatura o con promotores muy fuertes y que no requiere la utilización de inductores tóxicos, tales como IPTG, que requerirían procedimientos de validación costosos para verificar su ausencia cuando se producen proteínas terapéuticas.
60

De acuerdo con esto, un primer aspecto de la presente invención proporciona un vector, que comprende un promotor que puede unirse operativamente a un gen que codifica una proteína heteróloga, en el que el promotor induce la expresión de cualquier proteína heteróloga unida operativamente, en presencia de nucleótidos, pudiendo tener o no dichos nucleótidos un grupo fosfato y en el que la secuencia promotora es un promotor XapA que compren-
65 de un sitio de unión ribosomal que tiene lo siguiente AGGAGGxxxxx o AGGAGGxxxxxx o AGGAGAxxxxxx o AGGAGAxxxxxx.

Puede que el grupo fosfato de la molécula sea un éster fosfato. Puede haber uno o más grupos fosfato unidos.

El vector puede ser cualquiera, incluyendo un plásmido o un bacteriófago. Preferiblemente, el vector es un vector de expresión inducible.

La inducción del promotor se realiza mediante la presencia de nucleótidos en el medio (habitualmente medio de cultivo).

El término “nucleótidos sin o con un grupo fosfato” incluye tanto nucleótidos como nucleósidos.

Según la presente invención, los nucleótidos incluyen una o más de cualquiera de las moléculas de nucleósido que contiene uno o más grupos fosfato unidos covalentemente a una molécula de azúcar, que puede ser de la forma oxi, desoxi o didesoxi. Los ejemplos de éstas incluyen: adenosina trifosfato (ATP), guanosina trifosfato (GTP), citidina trifosfato (CTP), timidina trifosfato (TTP) y uridina trifosfato (UTP). Las formas didesoxi de éstas (escritas habitualmente como dATP, dCTP, dGTP y dTTP y conocidas habitualmente como bases) forman la estructura básica del ADN. Las formas oxi de éstas (escritas habitualmente como ATP o rATP, CTP o rCTP, GTP o rGTP, UTP o rUTP y conocidas habitualmente como bases de ARN) forman la estructura básica del ARN.

Según la presente invención, los nucleósidos incluyen uno o más de cualquier compuesto que comprende una purina o pirimidina unida mediante un enlace N-glicosídico a un azúcar, particularmente una molécula de azúcar de oxi o desoxi o didesoxi-ribosa. Los ejemplos de dichos nucleósidos incluyen las formas oxi, desoxi y didesoxi de adenosina, guanosina, inosina, citidina, timidina, uridina, xantosina y derivados de éstas.

La xantosina se conoce en la clasificación USPTO como: 536/27.8. También se conoce como xantosina ó 9-beta-D-ribofuranosil xantina (número del Chemical Abstract Service 5968-90-1) e incluye el compuesto bien como una sal anhidra o como una sal dihidrato. Los nombres sinónimos para la xantosina son: Xantina ribósido, 9-beta-D-ribofuranosil-9H-purina-2,6-diol y 9-beta-D-ribofuranosil-9H-purina-2,6-(1H,3H)-diona.

Las concentraciones preferidas de nucleótidos como inductor están en el intervalo de 0,01 a 10 mg/ml, más preferiblemente de 0,1 a 1 mg/ml.

Puede utilizarse cualquier promotor que pueda unirse operativamente al gen que codifica una proteína heteróloga y en el que el promotor induce la expresión de cualquiera de dichas proteínas heterólogas unidas operativamente, en presencia de nucleótidos.

Las etapas para unir operativamente secuencias de ácido nucleico son muy conocidas en la técnica y actualmente se basan mayoritariamente en la utilización de enzimas de restricción para cortar el ácido nucleico y de polimerasas para unir el ácido nucleico. Los ejemplos de dichas etapas se muestran en la sección de los ejemplos.

El promotor puede inducirse directamente o indirectamente mediante la presencia de nucleótidos. Los promotores adecuados incluyen los de un gen xapA. Es aceptable cualquier gen xapA. La región promotora xapA de *E. coli* se muestra en la Figura 1a. Las regiones modificadas del promotor xapA, incluidas en la presente invención, se muestran en las Figuras 1b y 1c. Estas regiones modificadas incluyen sitios útiles de enzimas de restricción.

Además de esta secuencia, puede utilizarse cualquier secuencia modificada si dicha secuencia induce la expresión de una secuencia de ácido nucleico unida operativamente en presencia de nucleótidos o nucleósidos. Por ejemplo, la secuencia modificada puede ser una secuencia sustancialmente homóloga. Una secuencia sustancialmente homóloga tiene preferiblemente al menos 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 ó 100% de identidad de secuencia con la secuencia definida.

“% de identidad” es una medida de la relación entre dos secuencias de ácido nucleico o polipéptido, determinada comparando sus secuencias. En general, las dos secuencias que se quieren comparar se alinean para proporcionar una correlación máxima entre las secuencias. El alineamiento de las dos secuencias se examina y se determina el número de posiciones que proporciona una correspondencia exacta de aminoácidos o nucleótidos y se divide por la longitud total del alineamiento y el resultado se multiplica por 100 para proporcionar un % de identidad. El % de identidad puede determinarse sobre la longitud completa de la secuencia que se quiere comparar, lo que es particularmente adecuado para secuencias con la misma o similar longitud o para secuencias que son altamente homólogas o sobre longitudes más cortas definidas lo que es más adecuado para secuencias de longitudes diferentes y con una menor homología.

Los métodos para comparar la identidad de dos o más secuencias son conocidos en la técnica. Por ejemplo, pueden utilizarse los programas disponibles en el Wisconsin Sequence Analysis Package versión 9.1 (Devereux J *et al.*, *Nucl Acid Res* 12 387-395 (1984), disponible en Genetics Computer Group, Madison, Wisconsin, EEUU) tales como BESTFIT y GAP.

BESTFIT utiliza el algoritmo de “homología local” de Smith y Waterman (Advances in Applied Mathematics, 2:482-489, 1981) y encuentra la mejor región única de similitud entre dos secuencias. BESTFIT es más adecuado para comparar dos secuencias de polinucleótidos o dos secuencias de polipéptidos que tienen una longitud diferente,

asumiendo el programa que la secuencia más corta representa una parte de la más larga. En comparación, GAP alinea dos secuencias encontrando una “similitud máxima” según el algoritmo de Neddleman y Wunsch (J. Mol. Biol. **48**:443-354, 1970).

GAP es más adecuado para comparar secuencias que tienen aproximadamente la misma longitud y se espera un alineamiento sobre la longitud total. Preferiblemente, los parámetros “Peso del Hueco” y “Peso de la Longitud” utilizados en cada programa son 50 y 3 para secuencias de polinucleótidos y 12 y 4 para secuencias de polipéptidos, respectivamente. Preferiblemente, los % de identidades y similitudes se determinan cuando las dos secuencias que se quieren comparar están alineadas de manera óptima.

También se conocen en la técnica otros programas para determinar la identidad y/o similitud entre dos secuencias, por ejemplo la familia BLAST de programas (Altschul *et al.*, *J. Mol. Biol.*, **215**:403-410, (1990) y Altschul *et al.*, *Nuc Acids Res.*, **25**:289-3402 (1997), disponibles en el National Center for Biotechnology Information (NCBI, Bethesda, Maryland, EEUU y accesible a través de la página principal de NCBI en www.ncbi.nlm.nih.gov) y FASTA (Pearson W.R. y Lipman D.J., *Proc. Nat. Acad. Sci.*, USA, **85**:2444-2448 (1988), disponible como parte del Wisconsin Sequence Analysis Package).

En relación con la presente invención, “condiciones astringentes” se refiere a las condiciones de lavado utilizadas en un protocolo de hibridación. En general, las condiciones de lavado deben ser una combinación de temperatura y concentración de sal de manera que la temperatura de desnaturalización es aproximadamente 5 a 20°C por debajo de la T_m calculada del ácido nucleico que se está estudiando. La T_m de una sonda de ácido nucleico de 20 bases o menos se calcula en condiciones estándar (1M NaCl) como $[4^\circ\text{C} \times (\text{G}+\text{C}) + 2^\circ\text{C} \times (\text{A}+\text{T})]$, según las reglas de Wallace para oligonucleótidos cortos. Para fragmentos de ADN más largos, puede utilizarse el método del vecino más próximo, que combina termodinámica sólida y datos experimentales, según los principios mostrados en Breslauer *et al.*, *PNAS* **83**:3746-3750 (1986). Las condiciones óptimas de sal y temperatura para la hibridación pueden determinarse fácilmente en experimentos preliminares en los que se inmovilizan muestras de ADN en filtros y se hibridan con la sonda de interés y se lavan en condiciones de diferentes astringencias. Aunque las condiciones para PCR pueden ser diferentes de las condiciones estándar, la T_m puede utilizarse como guía para la estabilidad relativa esperada de los cebadores. Para cebadores cortos de aproximadamente 14 nucleótidos, se utilizan temperaturas de hibridación bajas de aproximadamente 44°C a 50°C. La temperatura puede ser mayor dependiendo de la composición de bases de la secuencia de cebador utilizada. Las condiciones astringentes adecuadas son aquellas en las que se evita la hibridación no específica. Las condiciones astringentes adecuadas son 0,5xSSC/1%SDS/58°C/30 min para una sonda de oligonucleótido de 21mer.

El vector puede comprender además un elemento regulador con el que interaccionan los nucleótidos con el fin de regular el promotor. El elemento regulador es preferiblemente una secuencia de ácido nucleico. Cuando el elemento regulador es un gen *xapR*, puede ser necesario que se exprese con el fin de que la proteína expresada interaccione con los nucleótidos y regule el promotor, para una expresión incrementada de cualquier proteína heteróloga.

Lo más preferiblemente, según la invención, el elemento regulador es de un gen *xapR*. Es aceptable cualquier gen *xapR*. El gen *xapR* de *E. coli* se muestra en la Figura 2.

Además de esta secuencia *xapR* dada, puede utilizarse cualquier secuencia *xapR* modificada que actúe de la misma manera sobre el promotor. Por ejemplo, la secuencia modificada puede ser una secuencia *xapR* sustancialmente homóloga (como se ha definido anteriormente). Puede necesitarse que la secuencia *xapR* se exprese. Las secuencias *xapR* modificadas adecuadas están descritas en Jorgensen y Dandanell, 1999.

Se piensa que la proteína *xapR* expresada se une a regiones específicas del promotor *xapA*, permitiendo la unión posterior de la Polimerasa de ARN y la transcripción posterior de cualquier gen heterólogo unido operativamente.

Con el fin de que *xapR* u otros elementos reguladores se expresen puede ser necesario que el vector comprenda un promotor más que está unido operativamente al elemento regulador para controlar su expresión. Preferiblemente, dicho promotor es un promotor constitutivo o inducible, de manera que el elemento regulador se expresa constitutivamente. El promotor adicional puede ser un promotor *xapR* o cualquier otro promotor constitutivo o inducible conocido, tal como los promotores *lac*, *trc*, *lambda pR*, *lambda pL* o *trp*.

En una alternativa, el elemento regulador que regula la expresión por el promotor, puede estar presente pero no como parte del vector en el que está presente el promotor inducible. El elemento regulador (preferiblemente con promotor constitutivo o inducible) puede estar presente en una célula como parte de otro vector o integrado en el genoma de la célula.

El vector del primer aspecto de la invención puede tener o no un gen unido operativamente que codifica una proteína heteróloga. La forma del vector sin dicho gen puede denominarse un “casete vacío” que permite la adición de cualquiera de dichos genes para utilizarse en la expresión de ese gen. La adición es por etapas conocidas en la técnica (predominantemente la utilización de enzimas de restricción de corte y polimerasas de unión), ejemplos de las cuales se proporcionan en los ejemplos.

Cualquier gen que codifica cualquier proteína heteróloga puede insertarse en el vector del primer aspecto de la invención. Los ejemplos de dichos genes incluyen los que codifican citoquinas, hormonas, quimioquinas, enzimas,

antígenos, etc (preferiblemente formas humanas). Las citoquinas incluyen interleuquinas, por ejemplo, interleuquina-4 humana. Las hormonas incluyen la hormona de crecimiento humana. La proteína heteróloga producida puede ser una proteína de fusión, por ejemplo, unida a un péptido líder, para la secreción en el espacio periplásmico o en el medio extracelular, utilizando péptidos líder como *ompA* o *pelB* o un polipéptido secundario, unido operativamente a la proteína heteróloga, cuya función puede ser evitar la degradación proteolítica de la proteína heteróloga o proporcionar una etiqueta de afinidad para la purificación o ayudar en la solubilización de la proteína heteróloga. Los ejemplos de dichos polipéptidos secundarios incluyen tioredoxina, proteína de unión a maltosa, etiquetas de histidina y dichos polipéptidos secundarios pueden incluir o no sitios de escisión para la eliminación del polipéptido secundario mediante escisión selectiva utilizando medios químicos o enzimáticos. Los ejemplos de éstos incluyen bromuro de cianógenos, tripsina, enteroquinasa y Factor Xa.

Un segundo aspecto de la invención proporciona un vector de expresión, que comprende un ácido nucleico aislado, que comprende un gen regulador de *xapR* del operón de la xantosina junto con un promotor de un gen *xapA*.

El promotor *xapA* puede ser como se describe en la presente memoria en la Figura 1a e incluye cualquier secuencia sustancialmente homóloga, también como se ha descrito en la presente memoria anteriormente.

El gen *xapR* puede ser como se describe en la presente memoria en la Figura 2 e incluye cualquier secuencia sustancialmente homóloga, también como se ha descrito en la presente memoria anteriormente.

El ácido nucleico puede comprender también un promotor más que está unido operativamente al gen *xapR*. Dicho promotor es como se ha descrito en el primer aspecto de la invención. Puede ser del promotor *xapR* o puede ser cualquier promotor inducible o constitutivo. El promotor *xapR* puede ser como se describe en la presente memoria en la Figura 2.

El vector o vector de expresión del primer y segundo aspecto puede comprender, por supuesto, otros elementos, tales como marcador de selección fenotípico, tal como resistencia a antibióticos, gen o genes de replicación, etc.

Un tercer aspecto de la invención proporciona una secuencia promotora *xapA* que comprende un sitio de unión ribosomal que tiene una cualquiera de las secuencias siguientes:

AGGAGG xxxxx

AGGAGG xxxxxx

AGGAGA xxxxx

AGGAGA xxxxxx

en las que x es cualquier base.

Las bases opcionales (x) pueden ser las descritas en el Ejemplo 2 (taccx o tatccx).

Las secuencias de ácido nucleico del tercer aspecto pueden comprender las secuencias 5', también como se muestra en el Ejemplo 2. Las secuencias de ácido nucleico del tercer aspecto pueden ser parte de una secuencia promotora. La secuencia promotora puede ser de un promotor *xapA*. Las secuencias de ácido nucleico del tercer aspecto de la invención pueden ser parte de un vector, como se ha descrito según el primer aspecto de la invención.

Un cuarto aspecto de la invención proporciona una célula anfitriona, que comprende uno o más vectores o secuencias de ácido nucleico, según una cualquiera del primer, segundo o tercer aspecto de la invención. Dicha célula anfitriona puede ser cualquiera, preferiblemente aquellas que pueden utilizarse en cultivo para proporcionar una producción de proteínas a escala comercial. Los ejemplos de dichas células anfitrionas incluyen *E. coli*, *Bacillus sp.* y levaduras tales como *Saccharomyces* y *Pichia*.

El vector y/o ácido nucleico se introduce en la célula anfitriona mediante cualquier técnica, tal como transformación, electroporación etc.

Un quinto aspecto de la invención proporciona un método para la expresión de una proteína heteróloga, comprendiendo el método cultivar las células anfitrionas, según el cuarto aspecto de la invención en condiciones que inducen la expresión de la proteína heteróloga. Dichos métodos son muy conocidos en la técnica (véase, por ejemplo, "Manual of Industrial Microbiology and Biotechnology" 2a Edición, Demain A, y Davies J, 1999). Las condiciones adecuadas incluyen cultivar dichas células anfitrionas transformadas en un medio de cultivo que contiene nutrientes que cumplen los requerimientos de dichas células anfitrionas, tales como fuentes de carbono y nitrógeno, vitaminas y elementos traza, junto con un compuesto adecuado para seleccionar las células que contienen el vector de expresión, que puede contener un marcador selectivo, tal como un gen de resistencia a antibióticos. Las células se cultivan en condiciones en las que se consigue un crecimiento óptimo, respecto al pH, que varía típicamente de pH6-8, temperatura, que varía típicamente de 20-42°C y también con un suministro de oxígeno, que varía de 10-50%, siendo 30% un valor óptimo típico. Las células se crecen en dichas condiciones hasta que se alcanza una densidad adecuada, momento en el que se

añade el inductor en una cantidad suficiente para conseguir la inducción máxima del vector de expresión en las células y, consecuentemente, permitiendo la expresión de la proteína de interés. Se permite que la inducción continúe durante el tiempo óptimo, definido empíricamente, momento en el que se puede recoger la proteína de interés.

5 El método puede comprender además la purificación de la proteína heteróloga. Dichas etapas de purificación también son conocidas en la técnica (véase, por ejemplo, "Protein Purification", 2a Edición, Janson, J-C y Ryden, L, 1998). Las etapas básicas para la extracción de la proteína heteróloga producida intracelularmente implican típicamente una etapa de rotura de las células, mediante varios métodos, por ejemplo, homogeneización o lisis química o congelación-descongelación o ultrasonidos. La proteína intracelular que estaba presente en el extracto soluble podría capturarse
10 mediante una amplia variedad de etapas cromatográficas muy conocidas en la técnica, por ejemplo, cromatografía de intercambio iónico, de interacción hidrofóbica, de afinidad, cromatografía de fase reversa, de exclusión por tamaño o de filtración en gel y ultrafiltración. Típicamente, se utilizan etapas cromatográficas adicionales para purificar más la proteína heteróloga de proteínas e impurezas contaminantes de la célula anfitriona, lo que resulta en una proteína altamente purificada. La proteína intracelular que estaba presente en la parte insoluble del extracto, por ejemplo, presente
15 como cuerpos de inclusión, podría solubilizarse utilizando varios métodos conocidos en la técnica, utilizando compuestos tales como guanidina, urea o dodecil sulfato sódico, por ejemplo y purificarse posteriormente utilizando una o más etapas cromatográficas como se ha mencionado anteriormente. La proteína heteróloga que se ha producido en el espacio periplásmico, utilizando un vector de expresión que utiliza un péptido líder como se ha descrito anteriormente, podría liberarse de dicho espacio utilizando un choque osmótico, para liberar los contenidos del espacio periplásmico
20 en el medio, utilizando compuestos tales como sacarosa o sulfato de magnesio o lisozima/EDTA, después de lo cual se utiliza una o más etapas cromatográficas para purificar dicha proteína.

Un séptimo aspecto de la invención proporciona la utilización de uno o más vectores o secuencias de ácido nucleico, según uno cualquiera del segundo o tercer aspecto de la invención en la producción de una proteína heteróloga.

25 Todas las características preferidas de los diferentes aspectos de la invención se aplican entre sí, *mutatis mutandis*.

Es el objeto de esta invención proporcionar vectores de expresión, que comprenden un promotor y un represor regulador, que procede de un operón para el metabolismo de la xantosina en *E. coli* para la expresión de proteínas heterólogas con valor comercial. En una realización particular de la invención, la proteína reguladora xapR, junto con su promotor, se aíslan e introducen en un vector de expresión junto con un gen heterólogo unido operativamente al promotor del gen xapA. El vector de expresión resultante se transforma en una célula anfitriona adecuada, que se
30 crece posteriormente en condiciones adecuadas para conseguir un crecimiento óptimo. La expresión de la proteína heteróloga no se produce hasta la adición de xantosina, que posteriormente activa la proteína xapR, permitiendo que se una a la región promotora xapA, permitiendo la expresión de la proteína heteróloga a partir del promotor xapA.

Es un objeto adicional de esta invención proporcionar nuevas secuencias del sitio de unión ribosomal para el operón de la xantosina, que tienen unos niveles incrementados de expresión de proteínas heterólogas, en los que se han hecho cambios específicos respecto al sitio natural de unión ribosomal del operón xapA, como se detalla en la primera realización, y en el que se ha utilizado un método de cribado para identificar aquellas mutaciones que resultan
40 en unos niveles de expresión incrementados.

Es un objeto adicional de esta invención proporcionar un método para la producción de proteínas heterólogas utilizando un sistema de expresión inducible con nucleótidos, tales como xantosina. En los ejemplos siguientes, se describe la construcción de los vectores de expresión inducibles con xantosina y la utilidad de la invención se ilustra
45 utilizando la producción de la hormona de crecimiento humana (hGH) y la interleuquina-4 humana en *E. coli*.

Es evidente para un experto en la técnica que pueden expresarse otros genes en el sistema descrito en la presente memoria. Además, dicho sistema podría estar integrado en el genoma de *E. coli*, u otros organismos, o podría construirse un sistema de expresión similar utilizando genes homólogos de operones de la xantosina de otros organismos, incluyendo géneros tales como *Salmonella*, *Pseudomonas*, *Bacillus* o *Streptococcus*, etc.

La presente invención se describe con referencia a las figuras adjuntas, en las que:

55 La Figura 1a muestra la secuencia de nucleótidos de la región promotora xapA de *E. coli*.

(RBS - Sitio de Unión Ribosomal)

La Figura 1b muestra la secuencia de nucleótidos de la región promotora xapA de *E. coli* junto con sitios de
60 restricción nuevos *HindIII* y *NdeI*.

La Figura 1c muestra la secuencia de nucleótidos de la región promotora xapA de *E. coli* junto con sitios de restricción nuevos *HindIII* y *NcoI*.

65 La Figura 2 muestra la secuencia de nucleótidos y de aminoácidos del gen xapR de *E. coli* (mayúscula) y la región promotora xapR (minúscula cursiva) junto con nuevos sitios de restricción *BamHI* y *KpnI*.

La Figura 3 muestra los vectores pUC19x, pXap1a y pXap1b.

ES 2 295 433 T3

La Figura 4 muestra el mapa de pXap1b- β gal.

La Figura 5 muestra los mapas de Xap 1a-hGH y de pXap-IL4.

La Figura 6 muestra el análisis proteico de fracciones pXap-hGH:

(A)-SDS-PAGE; (B)-Transferencia Western.

Carriles:

1) proteína hGH estándar; 2) Marcador de Peso Molecular;

Carriles 3-7 Medio LB

3) inducción 0 h; 4) inducción 2 h; 5) inducción 3 h; 6) inducción 4 h; 7) inducción 6 h

Carriles 8-12 Medio Definido

8) inducción 0 h; 9) inducción 2 h; 10) inducción 3 h; 11) inducción 4 h; 12) inducción 6 h.

En la parte A de la figura, la proteína de interés es la que está presente más claramente y abundantemente (indicada con una flecha) y que ha migrado con el estándar de proteína del carril 1.

En la parte B, la Transferencia Western muestra (mediante un anticuerpo específico) que la proteína esperada está presente.

La Figura 7 muestra el análisis proteico de las fracciones pXap-IL4 a 20°C y 28°C:

(A) SDS-PAGE; (B) Transferencia Western.

Carril M: Marcadores Proteicos.

Carriles 1-4: 20C, inducción 0h, 2 h, 5 h y toda la noche.

Carriles 5-7: 28C, después de una inducción de 2 h, 5 h y toda la noche. Carril Std: IL4 estándar.

En la parte A de la figura, la proteína de interés es la que está presente más claramente y abundantemente (indicada con una flecha) y que ha migrado con el estándar de proteína del carril marcado Std en la parte derecha.

En la parte B, la Transferencia Western muestra (mediante un anticuerpo específico) que la proteína esperada está presente.

La presente invención se describe ahora con referencia a los ejemplos siguientes no limitativos:

Ejemplos

Los materiales y métodos utilizados se describen a continuación y la invención se ilustra en los ejemplos. Como apoyo a la mayoría de los métodos, se hace referencia a los libros siguientes: Sambrook *et al.*, (1989) "Molecular Cloning: A Laboratory Manual", 2ª Edición y Miller, J. (1972) "Experiments in Molecular Genetics", ambos de Cold Spring Harbor Press, Cold Spring Harbor, NY, EEUU y Dieffenbach C.W. y Dveksler G.S. (1995); *PCR Primer: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York.

Ejemplo 1

Construcción de un Vector de Expresión Inducible por Xantosina

Amplificación por PCR del gen XapR (proteína reguladora)

El gen XapR, junto con su propia secuencia promotora, se amplificó del ADN genómico de *E. coli* a partir de la cepa MG1655 (ATCC # 700926) utilizando el cebador directo xap2 (SEQ_ID1) 5'-ACGGTACCTTTTGCTATCTGC GATTTGCG-3' y el cebador inverso xap5 (SEQ_ID2) 5'-CTCATTAAGGATCCGCGGCTCTGCTCTTCAG-3'. La mezcla de reacción contenía Tampón *Pfu* (20 mM Tris-HCl pH8,8 a 25°C; 10 mM sulfato de amonio; 10 mM cloruro de potasio; 0,1% Tritón X-100; 0,1 mg/ml albúmina de suero bovino y 2 mM sulfato de magnesio), 0,25 mM de cada dATP, dCTP, dGTP y dTTP, 500 ng de ADN genómico de *E. coli*, 50 pmoles de cada cebador, 1,25 unidades de polimerasa de ADN *Pfu* (MBI Fermentas) y agua hasta un volumen final de 50 μ l. El método de amplificación por PCR utilizado fue el siguiente:

ES 2 295 433 T3

1 ciclo de:

94°C 5 minutos de desnaturalización inicial

5 72°C parada (adición de la polimerasa)

14 ciclos de:

10 65°C 1 minuto de hibridación (-1°C descenso por ciclo)

72°C 5 minutos de elongación

15 94°C 1 minuto de desnaturalización

19 ciclos de:

20 55°C 1 minuto de hibridación

72°C 5 minutos de elongación (+0,02 minutos de incremento por ciclo)

94°C 1 minuto de desnaturalización

25 55°C 1 minuto de hibridación

72°C 7 minutos de elongación final

30 Se amplificó un fragmento de 1,2 kb correspondiente al gen xapR. El fragmento de ADN resultante de esta reacción de PCR se analizó en un gel de 0,8% TAE/agarosa teñido con 0,1 µg/ml de bromuro de etidio. El fragmento se purificó posteriormente del gel.

35 *Amplificación por PCR del promotor xapA*

La región promotora de xapA se amplificó por PCR del ADN genómico de *E. coli* a partir de la cepa MG1655 utilizando el cebador directo xap1 (SEQ_ID3) 5'-CCAAGCTTAGCATAATTCCCTATGCCGATC-3' y el cebador inverso fosforilado xp4 (SEQ_ID4) 5'-GAACCTGAGACATATGTATCCTTTTG-3' o xapNco R fosforilado (SEQ_ID5) 5'-GTGGTCACCATGGGTATCCTTTTCTG TAG G-3' utilizando la mezcla de reacción descrita anteriormente. El método de amplificación utilizado fue el siguiente:

1 ciclo de:

45 94°C 5 minutos de desnaturalización inicial

72°C Parada (adición de la polimerasa)

50 35 ciclos de:

65°C 1 minuto de hibridación

72°C 0,5 minutos de elongación

55 94°C 1 minuto de desnaturalización

1 ciclo de:

60 65°C 1 minuto de hibridación

72°C 3 minutos de elongación final

65 Se amplificó un fragmento de 270 pb correspondiente al promotor xapA. El fragmento de ADN resultante de estas reacciones de PCR se analizó en un gel de 1% TAE/agarosa teñido con 0,1 µg/ml de bromuro de etidio y de nuevo se purificó del gel.

Construcción del Vector de Expresión

El primer vector que se construyó fue pUC19X. Para construir este vector, pUC19 se digirió con *SacI* y *NdeI* y las protuberancias resultantes se rellenaron utilizando la polimerasa de ADN T4 (Promega) según las instrucciones del fabricante. El fragmento de 2,4 kb resultante se autoligó y se transformó en células XL1-Blue competentes (Stratagene). Los productos de PCR purificados, el gen *xapR* y el promotor *xapA*, como se ha descrito anteriormente, se insertaron en pUC19X utilizando las endonucleasas de restricción *BamHI* y *KpnI* del gen *xapR* y *HindIII* y *HincII* para el promotor *xapA*. Los plásmidos resultantes se denominaron pXap1a y pXap1b y se diferencian sólo por la presencia de un único sitio *NcoI* o *NdeI*. Los mapas de la construcción inicial y de las construcciones del vector de expresión se proporcionan en la Figura 3.

La construcción se diseña de manera que la proteína *xapR*, bajo el control de su propio promotor (incorporado en el producto de PCR como se ha descrito anteriormente), se expresa constitutivamente. El promotor *xapA* se inserta en una región adecuada del vector y permite la clonación de genes heterólogos en el sitio *NcoI* (pXap1a) o en el sitio *NdeI* (pXap1b). Esto garantiza que dicho gen heterólogo esté posicionado correctamente respecto a la región promotora del promotor *xapA*, respecto al sitio de inicio de la transcripción, el sitio de unión ribosomal y el sitio de unión de la Polimerasa de ARN del promotor, lo que permite la expresión correcta de la proteína heteróloga después de la inducción con xantosina, pero no la expresión no inducida en ausencia de xantosina.

Ejemplo 2

*Creación de mutantes del sitio de unión ribosomal y análisis de actividad**Análisis de la expresión de pXap1b utilizando β -galactosidasa*

El gen de la β -galactosidasa de *E. coli* (β -Gal) se obtuvo inicialmente mediante la amplificación por PCR del gen β -gal del ADN genómico de *E. coli* a partir de la cepa MG1655 utilizando los métodos de amplificación descritos anteriormente, con los cebadores siguientes - β -Gal-1 (SEQ_ID6) 5'-ATATGGGCCCCATGGATCCCGTCGTTTTAC-3' y β -Gal-2 (SEQ_ID7) 5'-AGTGTGAAGCTTATTATTTTGTACACCAG-3'. El fragmento de PCR se ligó entonces en el vector de expresión pSE420 (Invitrogen) utilizando digestiones dobles *NcoI*/*HindIII* y se utilizaron para transformar células DH5 químicamente competentes (Gibco-BRL). Los clones positivos se identificaron mediante análisis de restricción.

Posteriormente, el gen de la β -galactosidasa se clonó en uno de los vectores pXap, concretamente pXap1b, mediante digestión de pSE420- β -Gal con *HindIII* seguido de la obtención de extremos romos del fragmento utilizando la enzima de Klenow y una digestión posterior con *NcoI*.

El fragmento purificado se introdujo entonces en pXap1b digerido de manera similar, para estar en marco con el promotor *xapA*. Después de ligar y de transformar, los clones positivos se identificaron por la aparición de color azul cuando se plaquean en placas LXXX (placas de Agar LB/Ampicilina suplementadas con 1 mg/ml de Xantosina, 80 μ g/ml de X-gal). La xantosina induce la expresión del gen β -gal que forma posteriormente un producto de color azul por la hidrólisis del sustrato X-gal. Las colonias azules positivas se cribaron para detectar actividad β -Gal mediante el ensayo de Miller (1972) y una de dichas colonias se denominó plásmido pXap1b- β -Gal (Figura 4).

Mutagénesis de la región del sitio de unión ribosomal de pXap1b- β -Gal

Con el fin de desarrollar vectores nuevos con una expresión incrementada, se realizó PCR mutagénica en el plásmido pXap1b- β -Gal, con el fin de cambiar el sitio de unión ribosomal (RBS) natural y analizar si existe una expresión incrementada. Se construyeron dos de dichos mutantes, RBS1 y RBS2. RBS1 se diseñó para reemplazar el RBS natural de *xapA*, AGGA, con la secuencia RBS consenso más fuerte AGGAGG. RBS2 se diseñó para introducir una secuencia RBS igual de fuerte AGGAGA, más una base adicional entre el RBS y el codón de inicio, ya que se ha observado que la alteración del espacio entre el RBS y el codón de inicio puede afectar los niveles de expresión, bien de manera positiva o negativa, en otros sistemas de expresión (Marquis *et al.*, 1986; Chen *et al.*, 1994).

El plásmido pXap1b- β -Gal se amplificó por PCR utilizando los cebadores siguientes:

pXapRBS1A

5'-CCTACAGAAAAAGGAGGTACCCATGGATCCCGTCG-3' (SEQ_ID12) junto con

pXapRBS1B

5'-CGGGATCCATGGGTACCTCCTTTTCTGTAGGGTGG-3' (SEQ_ID13) para el mutante RBS1 y

ES 2 295 433 T3

pXapRBS2A

5'-CCCTACAGAAAAAGGAGATATCCCATGGATCCCGTCGTTTTACAACG-3' (SEQ_ID14) junto con

pXapRBS2B

5'-CGACGGGATCCATGGGATATCTCCTTTTTCTGTAGGGTGGAATCTAACG-3' (SEQ_ID15) para el mutante RBS2, en reacciones independientes, utilizando las condiciones siguientes.

La mezcla de reacción fue como se ha descrito en el ejemplo 1, excepto en que se utilizó 1 ng de pXap1b- β -Gal purificado como molde. Las condiciones de reacción utilizadas fueron:

1 ciclo de:

94°C 5 minutos de desnaturalización inicial

72°C parada (adición de la polimerasa)

18 ciclos de:

55°C 1 minuto de hibridación

72°C 15 minutos de elongación

94°C 1 minuto de desnaturalización

La endonucleasa de restricción *DpnI* (Fermentas) se añadió a la mezcla de reacción y se incubó a 37°C durante 2 horas. Se utilizaron 5 μ l de la mezcla de reacción para transformar células DH5 químicamente competentes de *E. coli* y las células transformadas se plaquearon en placas LAXX como se ha descrito anteriormente. Se cribaron varias colonias positivas azules y se identificaron por digestión de restricción.

Cribado de la actividad de expresión de pXap1b- β -Gal y de mutantes RBS

Se crecieron cultivos de pXap1b- β -Gal y de mutantes RBS positivos, transformados en células DH5 competentes, a 37°C en medio definido suplementado con ampicilina y se indujeron con 1 mg/ml de xantosina (concentración final) con el fin de determinar su actividad β -galactosidasa, determinada por el ensayo ONPG (Miller, 1972) (Tabla 1).

TABLA 1

Comparación de la actividad β -galactosidasa de pXap1b- β -Gal y de los mutantes RBS pXap1b- β -Gal-RBS1 y pXap1b- β -Gal-RBS2

Plásmido/Cepa	Actividad β -Galactosidasa, en Unidades Miller
pXap1b- β -Gal / DH5	16.900
pXap1b- β -Gal-RBS1 / DH5	21.500
pXap1b- β -Gal-RBS2 / DH5	24.600

Se piensa (Chen *et al.*, 1994) que un espaciado de 5 nucleótidos (nt) entre el RBS y el codón de inicio ATG es el espaciado óptimo en los promotores de *E. coli*. Este es el caso en el RBS de xapA natural. La alteración del RBS de pXap1b- β -Gal con la secuencia consenso AGGAGG resultó en unos niveles de expresión mayores (RBS1) y, sorprendentemente, un incremento del espaciado entre la secuencia consenso y el codón de inicio (RBS2) incrementó aún más los niveles de expresión.

Regiones RBS de los plásmidos RBS1 y RBS2, comparadas con el plásmido pXap1b- β -Gal.

CCACCCTACAGAAAAAGGA_{tatccc}ATGGATC-pXap1b- β -Gal (espaciado de 5nt)

CCACCCTACAGAAAAAGGAGG_{tatccc}ATGGATC-pXap1b- β -Gal-RBS1 (espaciado de 5nt)

CCACCCTACAGAAAAAGGAGAtatcccATGGATC-pXap1b- β -Gal-RBS2 (espaciado de 6nt)

Ejemplo 3

*Expresión de genes heterólogos**Expresión de la hormona de crecimiento humana*

El gen de la hormona de crecimiento humana (hGH) se amplificó por PCR a partir del plásmido phg107 (ATCC#40011) utilizando cebadores hGH-Nde (SEQ_ID8) 5'-AAGAATCCCATATGTTCCCAACCATTCCCTTAT CC-3' y hGH-Rev (SEQ_ID9) 5'-CGCGGATCCAAGCTTATTAGAAGCCACAGCTGCCCTCC-3'.

Las condiciones y los parámetros de amplificación fueron como se ha descrito en el ejemplo 1. El fragmento amplificado se clonó en pXap1a utilizando los sitios de restricción *NdeI/BamHI* (Figura 5). La producción de la hormona de crecimiento humana por este nuevo clon se ensayó en la cepa TG1 de *E. coli*, a 30°C tanto en medio complejo (LB) como definido (Miller, 1992).

La expresión de hGH se indujo mediante la adición de 1 mg/ml de xantosina a los cultivos. La presencia y la identidad de la hormona de crecimiento humana producida se confirmó mediante SDS-PAGE y mediante análisis por transferencia Western utilizando un anticuerpo monoclonal y se detectó utilizando IgG-AP de cabra anti-ratón de BIO-RAD según las instrucciones del fabricante (Figura 6). Los análisis de densitometría mostraron que la proteína hGH recombinante representaba más del 20% de las proteínas celulares totales, incluso después de sólo 2 horas de inducción, en los dos medios.

Expresión de la interleuquina-4 humana

El gen de la interleuquina-4 (IL-4) se amplificó por PCR a partir de un plásmido hIL-4 - pPlc299hIL4 (Dr. Nico Mertens, VIB, Bélgica) utilizando cebadores IL4-Nde (SEQ_ID10) 5'-AAGAATCCCATATGCACAAGTGCGATAT CACC-3' e IL4-Rev (SEQ_ID11) 5'-AAGGATCCCAAGCTTAGCTCGAACACTTTGAATATTTC-3'.

Las condiciones y los parámetros de amplificación fueron como se ha descrito anteriormente (Sección: amplificación por PCR del gen XapR). El fragmento amplificado se clonó en pXap2a utilizando los sitios de restricción *NdeI/BamHI* (Figura 5). La presencia y la identidad de la interleuquina-4 se evaluó a 20°C y 28°C en medio LB. La expresión de IL-4 se indujo mediante la adición de 1 mg/ml de xantosina a los cultivos y se confirmó mediante SDS-PAGE y mediante análisis por transferencia Western utilizando un anticuerpo monoclonal y se detectó utilizando IgG-AP de cabra anti-ratón de BIO-RAD según las instrucciones del fabricante (Figura 7). Los análisis de densitometría mostraron que la IL-4 recombinante constituía más del 30% de las proteínas celulares totales, a las dos temperaturas.

Referencias

Bohannon, D y Soneshin, A (1989); *J Bact* 171: 4718-27

Chen E, et al. (1994); *Nuc Acid Res* 22: 4953-7

Demain A, y Davies J (1999) "Manual of Industrial Microbiology and Biotechnology" 2a Edición. ASM Press, Washington DC, EEUU.

Hammer-Jespersen, K et al. (1980); *Mol Gen Genet* 179: 3414

Henikoff, S et al. (1988); *PNAS* 85: 6602-6

Janson, J-C y Rydén, L (1998) "Protein Purification: Principles. High-Resolution Methods, and Applications". John Wiley & Sons, NY, EEUU.

Jorgensen, C. y Dandanell, G., (1999); *J. Bact.* 181: 14, 4397-4403

Maquis D, et al. (1986); *Gene* 42: 175-83

Miller J.H. (1972); *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor. Nueva York.

Miller J. H. (1992); *A Short Course in Bacterial Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York.

Nygaard, (1983); p 27.93; *Metabolism of nucleotides, nucleosides and nucleobases in microorganisms*. Academic Press, Londres.

Seeger, C. et al. (1995); *J Bact* 177:19 5506-16

Apéndice 1- Secuencias de ADN

SEQ_ID1

Cebador Xap2

ACGGTACCTTTTGCTATCTGCGATTGCG

SEQ_ID2

Cebador Xap5

CTCATTTAAAAGGATCCGCGGCTCTGCTCTTCAG

SEQ_ID3

Cebador Xap1

CCAAGCTTAGCATAATTCCCTATGCCGATC

SEQ_ID4

Cebador Xap4

GAACCTGAGACATATGTATCCTTTTG

SEQ_ID5

Cebador xapNcoR

GTGGTCACCATGGGTATCCTTTTTCTGTAGG

SEQ_ID6

Cebador B-Gal-1

ATATGGGCCCATGGATCCCGTCGTTTTAC

SEQ_ID7

Cebador B-Gal-2

AGTGTGAAGCTTATTATTTTTGACACCAG

SEQ_ID8

Cebador hGH-Nde

5 **AAGAATCCCATATGTTCCCAACCATTCCTTATCC**

SEQ_ID9

Cebador hGH-Rev

10 **CGCGGATCCAAGCTTATTAGAAGCCACAGCTGCCCTCC**

SEQ_ID10

Cebador IL4-Nde

20 **AAGAATCCCATATGCACAAGTGCGATATCACC**

SEQ_ID11

Cebador IL4-Rev

25 **AAGGATCCCAAGCTTAGCTCGAACACTTTGAATATTTC**

SEQ_ID12

Cebador pXapRBS1A

35 **CCTACAGAAAAAGGAGGTACCCATGGATCCCGTCG**

SEQ_ID13

Cebador pXapRBS1B

40 **CGGGATCCATGGGTACCTCCTTTTCTGTAGGGTGG**

SEQ_ID14

Cebador pXapRBS2A

50 **CCCTACAGAAAAAGGAGATATCCCATGGATCCCGTCGTTTTACAACG**

SEQ_ID15

Cebador pXapRBS2B

55 **CGACGGGATCCATGGGATATCTCCTTTTCTGTAGGGTGGGAATCTAACG**

60

65

REIVINDICACIONES

1. Un vector que comprende un promotor que puede unirse operativamente a un gen que codifica una proteína heteróloga, en el que el promotor induce la expresión de cualquier proteína heteróloga unida operativamente, en presencia de nucleótidos, pudiendo tener o no dichos nucleótidos un grupo fosfato y en el que la secuencia promotora es un promotor xapA, que comprende un sitio de unión ribosomal que tiene la secuencia siguiente:
AGGAGG xxxxx o
AGGAGG xxxxxx o
AGGAGA xxxxx o
AGGAGA xxxxxx
2. Un vector, según la reivindicación 1, que es un plásmido o un bacteriófago.
3. Un vector, según la reivindicación 1, que es un vector de expresión inducible.
4. Un vector, según una cualquiera de las reivindicaciones 1 a 3, en el que el promotor induce la expresión de la proteína heteróloga unida operativamente en presencia de xantosina.
5. Un vector, según una cualquiera de las reivindicaciones 1 a 4, que comprende además un elemento regulador que regula la expresión por el promotor.
6. Un vector, según la reivindicación 5, en el que el elemento regulador es una secuencia de ácido nucleico.
7. Un vector, según la reivindicación 6, en el que el elemento regulador es de un gen xapR.
8. Un vector, según cualquiera de las reivindicaciones 6 ó 7, en el que un promotor adicional se une operativamente al elemento regulador.
9. Un vector, según una cualquiera de las reivindicaciones 6 a 8, en el que el promotor adicional es un promotor inducible o constitutivo.
10. Un vector, según la reivindicación 8, en el que el promotor adicional es de un promotor xapR.
11. Un vector, según una cualquiera de las reivindicaciones 1 a 10, que comprende un gen que codifica una proteína heteróloga.
12. Un vector, según la reivindicación 11, en el que el gen que codifica una proteína heteróloga codifica una citoquina, quimioquina, hormona, enzima o antígeno.
13. Un vector de expresión, que comprende un ácido nucleico aislado, que comprende el gen regulador xapR del operón de la xantosina junto con un promotor de un gen xapA.
14. Una secuencia promotora de xapA, que comprende un sitio de unión ribosomal que tiene la secuencia siguiente:
AGGAGG xxxxx o
AGGAGG xxxxxx o
AGGAGA xxxxx o
AGGAGA xxxxxx
15. Una célula anfitriona, que comprende uno o más vectores o secuencias de ácido nucleico, según una cualquiera de las reivindicaciones 1 a 14.
16. Un método para la expresión de una proteína heteróloga, comprendiendo el método cultivar las células anfitrionas, según la reivindicación 15, en condiciones que inducen la expresión de la proteína heteróloga.
17. Un método según la reivindicación 16, que comprende además purificar la proteína heteróloga.
18. Un método según cualquiera de las reivindicaciones 16 ó 17, en el que la proteína heteróloga es una citoquina, quimioquina, hormona, enzima o antígeno.

ES 2 295 433 T3

19. Utilización de uno o más vectores o secuencias de ácido nucleico, según una cualquiera de las reivindicaciones 1 a 14 en la producción de una proteína heteróloga.

5

10

15

20

25

30

35

40

45

50

55

60

65

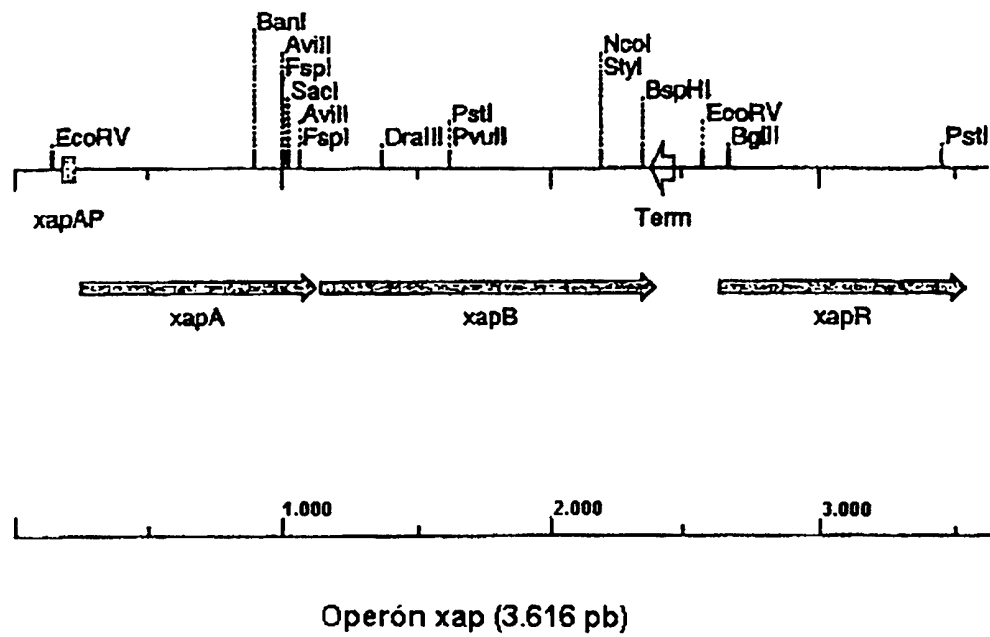


Figura 1

agcataattccctatgccgatcggtattagtagattgcttaataaagtggtttgctaaa

acgacatattgtaatgcaaaggaattaatatcccaatacagttttaccgtggattaa
región de unión 1 de xapR

tagagaaaagatatataatccatcgatatcgaatcgctattgaatgccgctgagaaaagt
región de unión 2 de xapR

gtattggtaagagccgggacaacctcgctataaaaaggcgtagattccaccctacagaa

aaaggatatgtATGTCTCAGGTTCAATTT
RBS Inicio del gen xapA

Figura 1a

1b

Sitio HindIII
ccAAGCTTagcataattccctatgccgatcggtattagtagattgcttaataaagtggtttgct
 aaaaacgacatatgtttaatgcaaaggaattaatatcg**ccaatacagttttt**accgtggattaa
región de unión 1 de xapR
 tagagaaaagatatataatccatcgatatcgaatcgccattgaatgccgctgag**aaaagtgtat**
región de unión 2 de xapR
tggtaagagccgggacaacctcgctataaaaaggcggttagattccaccctacagaaaaaggata
RBS
 Sitio NdeI
CATATGtctcaggttc

1c

Sitio Hind III
ccAAGCTTagcataattccctatgccgatcggtattagtagattgcttaataaagtggtttgct
 aaaaacgacatatgtttaatgcaaaggaattaatatcg**ccaatacagttttt**accgtggattaa
región de unión 1 de xapR
 tagagaaaagatatataatccatcgatatcgaatcgccattgaatgccgctgag**aaaagtgtat**
región de unión 2 de xapR
tggtaagagccgggacaacctcgctataaaaaggcggttagattccaccctacagaaaaaggata
RBS
 Sitio NcoI
ccCATGGTgaccac

Figuras 1b y 1c

ES 2 295 433 T3

sitio *Bam*HI

1 tca tta aaa gga tcc gcg gct ctg ctc ttc aga gct gct ttt
atg ata aag gtt aat tag
61 tga aag ata ttt tat ttt ctt ccg gcg ttg gtt ttg gta ttg
gtg ctt tgt tta caa tcg
121 tgc gct tgc cta ttc ctg tac caa atg tat tac cgg gta tat
tat cca ttg tgt tta tgt
181 atg tcg gat atc tgg tgg tga aat act tta tgc cat gat aat
tta ata cga tgt att tat

241 tat ATG GAG CAC TTA ATT ATG GAA CGC GTA TAC AGA ACA GAT
CTT AAG TTG CTC CGT TAT

Met Glu His Leu Ile Met Glu Arg Val Tyr Arg Thr Asp
Leu Lys Leu Leu Arg Tyr

301 TTT CTT GCC GTA GCG GAA GAG TTG CAT TTT GGC CGC GCA GCA
GCG CGT TTA AAT ATG TCT

Phe Leu Ala Val Ala Glu Glu Leu His Phe Gly Arg Ala Ala
Ala Arg Leu Asn Met Ser

361 CAG CCT CCG CTC AGC ATT CAT ATT AAA GAG CTG GAA AAT CAA
CTC GGC ACG CAG CTT TTT

Gln Pro Pro Leu Ser Ile His Ile Lys Glu Leu Glu Asn Gln
Leu Gly Thr Gln Leu Phe

421 ATT CGC CAT TCG CGC AGC GTC GTA CTG ACA CAC GCG GGC AAA
ATC TTG ATG GAA GAA TCG

Ile Arg His Ser Arg Ser Val Val Leu Thr His Ala Gly Lys
Ile Leu Met Glu Glu Ser

481 CGT CGA TTG CTG GTG AAT GCT AAT AAT GTA TTG GCT CGG ATT
GAA CAA ATA GGT CGG GGA

Arg Arg Leu Leu Val Asn Ala Asn Asn Val Leu Ala Arg Ile
Glu Gln Ile Gly Arg Gly

541 GAA GCA GGG CGG ATT GAA CTC GGC GTA GTG GGA ACG GCA ATG
TGG GGA CGG ATG CGC CCG

Glu Ala Gly Arg Ile Glu Leu Gly Val Val Gly Thr Ala Met
Trp Gly Arg Met Arg Pro

601 GTT ATG CGG CGA TTC CTC AGG GAA AAT CCT AAC GTT GAC GTT
CTT TTT CGC GAA AAG ATG

Val Met Arg Arg Phe Leu Arg Glu Asn Pro Asn Val Asp Val
Leu Phe Arg Glu Lys Met

661 CCC GCG ATG CAA ATG GCC TTG CTG GAA CGC CGC GAA CTT GAT
GCC GGG ATC TGG CGA ATG

Pro Ala Met Gln Met Ala Leu Leu Glu Arg Arg Glu Leu Asp
Ala Gly Ile Trp Arg Met

721 GCG ACA GAA CCA CCG ACT GGT TTT ACC AGC TTA CGG TTG CAT
GAA TCG GCG TTT CTG GTG

Ala Thr Glu Pro Pro Thr Gly Phe Thr Ser Leu Arg Leu His
Glu Ser Ala Phe Leu Val

781 GCG ATG CCT GAA GAG CAT CAT CTC TCA TCA TTT TCC ACC GTC
CCG CTG GAA GCG CTA CGT

Ala Met Pro Glu Glu His His Leu Ser Ser Phe Ser Thr Val
Pro Leu Glu Ala Leu Arg

841 GAC GAG TAT TTT GTT ACA ATG CCG CCC GTT TAC ACT GAC TGG
GAT TTT TTG CAG CGA GTT

Asp Glu Tyr Phe Val Thr Met Pro Pro Val Tyr Thr Asp Trp
Asp Phe Leu Gln Arg Val

901 TGC CAG CAG GTG GGA TTT TCA CCG GTT GTT ATT CGC GAA GTT
AAT GAA CCG CAA ACG GTA

Cys Gln Gln Val Gly Phe Ser Pro Val Val Ile Arg Glu Val
Asn Glu Pro Gln Thr Val

961 CTC GCT ATG GTC AGT ATG GGC ATT GGT ATC ACA TTG ATA GCG
GAC AGC TAC GCA CAA ATG

Leu Ala Met Val Ser Met Gly Ile Gly Ile Thr Leu Ile Ala
Asp Ser Tyr Ala Gln Met

1.021 AAC TGG CCA GGT GTC ATT TTC CGT CCG CTC AAA CAG CGC ATC
CCT GCA GAT TTA TAT ATT

Asn Trp Pro Gly Val Ile Phe Arg Pro Leu Lys Gln Arg Ile
Pro Ala Asp Leu Tyr Leu

1.081 GTT TAT GAA ACA CAG CAG GTG ACG CCC GCG ATG GTT AAG CTG
TTG GCT GCA TTG ACA CAG

Ile Val Tyr Glu Thr Gln Gln Val Thr Pro Ala Met Val Lys
Leu Ala Ala Leu Thr Gln

1.141 tag gtg agg aga cag gaa ttg ccg gaa ata gta gtc tgc taa
gtg aag cgc atc agg caa

sitioKpnI

1.201 ttt cgc aaa tcg cag ata gca aaa ggt acc gt

Figura 2

Figura 3

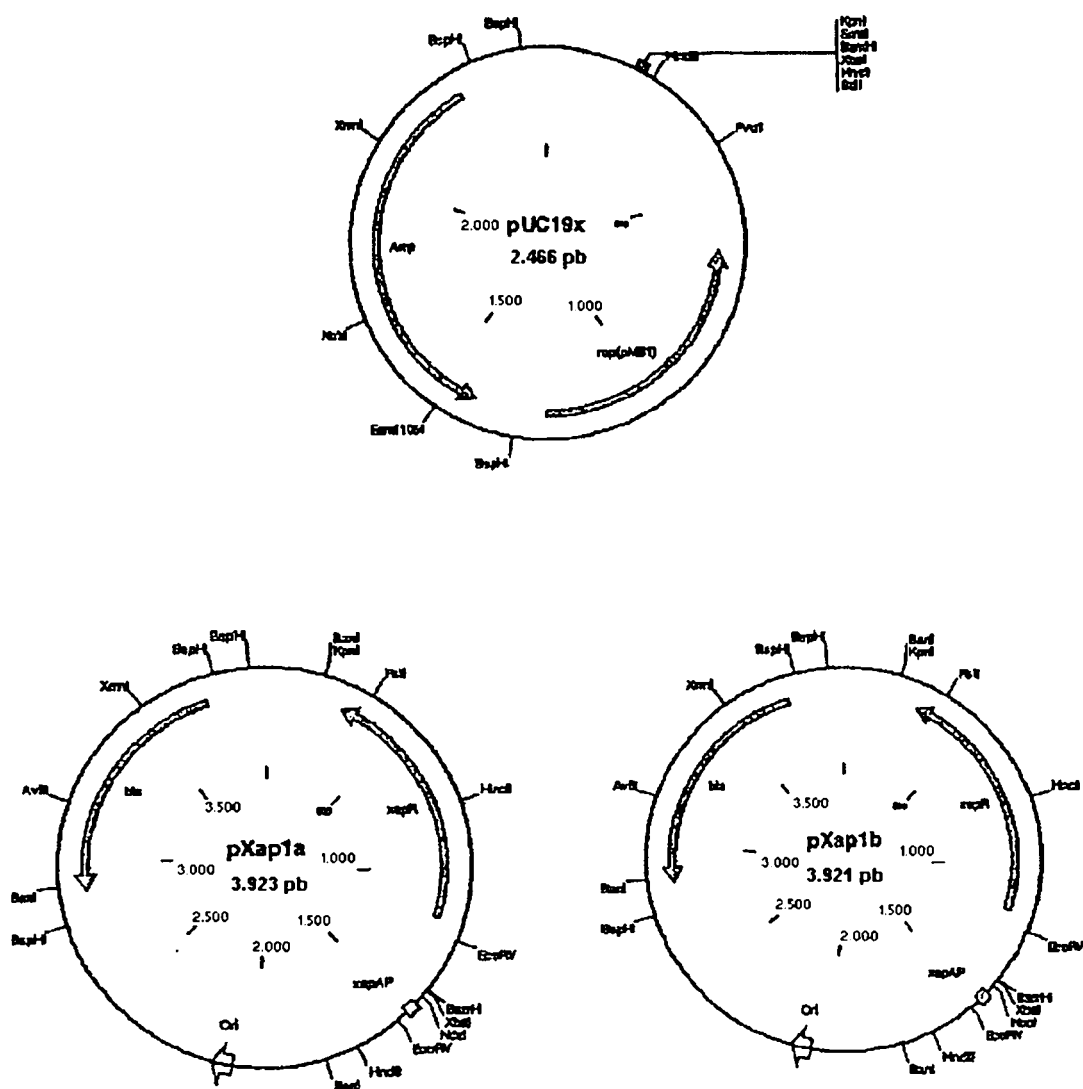
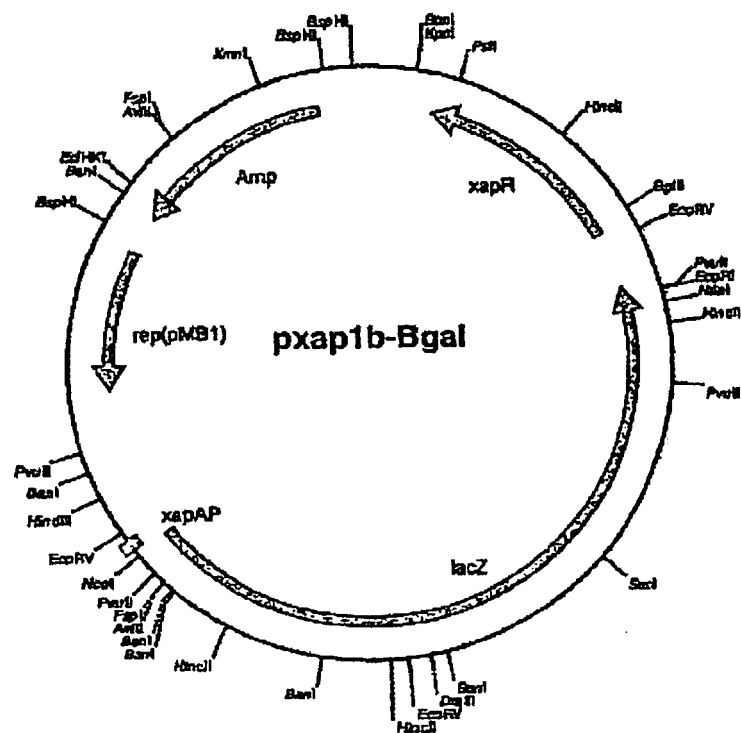


Figura 4



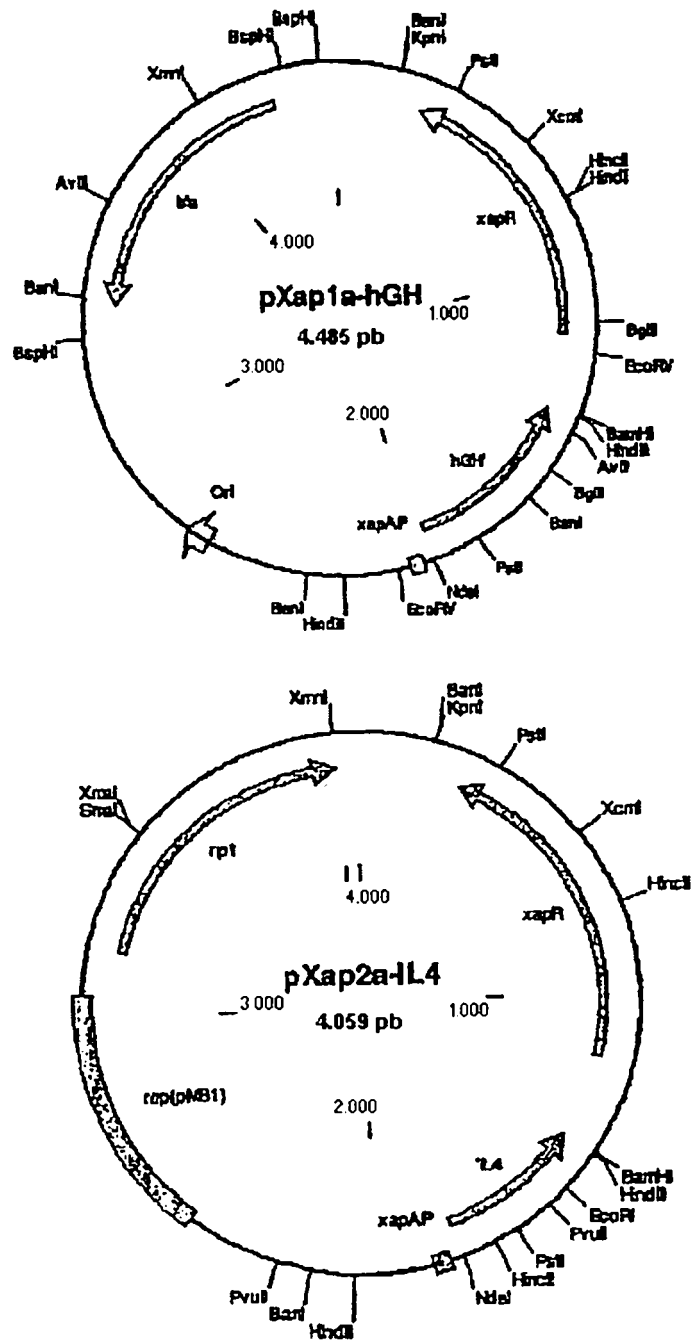


Figura 5

Figura 6

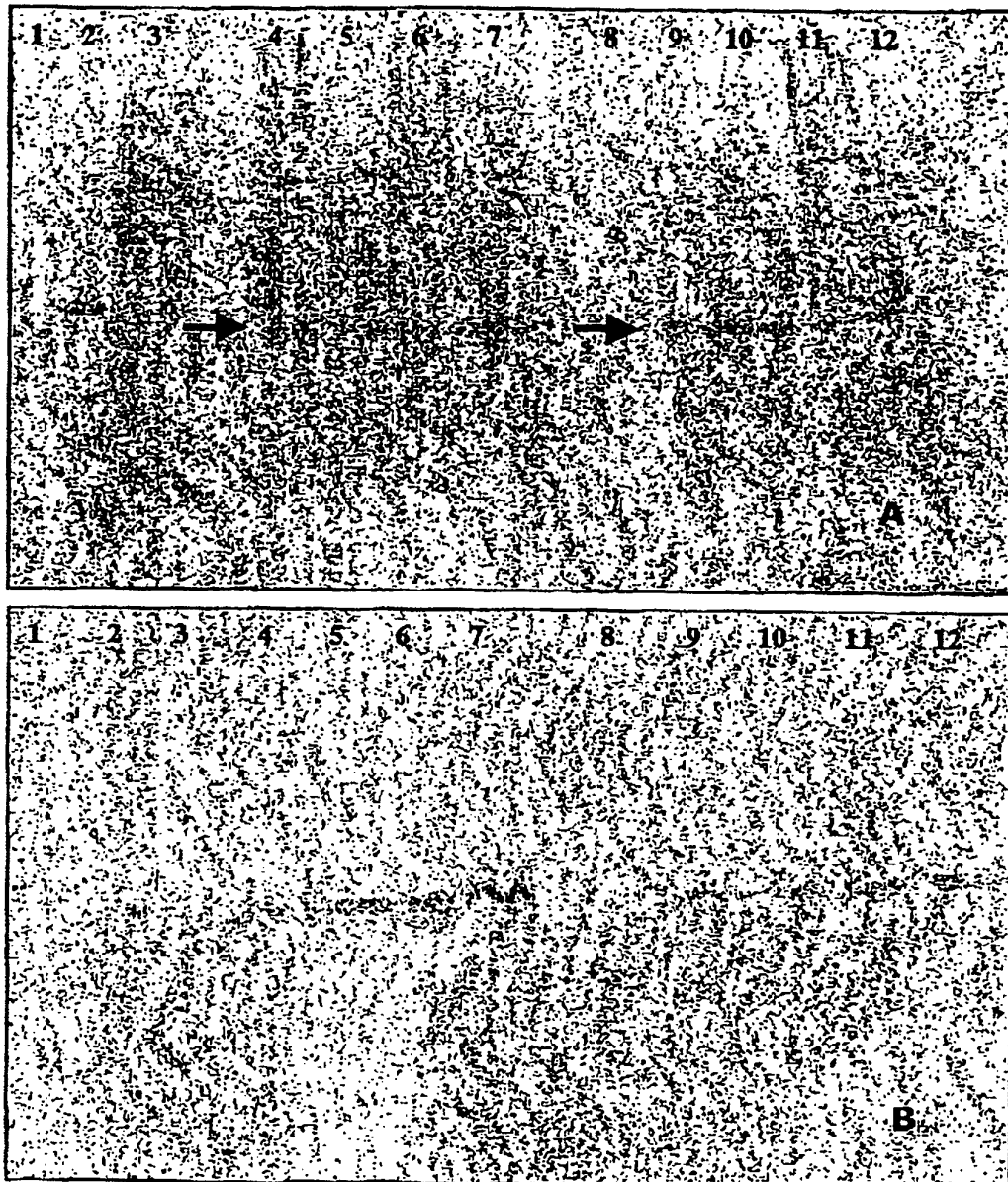
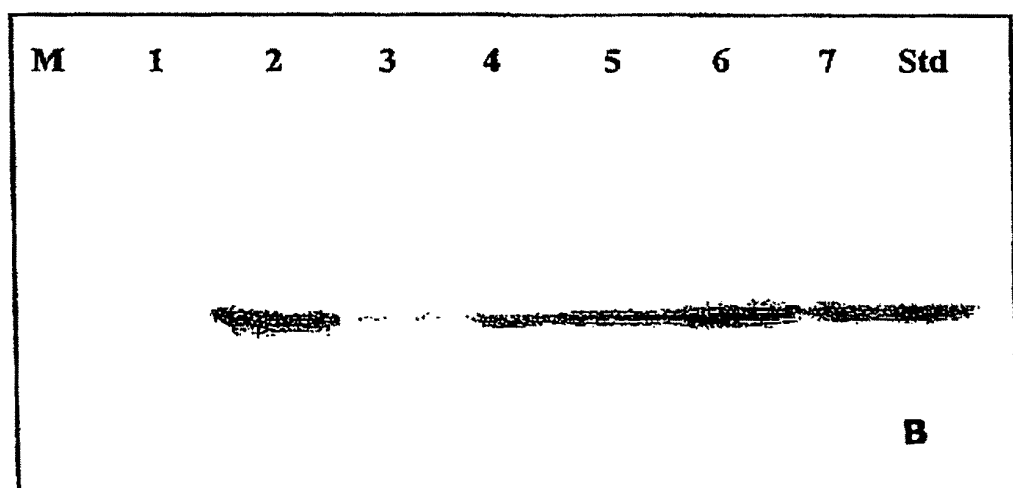
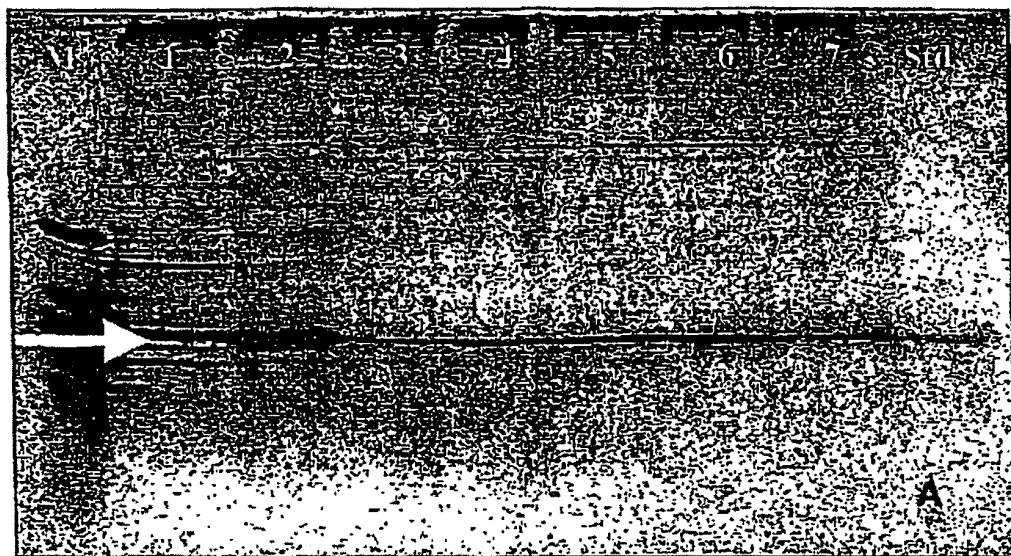


Figura 7



ES 2 295 433 T3

LISTA DE SECUENCIAS

	<110> Biotecnol S.A.	
5	<120> vectores y promotores para expresión génica	
	<130> -	
10	<140> PCT/GB2002/005888	
	<141> 2002-12-23	
	<150> PT102704	
15	<151> 21-12-2001	
	<160> 20	
20	<170> PatentIn version 3.2	
	<210> 1	
	<211> 29	
25	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
30	<223> cebador oligonucleotídico para PCR	
	<400> 1	
35	acggtacctt ttgctatctg cgatttgcg	29
	<210> 2	
	<211> 33	
40	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
45	<223> cebador oligonucleotídico para PCR	
	<400> 2	
50	ctcattaaaa ggatccgcgg ctctgctctt cag	33
	<210> 3	
	<211> 30	
55	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
60	<223> cebador oligonucleotídico para PCR	
	<400> 3	
65	ccaagcttag cataattccc tatgccgatc	30

ES 2 295 433 T3

	<210> 4	
	<211> 26	
	<212> ADN	
5	<213> Secuencia Artificial	
	<220>	
	<223> cebador oligonucleotídico para PCR	
10	<400> 4	
	gaacctgaga catatgtatc cttttg	26
15	<210> 5	
	<211> 31	
	<212> ADN	
20	<213> Secuencia Artificial	
	<220>	
	<223> cebador oligonucleotídico para PCR	
25	<400> 5	
	gtggtcacca tgggtatcct tttctgtag g	31
30	<210> 6	
	<211> 29	
	<212> ADN	
35	<213> Secuencia Artificial	
	<220>	
	<223> cebador oligonucleotídico para PCR	
40	<400> 6	
	atatgggccc atggatcccg tcgttttac	29
45	<210> 7	
	<211> 29	
	<212> ADN	
50	<213> Secuencia Artificial	
	<220>	
	<223> cebador oligonucleotídico para PCR	
55	<400> 7	
	agtggaagc ttattatttt tgacaccag	29
60	<210> 8	
	<211> 35	
	<212> ADN	
65	<213> Secuencia Artificial	

ES 2 295 433 T3

	<220>	
	<223> cebador oligonucleotídico para PCR	
5	<400> 8	
	aagaatccca tatgttccca accattccct tatcc	35
10	<210> 9	
	<211> 38	
	<212> ADN	
	<213> Secuencia Artificial	
15	<220>	
	<223> cebador oligonucleotídico para PCR	
20	<400> 9	
	cgcggatcca agcttattag aagccacagc tgcctcc	38
25	<210> 10	
	<211> 32	
	<212> ADN	
	<213> Secuencia Artificial	
30	<220>	
	<223> cebador oligonucleotídico para PCR	
35	<400> 10	
	aagaatccca tatgcacaag tgcgatatca cc	32
40	<210> 11	
	<211> 38	
	<212> ADN	
	<213> Secuencia Artificial	
45	<220>	
	<223> cebador oligonucleotídico para PCR	
50	<400> 11	
	aaggatccca agcttagctc gaacacttg aatatttc	38
55	<210> 12	
	<211> 35	
	<212> ADN	
	<213> Secuencia Artificial	
60	<220>	
	<223> cebador oligonucleotídico para PCR	
65	<400> 12	
	cctacagaaa aaggaggtac ccatggatcc cgtcg	35

ES 2 295 433 T3

<210> 13
 <211> 36
 <212> ADN
 5 <213> Secuencia Artificial

 <220>
 <223> cebador oligonucleotídico para PCR
 10
 <400> 13

 cgggatccat gggtagctcc ttttctgta gggtagg 36
 15
 <210> 14
 <211> 47
 <212> ADN
 20 <213> Secuencia Artificial

 <220>
 <223> cebador oligonucleotídico para PCR
 25
 <400> 14

 ccctacagaa aaaggagata tcccatggat cccgtcggtt tacaacg 47
 30
 <210> 15
 <211> 49
 <212> ADN
 35 <213> Secuencia Artificial

 <220>
 <223> cebador oligonucleotídico para PCR
 40
 <400> 15

 cgacgggatc catgggatat ctctttttc ttagggtagg aatctaacg 49
 45
 <210> 16
 <211> 1232
 <212> ADN
 50 <213> *Escherichia coli*

 <220>
 <221> CDS
 55 <222> (241)..(1140)

 60

 65

ES 2 295 433 T3

<400> 16

	tcattaaag gatccgcggc tctgctcttc agagctgctt ttatgataaa ggttaattag	60
5	tgaaagatat tttattttct tccggcggtg gttttgggtat tgggtgctttg tttacaatcg	120
	tgcgcttgcc tattcctgta ccaaattgtat taccgggtat attatccatt gtgtttatgt	180
	atgtcggata tctgggtggtg aaatacttta tgccatgata atttaataacg atgtatttat	240
10	tat atg gag cac tta att atg gaa cgc gta tac aga aca gat ctt aag	288
	Tyr Met Glu His Leu Ile Met Glu Arg Val Tyr Arg Thr Asp Leu Lys	
	1 5 10 15	
15	ttg ctc cgt tat ttt ctt gcc gta gcg gaa gag ttg cat ttt ggc cgc	336
	Leu Leu Arg Tyr Phe Leu Ala Val Ala Glu Glu Leu His Phe Gly Arg	
	20 25 30	
	gca gca gcg cgt tta aat atg tct cag cct ccg ctc agc att cat att	384
20	Ala Ala Ala Arg Leu Asn Met Ser Gln Pro Pro Leu Ser Ile His Ile	
	35 40 45	
	aaa gag ctg gaa aat caa ctc ggc acg cag ctt ttt att cgc cat tcg	432
	Lys Glu Leu Glu Asn Gln Leu Gly Thr Gln Leu Phe Ile Arg His Ser	
	50 55 60	
25	cgc agc gtc gta ctg aca cac gcg ggc aaa atc ttg atg gaa gaa tcg	480
	Arg Ser Val Val Leu Thr His Ala Gly Lys Ile Leu Met Glu Glu Ser	
	65 70 75 80	
	cgt cga ttg ctg gtg aat gct aat aat gta ttg gct cgg att gaa caa	528
30	Arg Arg Leu Leu Val Asn Ala Asn Asn Val Leu Ala Arg Ile Glu Gln	
	85 90 95	
	ata ggt cgg gga gaa gca ggg cgg att gaa ctc ggc gta gtg gga acg	576
35	Ile Gly Arg Gly Glu Ala Gly Arg Ile Glu Leu Gly Val Val Gly Thr	
	100 105 110	
	gca atg tgg gga cgg atg cgc ccg gtt atg cgg cga ttc ctc agg gaa	624
	Ala Met Trp Gly Arg Met Arg Pro Val Met Arg Arg Phe Leu Arg Glu	
	115 120 125	
40	aat cct aac gtt gac gtt ctt ttt cgc gaa aag atg ccc gcg atg caa	672
	Asn Pro Asn Val Asp Val Leu Phe Arg Glu Lys Met Pro Ala Met Gln	
	130 135 140	

45

50

55

60

65

ES 2 295 433 T3

	atg gcc ttg ctg gaa cgc cgc gaa ctt gat gcc ggg atc tgg cga atg	720
	Met Ala Leu Leu Glu Arg Arg Glu Leu Asp Ala Gly Ile Trp Arg Met	
	145 150 155 160	
5	gcg aca gaa cca ccg act ggt ttt acc agc tta cgg ttg cat gaa tcg	768
	Ala Thr Glu Pro Pro Thr Gly Phe Thr Ser Leu Arg Leu His Glu Ser	
	165 170 175	
10	gcg ttt ctg gtg gcg atg cct gaa gag cat cat ctc tca tca ttt tcc	816
	Ala Phe Leu Val Ala Met Pro Glu Glu His His Leu Ser Ser Phe Ser	
	180 185 190	
	acc gtc ccg ctg gaa gcg cta cgt gac gag tat ttt gtt aca atg ccg	864
	Thr Val Pro Leu Glu Ala Leu Arg Asp Glu Tyr Phe Val Thr Met Pro	
	195 200 205	
15	ccc gtt tac act gac tgg gat ttt ttg cag cga gtt tgc cag cag gtg	912
	Pro Val Tyr Thr Asp Trp Asp Phe Leu Gln Arg Val Cys Gln Gln Val	
	210 215 220	
20	gga ttt tca ccg gtt gtt att cgc gaa gtt aat gaa ccg caa acg gta	960
	Gly Phe Ser Pro Val Val Ile Arg Glu Val Asn Glu Pro Gln Thr Val	
	225 230 235 240	
	ctc gct atg gtc agt atg ggc att ggt atc aca ttg ata gcg gac agc	1008
	Leu Ala Met Val Ser Met Gly Ile Gly Ile Thr Leu Ile Ala Asp Ser	
	245 250 255	
25	tac gca caa atg aac tgg cca ggt gtc att ttc cgt ccg ctc aaa cag	1056
	Tyr Ala Gln Met Asn Trp Pro Gly Val Ile Phe Arg Pro Leu Lys Gln	
	260 265 270	
30	cgc atc cct gca gat tta tat att gtt tat gaa aca cag cag gtg acg	1104
	Arg Ile Pro Ala Asp Leu Tyr Ile Val Tyr Glu Thr Gln Gln Val Thr	
	275 280 285	
	ccc gcg atg gtt aag ctg ttg gct gca ttg aca cag taggtgagga	1150
	Pro Ala Met Val Lys Leu Leu Ala Ala Leu Thr Gln	
	290 295 300	
35	gacaggaatt gccggaata gtagtctgct aagtgaagcg catcaggcaa tttcgcaa	1210
	at	
	cgcagatagc aaaaggtacc gt	1232

<210> 17

<211> 300

45 <212> PRT

<213> *Escherichia coli*

<400> 17

50

Tyr Met Glu His Leu Ile Met Glu Arg Val Tyr Arg Thr Asp Leu Lys
1 5 10 15

55

Leu Leu Arg Tyr Phe Leu Ala Val Ala Glu Glu Leu His Phe Gly Arg
20 25 30

60

Ala Ala Ala Arg Leu Asn Met Ser Gln Pro Pro Leu Ser Ile His Ile
35 40 45

65

ES 2 295 433 T3

Lys Glu Leu Glu Asn Gln Leu Gly Thr Gln Leu Phe Ile Arg His Ser
 50 55 60
 5 Arg Ser Val Val Leu Thr His Ala Gly Lys Ile Leu Met Glu Glu Ser
 65 70 75 80
 10 Arg Arg Leu Leu Val Asn Ala Asn Asn Val Leu Ala Arg Ile Glu Gln
 85 90 95
 15 Ile Gly Arg Gly Glu Ala Gly Arg Ile Glu Leu Gly Val Val Gly Thr
 100 105 110
 20 Ala Met Trp Gly Arg Met Arg Pro Val Met Arg Arg Phe Leu Arg Glu
 115 120 125
 25 Asn Pro Asn Val Asp Val Leu Phe Arg Glu Lys Met Pro Ala Met Gln
 130 135 140
 30 Met Ala Leu Leu Glu Arg Arg Glu Leu Asp Ala Gly Ile Trp Arg Met
 145 150 155 160
 35 Ala Thr Glu Pro Pro Thr Gly Phe Thr Ser Leu Arg Leu His Glu Ser
 165 170 175
 40 Ala Phe Leu Val Ala Met Pro Glu Glu His His Leu Ser Ser Phe Ser
 180 185 190
 45 Thr Val Pro Leu Glu Ala Leu Arg Asp Glu Tyr Phe Val Thr Met Pro
 195 200 205
 50 Pro Val Tyr Thr Asp Trp Asp Phe Leu Gln Arg Val Cys Gln Gln Val
 210 215 220
 55 Gly Phe Ser Pro Val Val Ile Arg Glu Val Asn Glu Pro Gln Thr Val
 225 230 235 240
 60 Leu Ala Met Val Ser Met Gly Ile Gly Ile Thr Leu Ile Ala Asp Ser
 245 250 255
 65 Tyr Ala Gln Met Asn Trp Pro Gly Val Ile Phe Arg Pro Leu Lys Gln
 260 265 270
 70 Arg Ile Pro Ala Asp Leu Tyr Ile Val Tyr Glu Thr Gln Gln Val Thr
 275 280 285
 75 Pro Ala Met Val Lys Leu Leu Ala Ala Leu Thr Gln
 290 295 300

<210> 18

<211> 270

<212> ADN

65 <213> *Escherichia coli*

ES 2 295 433 T3

<400> 18

	ccaagcttag cataattccc tatgccgatc gttattagta gattgcttaa taaagtgggt	60
5	tgctaaaaac gacatattgt taatgcaaag gaattaatat cgccaataca gtttttaccg	120
	tggtattaata gagaaaagat ataaatccat cgatatcgaa tcgcctattg aatgccgctg	180
10	agaaaagtgt attggtaaga gccgggacaa cctcgctata aaaaggcggt agattccacc	240
	ctacagaaaa aggataccca tggtgaccac	270

<210> 19

15 <211> 272

<212> ADN

<213> *Escherichia coli*

20 <400> 19

	ccaagcttag cataattccc tatgccgatc gttattagta gattgcttaa taaagtgggt	60
25	tgctaaaaac gacatattgt taatgcaaag gaattaatat cgccaataca gtttttaccg	120
	tggtattaata gagaaaagat ataaatccat cgatatcgaa tcgcctattg aatgccgctg	180
	agaaaagtgt attggtaaga gccgggacaa cctcgctata aaaaggcggt agattccacc	240
30	ctacagaaaa aggatacata tgtctcaggt tc	272

<210> 20

35 <211> 270

<212> ADN

<213> *Escherichia coli*

40 <400> 20

	agcataattc cctatgccga tcgttattag tagattgctt aataaagtgg ttgctaaaa	60
45	acgacatatt gttaatgcaa aggaattaat atcgccaata cagtttttac cgtggattaa	120
	tagagaaaag atataaatcc atcgatatcg aatcgcttat tgaatgccgc tgagaaaagt	180
	gtattggtaa gagccgggac aacctcgcta taaaaaggcg ttagattcca ccctacagaa	240
50	aaaggatatg tatgtctcag gttcaat	270

55

60

65