



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 350 706**

51 Int. Cl.:
C12N 1/19 (2006.01)
C12N 15/81 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06772695 .0**
96 Fecha de presentación : **09.06.2006**
97 Número de publicación de la solicitud: **1893628**
97 Fecha de publicación de la solicitud: **05.03.2008**

54 Título: **Región de control de genes sintética.**

30 Prioridad: **14.06.2005 US 690321 P**

45 Fecha de publicación de la mención BOPI:
26.01.2011

45 Fecha de la publicación del folleto de la patente:
26.01.2011

73 Titular/es: **MERCK SHARP & DOHME Corp.**
126 East Lincoln Avenue
Rahway, New Jersey 07065, US

72 Inventor/es: **Brownlow, Michelle, Kaupas;**
Schultz, Loren;
Bryan, Janine;
Losada, Maria, C.;
Kurtz, Myra y
Jansen, Kathrin

74 Agente: **Carpintero López, Mario**

ES 2 350 706 T3

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

ANTECEDENTES DE LA INVENCION

Saccharomyces cerevisiae, o la levadura del panadero, se usa ampliamente como huésped para la expresión de una diversidad de polipéptidos heterólogos. Se han expresado muchas proteínas diferentes de una diversidad de especies en *S. cerevisiae*, algunas a niveles de >10% de la proteína celular total. Típicamente, la expresión se ha mediado por un plásmido que contiene una secuencia de ADN que codifica el polipéptido heterólogo y la región de control del gen que controla expresiones de genes en *S. cerevisiae* así como otras secuencias necesarias para la sección y amplificación del plásmido tanto en *S. cerevisiae* como en *Escherichia coli*. Como alternativa, también es posible integrar la secuencia codificante y la región de control del gen en un cromosoma de *S. cerevisiae* y conseguir un alto nivel de expresión.

Las regiones de control de genes utilizadas en la expresión de polipéptidos heterólogos en *S. cerevisiae* típicamente son las que se producen de forma natural en *S. cerevisiae*, por ejemplo, la región de control de genes para la expresión de los genes divergentes *GAL1* y *GAL10*. Por el contrario, se ha descubierto que, generalmente, las regiones de control de genes heterólogas, cuando se usan en células de *S. cerevisiae*, son inactivas o conducen a un inicio de la transcripción aberrante. Se ha propuesto que el uso de regiones de control de genes de *S. cerevisiae* es esencial para la expresión eficaz de genes heterólogos en células de *S. cerevisiae*. (Romanos y col., YEAST 8: 423-488 (1992)).

El documento US-A-5407822 (Sanofi) desvela una secuencia descrita como un "promotor artificial" que contiene las secuencias de activación cadena arriba UAS1 y UAS2 del promotor del gen *GAL7* de *S. cerevisiae* fusionadas al componente TATA y la región de inicio de la transcripción del promotor de *ADH2* de *S. cerevisiae*. El documento EP-A-0258067 (Delta Biotechnology Limited) desvela una secuencia descrita como un promotor híbrido de levadura que contiene la secuencia de activación cadena arriba del promotor de *GAL1/GAL10* ("GAL10UAS") y la región 5' no codificante del gen *PGK*.

Las referencias citadas en el presente documento no se reconocen como técnica anterior para la invención reivindicada.

SUMARIO DE LA INVENCION

La presente invención se refiere a una región de control de genes sintética que comprende una secuencia reguladora de genes de una cepa de levadura y un promotor de una cepa de hongo filamentoso localizado cadena abajo de la secuencia reguladora de genes. El promotor de hongo filamentoso puede reconocerse por los factores de transcripción generales

y la ARN polimerasa de la cepa de levadura. La secuencia reguladora de genes sintética puede regular la transcripción iniciada por el promotor de hongo filamentoso en la cepa de levadura.

De acuerdo con una realización de la presente invención, la cepa de levadura se selecciona entre un grupo que consiste en *Saccharomyces cerevisiae*, *Pichia pastoris*, *Hansenula polymorpha* y *Yarrowia lipolytica*. De acuerdo con una realización preferida de la presente invención, la cepa de levadura es *Saccharomyces cerevisiae*. De acuerdo con una realización alternativa de la presente invención, la cepa de levadura es *Pichia pastoris*.

La cepa de hongo filamentoso puede seleccionarse entre el grupo que consiste en *Ustilago maydis*, *Aspergillus nidulans* y *Penicillium purpurogenum*. El promotor de hongo filamentoso puede seleccionarse entre el grupo que consiste en los promotores del gen *DMP1* de *Ustilago maydis*, el gen *ArgB* de *Aspergillus nidulans* y el gen *XynA* de *Penicillium purpurogenum*.

De acuerdo con una realización de la presente invención, el sitio de unión es para un activador de genes de la cepa de levadura. De acuerdo con una realización preferida de la presente invención, el sitio de unión es un sitio de unión para la proteína GAL4 de *S. cerevisiae*. De acuerdo con otra realización preferida de la presente invención, el sitio de unión comprende una secuencia seleccionada entre el grupo que consiste en las SEC ID N°: 10, 11, 12 y 15, sitios de unión a GAL4 sintéticos.

La secuencia reguladora de genes puede comprender además un sitio de unión para un represor de genes de la cepa de levadura. De acuerdo con una realización de la presente invención, el sitio de unión para un represor de genes es un sitio de unión para la proteína MIG1 de *S. cerevisiae*. De acuerdo con una realización preferida de la presente invención, el sitio de unión para la proteína MIG1 comprende una secuencia seleccionada entre el grupo que consiste en las SEC ID N°: 18, 19 y 20.

De acuerdo con una realización preferida, la región de control de genes sintética puede comprender una secuencia seleccionada entre el grupo que consiste en las SEC ID N°: 21, 22, 23, 24, 25 y 26.

La presente invención proporciona vectores de expresión de ADN que comprenden la región de control de genes sintética, una secuencia codificante que codifica una proteína, un polipéptido o un péptido bajo el control de la región de control, y un marcador de selección de levaduras. La secuencia codificante puede codificar una secuencia de aminoácidos eucariota, procariota o viral. Si la cepa de levadura es *S. cerevisiae*, el marcador de selección puede seleccionarse entre el grupo que consiste en *LEU2*, *TRP1*, *URA3* e *HIS3*.

El vector de expresión de ADN puede comprender además una secuencia señal de

poliadenilación localizada cadena abajo de la secuencia codificante. El vector de expresión de ADN puede comprender además un terminador de la transcripción localizado cadena abajo de la secuencia codificante. El vector de expresión de ADN puede comprender además un origen de replicación de levadura, tal como uno basado en la secuencia de ADN de 2 micrómetros de *S. cerevisiae*. El vector de expresión de ADN puede comprender además un origen de replicación bacteriano.

La presente invención proporciona además una cepa de levadura que contiene el vector de expresión de ADN. La cepa de levadura puede seleccionarse entre el grupo que consiste en *Saccharomyces cerevisiae*, *Pichia pastoris*, *Hansenula polymorpha* y *Yarrowia lipolytica*.

Otras características y ventajas de la presente invención serán evidentes tras las descripciones adicionales proporcionadas en el presente documento, incluyendo los diferentes ejemplos. Los ejemplos proporcionados ilustran diferentes componentes y la metodología útil para poner en práctica la presente invención. Los ejemplos no limitan la invención reivindicada. Basándose en la presente revelación, el experto en la materia puede identificar y emplear otros componentes y metodologías útiles para poner en práctica la presente invención.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Figura 1. Los promotores de hongos filamentosos. Posible caja TATA, sitio de inicio de la transcripción indicado subrayado (en forma de onda). El supuesto sitio de unión a MIG1 está subrayado (línea discontinua).

Figura 1A: secuencia promotora de *XynA* truncada (SEC ID N°: 1).

Figura 1B: secuencia promotora de *DPM1* (SEC ID N°: 2).

Figura 1C: secuencia promotora de *ArgB* (SEC ID N°: 3).

Figura 2: regiones de control de genes sintéticas que comprenden el promotor de *XynA* truncado y una región reguladora de genes. Los supuestos sitios de unión a la proteína GAL4 están subrayados. Los supuestos sitios de unión a la proteína MIG1 están subrayados (línea discontinua).

Las Figuras 2A-2F son las secuencias de ADN de las regiones reguladoras de genes sintéticas.

Figura 2A: EE2-*XynA* (SEC ID N°: 21).

Figura 2B: EE21-*XynA* (SEC ID N°: 22).

Figura 2C: EE22-*XynA* (SEC ID N°: 23).

Figura 2D: EE24-*XynA* (SEC ID N°: 24).

Figura 2E: EE25-*XynA* (SEC ID N°: 25).

Figura 2F: EE26-*XynA* (SEC ID N°: 26).

Figura 3. La expresión de un polipéptido heterólogo en *S. cerevisiae*, dirigida por diversas

regiones de control de genes sintéticas y la región de control de los genes *GAL1-GAL10* natural de levadura, respectivamente. Las calles 2-5 son muestras de células de levadura inducidas con galactosa. Las calles 6-9 son muestras de células de levadura cultivadas en presencia de glucosa y no inducidas con galactosa.

5 DESCRIPCIÓN DETALLADA DE LA INVENCION

Como se ha analizado anteriormente, las regiones de control de genes que se usan ampliamente en la expresión de proteínas heterólogas en levadura son típicamente las que se producen de forma natural en *S. cerevisiae* (por ejemplo, la región de control de genes para la expresión de los genes divergentes *GAL1* y *GAL10*) y que normalmente controlan la expresión de los genes de *S. cerevisiae* correspondientes. Por el contrario, la presente invención proporciona una región de control de genes sintética, que no se sabe que se produzca en la naturaleza. Como se usa en el presente documento, una “región de control de genes” es la secuencia de ADN que controla la transcripción de un gen, es decir, la velocidad de inicio de la transcripción. (Alberts, y col., Molecular Biology of the Cell, 3ª Edición, Garland Publishing; 1994). La región de control de genes sintética de la presente invención puede usarse en la regulación de la expresión de un gen heterólogo en una cepa de levadura.

1. La Cepa de Levadura

Como se usa en el presente documento, el término “levadura” se refiere a cualquier organismo eucariota unicelular que carece de clorofila y tejido vascular, y se reproduce por gemación o escisión, tal como el género *Saccharomyces*. El género *Saccharomyces* se compone de una diversidad de especies, incluyendo *cerevisiae*, *carlsbergensis*, *norbensis*, *diastaticus*, *oviformis*, *uvarum*, *rouxii*, *montanus*, *kluveri* y *elongisporus*.

De acuerdo con una realización preferida de la presente invención, la cepa de levadura es *Saccharomyces cerevisiae*. Como se ha analizado anteriormente, *S. cerevisiae* se usa comúnmente como un huésped para la expresión de una diversidad de polipéptidos heterólogos. La célula huésped de *S. cerevisiae* usada para la expresión recombinante puede seleccionarse o modificarse por ingeniería genética para facilitar la expresión de genes recombinantes. Como el fondo genético de una cepa puede influir en gran medida en las propiedades de la cepa para la expresión de proteínas heterólogas, se deseaba construir cepas de levadura con diferentes fondos genéticos que también contuvieran varios marcadores genéticos deseables: mutación *mn9* para prevenir la hiperglicosilación de proteínas secretadas, y mutaciones de proteasa *prb1* y/o *pep4* para reducir problemas con la proteólisis (Joyce y col., Patente de Estados Unidos Nº 5.820.870). Para las regiones de control de genes sintéticos que contienen sitios de unión a GAL4, puede conseguirse la sobreexpresión del factor de transcripción GAL4 en la cepa huésped de *S. cerevisiae* para aumentar la expresión a

partir de las regiones de control (Hopper y col., Patente de Estados Unidos Nº 5.068.185).

Además, se ha demostrado que varios géneros de levadura, tales como *Hansenula*, *Candida*, *Torulopsis* y *Pichia* contienen rutas metabólicas similares para la utilización de metanol como una única fuente de carbono para el crecimiento. Además, las especies de otros
5 géneros de levadura pueden utilizar una diversidad de fuentes de carbono, incluyendo galactosa, para el crecimiento. La cepa de levadura puede proceder de las familias *Saccharomycetaceae* y *Cryptococcaceae*, incluyendo, pero sin limitación, especies de los géneros *Pichia*, *Candida*, *Hansenula*, *Torulopsis*, *Kluyveromyces* y *Saccharomycopsis*.

Específicamente, la cepa de levadura también puede seleccionarse entre el grupo que
10 consiste en *Pichia pastoris*, *Hansenula polymorpha*, *Yarrowia lipolytica*, *Kluyveromyces lactis*, *Schizosaccharomyces pombe* y *Schwanniomyces occidentalis*. Estas cepas de levadura también se usaron como organismo huésped para la expresión de genes heterólogos, de forma similar a *Saccharomyces cerevisiae*. (Véase, por ejemplo, Buckholz y Gleeson, Bio/Technology 9:1067-1072 (1991); Gellisen y Hollenberg, Gene 190: 87-97 (1997);
15 Dominguez, y col., Int. Microbiol: 1:131-142 (1998)).

De acuerdo con una realización de la presente invención, la cepa de levadura es *Pichia pastoris*. (Véase, por ejemplo, Werten, y col., Yeast 15: 1087-1096 (1999); Cregg, y col., Mol. Biotechnol. 16, 23-52 (2000)).

De acuerdo con otra realización de la presente invención, la cepa de levadura es
20 *Hansenula polymorpha*.

De acuerdo con una realización de la presente invención, la cepa de levadura es *Yarrowia lipolytica*. (Muller, y col., Yeast 14: 1267-1283 (1998); Madzak, y col., Journal of Biotechnology, 109: 63-81 (2004)).

2. La Región de Control Sintética

La presente invención proporciona una región de control de genes sintética que
25 comprende una secuencia reguladora de genes sintética que comprende un sitio de unión para una proteína reguladora de genes en una cepa de levadura, y un promotor de una cepa de hongo filamentoso localizado cadena abajo de la secuencia reguladora de genes. La región de control de genes sintética es una secuencia de ADN específica. La secuencia del sitio de unión
30 preferentemente no es una secuencia de las regiones reguladoras de genes naturales de la cepa de levadura.

Como se usa en el presente documento, la expresión “hongo filamentoso” se refiere a cualquier organismo eucariota multicelular que carece de clorofila y tejido vascular, y forma una masa corporal de hifas filamentosas ramificadas que con frecuencia producen cuerpos
35 fructíferos especializados. Los ejemplos de hongos filamentosos incluyen *Ustilago maydis*,

Aspergillus nidulans y *Penicillium purpurogenum*.

2.1. El Promotor

La transcripción en células eucariotas requiere que la ARN polimerasa y factores de transcripción generales se ensamblen al promotor. Como se usa en el presente documento, un “promotor” es la secuencia de ADN en la que se ensamblan la ARN polimerasa y factores de transcripción generales. Un promotor puede comprender una caja TATA y el punto de inicio de la transcripción. Una caja TATA es una secuencia corta de pares de bases T-A y A-T que se reconoce por TFHD, un factor de transcripción general. El punto de inicio de la transcripción típicamente se localiza 25 pares de bases cadena abajo de la caja TATA en células de mamífero. *ibid.* En levadura, la distancia de la caja TATA al inicio de la transcripción típicamente es de aproximadamente 100 pares de bases.

El promotor de la región de control de genes sintética es un promotor de una cepa de hongo filamentoso, que puede reconocerse por los factores de transcripción generales y la ARN polimerasa de la cepa de levadura, tal como *S. cerevisiae* (Romanos y col., YEAST 8: 423-488, (1992)).

Los ejemplos de promotores de hongos filamentosos incluyen los del gen XynA, el gen *DPM1* y el gen *ArgB* (Fig. 1) (Zimmerman, y col., Yeast 12: 765-771(1996), Upshall, y col., Mol. Gen. Genet. 204: 349-354 (1986), Chavez, y col., Biol. Res. 34: 217-226 (2001)).

2.2. La Secuencia Reguladora de Genes

Una secuencia reguladora de genes contiene al menos un sitio de unión para una proteína reguladora de genes, cuya presencia en el ADN afecta a la velocidad de inicio de la transcripción. Las proteínas reguladoras de genes incluyen activadores de genes, que estimulan la expresión génica, y represores de genes, que reprimen la expresión génica. La velocidad de inicio de la transcripción puede aumentarse o reducirse por la unión de proteínas reguladoras de genes a las secuencias reguladoras de genes correspondientes. Como se usa en el presente documento, un “sitio de unión” se refiere a una secuencia de ADN en la región reguladora de genes a la que se une específicamente una proteína reguladora de genes.

Los sitios de unión para una proteína reguladora de genes se identificaron originalmente a partir de regiones reguladoras de genes naturales. A partir de los sitios de unión identificados para una proteína reguladora de genes dada, puede deducirse una supuesta secuencia de sitio de unión consenso. De acuerdo con una realización preferida de la presente invención, el sitio de unión es un sitio de unión sintético. Como se usa en el presente documento, un “sitio de unión sintético” se refiere al sitio de unión que no se identifica en regiones reguladoras de genes naturales, sino que más bien se construye teniendo cuenta la supuesta secuencia consenso para la proteína reguladora de genes.

Una secuencia reguladora de genes eucariota puede localizarse adyacente al promotor, a una distancia cadena arriba del mismo o incluso cadena abajo del promotor. (Alberts, y col., *Molecular Biology of the Cell*, 3ª Edición, Garland Publishing; 1994). Por ejemplo, las secuencias reguladoras de genes del *S. cerevisiae* normalmente se localizan unos pocos
5 cientos de pares de bases cadena arriba del promotor. Las secuencias reguladoras de genes de *S. cerevisiae* pueden estimular y/o reprimir la expresión del gen bajo su control por unión a factores de transcripción. Los ejemplos de secuencias reguladoras de genes incluyen las de los genes *GAP(TDH)*, *PGK*, *TPI*, *PHO5*, *ADH2* y *CUP1* (Romanos y col., YEAST 8: 423-488, (1992)).

10 Las secuencias reguladoras de genes también determinan si una región de control de genes es constitutiva, es decir, dirige la expresión continua del gen, o inducible, es decir, dirige la expresión del gen en respuesta a una señal. Por ejemplo, las regiones de control de genes para los genes *GAP(TDH)*, *PGK* y *TPI* son regiones de control de genes constitutivas (Romanos y col., YEAST 8: 423-488 (1992)).

15 La secuencia reguladora de genes preferentemente es una inducible, que regula la transcripción en respuesta a una o más señales. En una diversidad de sistemas de expresión de levadura recombinantes, tales como *S. cerevisiae*, se muestra que la expresión de muchos polipéptidos heterólogos diferentes con frecuencia es perjudicial para la célula huésped. Por consiguiente, puede haber una presión selectiva contra la expresión de dichos polipéptidos
20 heterólogos, de tal forma que las únicas células que se acumulan en un aumento a escala de dicho cultivo recombinante sean las que no expresan el polipéptido heterólogo o expresan una cantidad tan pequeña de polipéptido heterólogo que el cultivo se convierte en una fuente poco económica de ese polipéptido. Un esquema óptimo para el aumento a escala de dicho cultivo recombinante sería mantener una expresión mínima o nula del gen heterólogo durante la
25 expansión del cultivo a un gran volumen y alta densidad celular y después inducir la expresión máxima del gen heterólogo únicamente en la etapa final de crecimiento del cultivo antes del aislamiento del producto. De esta manera, la región de control de genes sintética para la expresión de un gen recombinante preferentemente es una región de control de genes inducible.

30 La secuencia reguladora génica de los genes *GAL1-GAL10* de *S. cerevisiae* es una secuencia reguladora de genes inducible. La secuencia reguladora de genes *GAL1-GAL10* responde tanto a galactosa como a glucosa. Está implicada en la regulación del metabolismo de galactosa en *S. cerevisiae*, por medio de la expresión controlada de las enzimas responsables para la utilización de galactosa como fuente de carbono, por ejemplo, *GAL1*
35 (galactoquinasa) y *GAL10* (uridina difosfogalactosa-4-epimerasa) (Lohr y col., FASEB J. 9:

777-787 (1995)). En ausencia de galactosa, se detecta muy poca expresión de estas enzimas. Si las células inicialmente crecen en medio que contiene glucosa, y se añade galactosa al cultivo, estas enzimas se inducen de forma coordinada al menos 1000 veces tras la reducción de glucosa del medio. Se ha mostrado que esta inducción se produce a nivel de la transcripción del ARN mensajero.

Ciertos experimentos han definido las regiones de control de genes que se necesitan y son suficientes para la inducción por galactosa y, por lo tanto, son útiles para dirigir la expresión de genes heterólogos en *Saccharomyces cerevisiae*. Los genes *GAL1* y *GAL10* se transcriben de forma divergente. La región de control de genes *GAL1-GAL10* es una secuencia de aproximadamente 606 pb localizada entre los dos genes, que contiene los promotores de *GAL1* y *GAL10* y la secuencia reguladora de genes inducible que responde a galactosa y glucosa. La secuencia de 606 pb se usa en pGAL110, un plásmido de expresión de levadura de aproximadamente 12,0 kpb (Hofmann, y col., Virology 209: 506-518 (1995)), para dirigir la expresión de genes heterólogos clonados cadena abajo en *S. cerevisiae*.

La secuencia reguladora de genes de *GAL1-GAL10* contiene los sitios de unión para la proteína de unión GAL4, un activador de genes de levadura, que responde a galactosa. En la tabla 1 se presentan ejemplos de sitios de unión para la proteína GAL 4.

Tabla 1-1

SEC ID N°:	Sitios de unión a GAL4	Referencias
4	CGGATTAGAAGCCGCCG	West, y col., Mol. Cell Biol. 4: 2467-2478 (1984)
5	CGGGTGACAGCCCTCC G	<i>ibid.</i>
6	AGGAAGACTCTCCTCCG	<i>ibid.</i>
7	CGCGCCGCACTGCTCC G	<i>ibid.</i>
8	CGGAGGACTGTCCTCCG	Bram, y col., EMBO J. 5: 603-608 (1986)
9	CGGAGCACTCTCCTCCG	Melcher, y col., Gene 247: 53-61 (2000)

Las secuencias de las SEC ID N°: 4-9 indicadas en la tabla 1-2 son sitios de unión a GAL4 identificados en las regiones reguladoras de genes naturales de *S. cerevisiae*.

Tabla 1-2

SEC ID N°:	Sitios de unión a GAL4	
10	CGGATGACACTCCTCCG	Supuesta nueva secuencia
11	CGGGCCACTGTCGTCCG	<i>ibid.</i>
12	GGTCGAGGCCATCCCC G	<i>ibid.</i>
13	CGGACGACTGTGGTCCG	Bram, y col., EMBO J. 5: 603-608 (1986)
14	CGGGCGACACTTCTCCG	Bram, y col., EMBO J. 5: 603-608 (1986)
15	AGGTCGAGGCCATCCCCG	<i>ibid.</i>

Las secuencias de las SEC ID N°: 10-12 y SEC ID N°: 15 indicadas en la tabla 1-2 son sitios de unión a GAL4 sintéticos, que no se identifican en las regiones reguladoras de genes naturales de *S. cerevisiae*. Las secuencias de las SEC ID N°: 13 y 14 son sitios de unión a GAL4 identificados en las regiones reguladoras de genes naturales de *S. cerevisiae*.

La secuencia reguladora de genes de *GAL1-GAL10* también contiene los sitios de unión para la proteína MIG1, un represor de gen de levadura, que responde a glucosa. En la tabla 2 se presentan ejemplos de sitios de unión para la proteína MIG1.

Tabla 2-1

SEC ID N°:	Sitios de unión para MIG1	Referencias
16	TATTTCTGGGGTA	Nehlin, y col. EMBO J. 10: 3373-3377 (1991)
17	GGTTTGTGGGGCC	<i>ibid.</i>

Las secuencias de las SEC ID N°: 16-17 son sitios de unión a MIG1 identificados en las regiones reguladoras de genes naturales de *S. cerevisiae*.

Tabla 2-2

SEC ID N°:	Sitios de unión para MIG1	
18	GCATACCGGGGCC	Supuesta nueva secuencia
19	ATTATGTGGGGTA	<i>ibid.</i>
20	AAAATCTGGGGAA	<i>ibid.</i>

Las secuencias de las SEC ID N°: 18-20 son sitios de unión a MIG1 sintéticos que no se identifican en las regiones reguladoras de genes naturales de *S. cerevisiae*.

Se construyeron regiones reguladoras de genes sintéticas para incluir un promotor de hongo filamentoso, y sitios de unión para factores de transcripción de levadura, tales como GAL4 y MIG1. De acuerdo con una realización preferida, las regiones reguladoras de genes

sintéticas se reprimen en presencia de glucosa y se activan en presencia de galactosa y en ausencia de glucosa.

En la Figura 2 se muestran algunos ejemplos de regiones reguladoras de genes sintéticas, incluyendo EE2-XynA (Fig. 2A, SEC ID N°: 21), EE21-XynA (Fig. 2B, SEC ID N°: 22). EE22-XynA (Fig. 2C, SEC ID N°: 23), EE24-XynA (Fig. 2D, SEC ID N°: 24), EE25-XynA (Fig. 2E, SEC ID N°: 25), y EE26-XynA (Fig. 2F, SEC ID N°: 26).

3. Los Vectores de Expresión

La presente invención proporciona un vector de expresión de ADN, que comprende una región de control de genes sintética, una secuencia de ADN que codifica un polipéptido bajo el control de la región de control y un marcador de selección de *S. cerevisiae*.

3.1 Los Componentes del Vector

El vector puede ser un vector de integración, que puede integrarse en un cromosoma de levadura, o un vector episómico. Para mantenerse en las células de *S. cerevisiae*, un vector episómico necesita contener un origen de replicación para replicarse. Los ejemplos de orígenes de replicación incluyen secuencias de replicación autónomas de levadura (*ARS*) y secuencias del plásmido circular 2 μ nativo de *S. cerevisiae*. El vector *ARS* puede estabilizarse por secuencias centroméricas (*CEN*). Típicamente, el número de copias de un vector *ARSICEN* es de aproximadamente 1 ó 2 por célula, mientras que el número de copias del vector basado en 2 μ es de aproximadamente 10 ó 40 por célula. (Romanos y col., YEAST 8:423-488 (1992)).

El vector de expresión puede comprender un marcador de selección de levadura, es decir, un gen que codifica un polipéptido para la selección fenotípica en levadura. El marcador de selección puede ser un marcador de selección auxotrófico, incluyendo *LEU2*, *TRP1*, *URA3* y *HIS3*, que se usan en cepas mutantes correspondientes que son auxotróficas para leucina, triptófano, uracilo e histidina, respectivamente. El marcador de selección puede ser un marcador de selección dominante, tal como *CUP1*, que confiere a la levadura resistencia al cobre. *Ibid*.

El vector de expresión puede necesitar replicarse en células bacterianas para la clonación molecular. De esta manera, el vector de expresión puede contener un origen de replicación bacteriano. El vector de expresión también puede comprender un marcador de selección bacteriano, es decir, un gen que codifica un polipéptido para la selección fenotípica en bacterias. El marcador de selección bacteriano puede ser un marcador de resistencia a antibióticos. Los ejemplos de marcadores de selección bacterianos incluyen genes que confieren resistencia a ampicilina, kanamicina, tetraciclina y cloranfenicol, respectivamente.

El vector de expresión puede contener un terminador de la transcripción localizado

cadena debajo del gen heterólogo para una formación eficaz del extremo 3' del ARNm. Los terminadores pueden ser los de *TRP1*, *ADH1*, *GAP*, *MF α 1* y *CYC1*. *Ibid.*

3.2. La Secuencia Codificante Heteróloga Expresada en *S. cerevisiae*

Una secuencia codificante que codifica una proteína, un polipéptido o un péptido puede ponerse bajo control de una región de control en el vector de expresión de ADN. La secuencia codificante puede codificar una secuencia de aminoácidos eucariota, procariota o viral.

Debido a la degeneración del código genético, puede usarse un gran número de secuencias de ácido nucleico codificantes diferentes para codificar una secuencia de aminoácidos particular. La degeneración del código genético se produce porque casi todos los aminoácidos se codifican por diferentes combinaciones de tripletes de nucleótidos o "codones". Los aminoácidos se codifican por los siguientes codones:

A=Ala=Alanina: codones GCA, GCC, GCG, GCU

C=Cis=Cisteína: codones UGC, UGU

D=Asp=Ácido aspártico: codones GAC, GAU

E=Glu=Ácido glutámico: codones GAA, GAG

F=Phe=Fenilalanina: codones UUC, UUU

G=Gli=Glicina: codones GGA, GGC, GGG, GGU

H=His=Histidina: codones CAC, CAU

I=Ile=Isoleucina: codones AUA, AUC, AUU

K=Lis=Lisina: codones AAA, AAG

L=Leu=Leucina: codones UUA, UUG, CUA, CUC, CUG, CUU

M=Met=Metionina: codón AUG

N=Asn=Asparagina: codones AAC, AAU

P=Pro=Prolina: codones CCA, CCC, CCG, CCU

Q=Gln=Glutamina: codones CAA, CAG

R=Arg=Arginina: codones AGA, AGG, CGA, CGC, CGG, CGU

S=Ser=Serina: codones AGC, AGU, UCA, UCC, UCG, UCU

T=Thr=Treonina: codones ACA, ACC, ACG, ACU

V=Val=Valina: codones GUA, GUC, GUG, GUU

W=Trp=Triptófano: codón UGG

I=Tir=Tirosina: codones UAC, UAU

Si se desea, la expresión de un polipéptido heterólogo en un huésped particular puede aumentarse por medio de la optimización de codones. La optimización de codones incluye el uso de codones más preferidos. En la técnica se conocen técnicas para la optimización de codones en diferentes huéspedes.

La optimización de codones para un huésped particular se realiza reemplazando codones que tienen un nivel de uso bajo o moderado con codones que tienen un alto nivel de uso. El porcentaje de codones óptimos presentes en una secuencia codificante puede variar. En diferentes realizaciones, el número de codones óptimos (incluyendo codones inicialmente presentes y codones introducidos) es de al menos un 50%, al menos un 75%, al menos un 95% o un 100% del número total de codones.

La optimización de codones puede realizarse como se indica a continuación:

1. Para un codón particular, comparar la frecuencia de codones de tipo silvestre con la frecuencia de uso de codones en general por genes de levadura.
2. Si el codón no es uno de los empleados comúnmente por las levaduras, reemplazarlo por un codón óptimo para una expresión elevada en células de levadura.
3. Repetir las etapas (1) y (2) para diferentes codones hasta que se consiga el nivel deseado de optimización de codones.
4. Inspeccionar la nueva secuencia codificante con respecto a secuencias indeseadas generadas tales como sitios de enzimas de restricción indeseados, sitios de corte y empalme, promotores, secuencias palindrómicas o repetidas indeseables, secuencias terminadoras de la transcripción y una alta frecuencia de bases GC. Retirar las secuencias indeseadas usando un codón alternativo.

El uso de codones alternativos se define por Lathe J. Molec. Biol., 183: 1-12, 1985. El uso de codones en diferentes huéspedes de levaduras es bien conocido en la técnica. Por ejemplo, Sharp y col., Yeast 7: 657-678, 1991, describe el uso de codones sinónimos en *Saccharomyces cerevisiae*.

La expresión en levadura puede conseguirse usando tanto secuencias optimizadas como secuencias no optimizadas para la expresión en levadura.

En la técnica son bien conocidas técnicas para la producción de genes recombinantes, la introducción en una célula y la expresión de genes recombinantes. Se proporcionan ejemplos de dichas técnicas en referencias tales como Ausubel, Current Protocols in Molecular Biology, John Wiley, 1987-2002, y Sambrook y col., Molecular Cloning, A Laboratory Manual, 2ª Edición, Cold Spring Harbor Laboratory Press, 1989.

EJEMPLOS

A continuación se proporcionan ejemplos para ilustrar adicionalmente diferentes características de la presente invención. Los ejemplos también ilustran la metodología útil para poner en práctica la invención. Estos ejemplos no limitan la invención reivindicada.

Ejemplo 1: Construcción de Vectores

Para analizar la capacidad de regiones de control sintéticas para dirigir la expresión de

genes en *S. cerevisiae*, se retiró la región de control de genes *GAL1-GAL10* de pGAL110, por digestión con *Bam*HI y *Xma*I, para formar un plásmido de 11,3 kpb, pFUNGL. En pFUNGL se insertaron las regiones de control de genes sintéticas y un gen heterólogo. Por ejemplo, se insertaron EE22-XynA (Fig. 2C, SEC ID N°: 23) y BB25-XynA (Fig. 2E, SEC ID N°: 25) en pFUNGL con o sin un gen optimizado que codificaba una proteína heteróloga (HP) para formar pF22MCS-HP, pF25MCS y pF25MCS-HP, respectivamente. El gen heterólogo también se inserta en pGAL110 para formar pGAL110-HP como control.

Ejemplo 2: Expresión en *S. cerevisiae*

Los vectores se usaron para transformar cepas de *S. cerevisiae* que contenían una mutación *leu2* en cepas con prototrofia de leucina (*Leu*⁺) usando un protocolo de transformación de esferoplastos (Hinnen y col., Proc. Natl. Acad. Sci. USA, 75: 1929-33, 1978).

Los transformantes se seleccionaron en medio con agar sintético que carecía de leucina y que contenía sorbitol 1,0 M. El medio de agar sintético superior e inferior que carecía de leucina y que contenía sorbitol 1,0 M se obtuvo en REMEL, Lenexa, KS (n° cat 09459 y 92155, respectivamente). Se obtuvieron aislados *Leu*⁺ clonales por crecimiento seriado en placas con SD sin leucina (KD MEDICAL, Columbia MD).

Para la producción en tubos, se transfirió una alícuota de 0,3 ml del cultivo de siembra a 5,0 ml de medio sin leucina 5X que contenía un 1,6% de glucosa y un 4% de galactosa o medio YEHDG durante 72 horas a 28-30°C hasta una DO₆₀₀ final de 5-16,0/ml. El medio YEHDG contiene, por litro: L-Hy-peptona de soja-Sheffield, 10 g; extracto de Levadura, 20 g; L-dextrosa, 16 g; D (+) galactosa, 40 g. Para la producción en matraces, se transfirió una alícuota de 1,5 ml del cultivo de siembra a 25 ml de medio y se cultivó como se ha descrito anteriormente con agitación a 220 rpm.

Después de la recolección de 10 unidades DO₅₀₀ de células por muestra, los sedimentos celulares se rompieron con perlas de vidrio en 0,3 ml de tampón de lisis (tampón fosfato sódico 0,1 M, pH 7,2, NaCl 0,5 M, PMSF 2 mM). El lisado se recuperó por centrifugación. La concentración de proteínas se determinó por Ensayo Pierce BCA de acuerdo con las instrucciones del fabricante. Los lisados celulares se analizaron con respecto a la expresión del gen heterólogo por análisis de inmunotransferencia después de electroforesis en geles de Tris-Glicina (Invitrogen, Carlsbad, CA) en tampones de Tris-glicina SDS 1X en condiciones reductoras y desnaturizantes. Las muestras contenían proteína celular total. Los geles se transfirieron por transferencia de Western a filtros de membrana de nitrocelulosa de 0,45 micrómetros (Invitrogen). Para estimar el tamaño de las proteínas, se procesaron en paralelo con los lisados patrones teñidos previamente.

Ejemplo 3: Expresión de Genes Heterólogos Dirigida por Regiones de control de Genes

Sintéticas

La proteína heteróloga se expresó en *E. coli* y *S. cerevisiae* y los productos de expresión se compararon.

Con la inducción por galactosa después de la reducción de glucosa, la proteína heteróloga se expresó en todas las cepas de *S. cerevisiae* transformadas ensayadas. La proteína principal producida por *S. cerevisiae* se detectó por análisis de transferencia de Western (500 ng de proteína/calle) con un anticuerpo monoclonal contra la proteína heteróloga (1:5000) y tenía un peso molecular de ~105-110-kDa como se muestra en la Fig. 3 (calles 2, 3 y 4). La proteína de ~105-110-kDa era ligeramente más pequeña que la banda de proteína de mayor tamaño detectada en la muestra de proteína heteróloga marcada con His producida en *E. coli* recombinante purificada (calle 1), ya que se añadió marcador His al peso molecular del control. No se observó señal detectable con un extracto de un transformante de control que contenía el vector pF25MCS solo (calle 5).

El nivel de expresión de genes heterólogos en la cepa transformada con pGAL110-HP es comparable al nivel de expresión obtenido a partir de pF25MCS-HP, pero mayor que el conseguido con pF22MCS-HP. Por lo tanto, la región reguladora de genes EE25-XynA es tan fuerte como la región reguladora de genes *GAL1-GAL10* natural en la dirección de la expresión de un gen heterólogo en *S. cerevisiae*.

Para las células cultivadas en glucosa, no se observaron bandas de proteína en las transferencias de Western, lo que indica que el gen heterólogo no se expresaba o se expresaba a niveles muy bajos en *S. cerevisiae* transformado (Fig. 3, calles 7, 8 y 9). De esta manera, la expresión se controla por las regiones reguladoras de genes sintéticas BB22-XynA y EE2S-XynA y se reprime en la presencia de glucosa. Esto es similar a la represión observada para la región reguladora natural del gen *GAL1-GAL10*.

Otras realizaciones están dentro de las siguientes reivindicaciones. Aunque se han mostrado y descrito varias realizaciones, pueden realizarse diversas modificaciones sin apartarse del espíritu y ámbito de la presente invención.

LISTA DE SECUENCIAS

<110> Merck & Co., Inc.

5

<120> REGIÓN DE CONTROL DE GENES SINTÉTICA

<130> 21624Y

<150> 60/690.321

10

<151> 14-05-2006

<160> 26

<170> FastSEQ para Windows Versión 4.0

15

<210> 1

<211> 318

<212> ADN

<213> *Penicillium purpurogenum*

20

<400> 1

```

ggatattcgt ttagttagcc cctctactat gacattactt ctctctgggat cctataagca 60
aatcaatcgg ggagatgata ctgtaatgaa gaggacccta ggtgactacg atcttgtata 120
agtatcaggg ttgatccctc caaatgatag ccactgatgc tagaacatat tcatagtccc 180
aactagggag gtttactatc ctttcttcaa catcgactct cataatcgat atacttgaaa 240
acccaacaca gaaagaagtt gtagctgaga gtattagcta tatgaacttt tgggttgtgt 300
tcactcaaaa gtgagttt                                     318

```

<210> 2

25

<211> 372

<212> ADN

<213> *Ustilago maydis*

<400> 2

```

ttgacgacac aaccacggag cgggtgttac ttcttctcac tcacgattct aactgctgtg 60
ttggtgcctc gccacaatg aagaagagtg agtgctaaga cgattctcga ttcacgggtc 120
tcgattctcg ccgtcttggg ctttgggaag gctaagagct aagtgccaaag agctaagagc 180
ggcagaacca gaaacccttc gatcagctgc cgagctggac caccgaccac catcggtctt 240
agctacaaca ctagtcgacg gctcgacctg gtggctgggt gtagccagaa tcgatgttgt 300
accgctccaa atcactcata cacgtctaga atctcatggc gaggtttagt gagtatgtgc 360
agatcttaga gt                                     372

```

30

<210> 3

<211> 1404

<212> ADN

<213> *Aspergillus nidulans*

<400> 3

```

aagctttatt tcgcggtttt ttggggtagt catctaataa aacagaccg ttcgaaataa 60
agcgccaaaa aaccccatca gtagattact ttgtctgggc gacgcagcag aggaagcccc 120
gcgatgactc tataaccaccg tacgccgata ctgcgtcgtc tccttcgggg cgctactgag 180
atatgggtggc atgcggctat tatcatcatc gcggcgatgg agaagtgggg ttgactccga 240
agacacttca atagtagtag cgccgctacc tcttcacccc aactgaggct tctgtgaagt 300
aaggagcgac gctgttgatt tgtagacgac gcttgatagg gagaagcatt ttcctcgctg 360

```

```

cgacaactaa acatctgctg cgaactatcc ctcttcgtaa attgtcgtga tgctcgccca 420
acagaggccg actcgctca tccgtcataa taacagcact acgagcgggt tgtctccggc 480
tgagcggagt aggcagtatt cgaacgctgt gtaaagcgga gtggggggga aagtgtggat 540
tgtggagagt gcttgcgaca catttcgcct cccccccct ttcacaccta acacctctca 600
atgcgatagt gttgaggctg atcagacggc gaatcggggc agatatgacc tacgetatca 660
caactccgac tagtctgccg cttagcccg tctatactgg agtttagagg cctcatttga 720
ctataattta cataaattag ataaatagag tcaaactctc ggagtaaact gatattaaat 780
gtattttaatc tatttatctc atgaacgcat gcaataattg cagcaaataat tgatgaagcg 840
agaggtagga tacttgcgta cgttattaac gtcgtttata actacttcgc tctccatcct 900
cgatgaagga ctgtgagcag ttcaaggat cagcagagtc aagggcctga gctacttcct 960
gacactcgtc aagttccata gtcgtctcag ttcccgact tgcaatggcg gtgatccgtg 1020
atcagcgaaac ggaagggcg ctaactctgt acgttaccgc cactaggcac tagtcgcttg 1080
ccttccccgc gattgagaca ttctttacca atgatcgga gctcctgctg gcggacttat 1140
gagtcattca aagaaatggt tactagcctt cgaggacgac cgcctgaata ctcagtaagt 1200
cgaatcattt ctcagttatt tgtggatgcc ctcgttctgt ccacaatttc gcttagtaaa 1260
gagtcataaa acacctacgg gagcaagaca ggtgttaaag tttccgcccc aagtctttta 1320
agttctttta catctatatt cttgcacttc aaaggcgggg ttcagaaaat tcaagaaatt 1380
gtagatataa gaacgtgaag cagt                                     1404

```

<210> 4

<211> 17

<212> ADN

<213> *Saccharomyces cerevisiae*

<400> 4

```

cggattagaa gccgccg 17

```

<210> 5

<211> 17

<212> ADN

<213> *Saccharomyces cerevisiae*

<400> 5

cggtgacag ccctccg 17

5

<210> 6

<211> 17

<212> ADN

<213> *Saccharomyces cerevisiae*

10

<400> 6

aggaagactc tcctccg . 17

15

<210> 7

<211> 17

<212> ADN

<213> *Saccharomyces cerevisiae*

20

<400> 7

cgcgccgcac tgctccg 17

25

<210> 8

<211> 17

<212> ADN

<213> *Saccharomyces cerevisiae*

<400> 8

cggaggactg tcctccg 17

30

<210> 9

<211> 17

<212> ADN

<213> *Saccharomyces cerevisiae*

35

<400> 9

	cgagcactc tcctccg	17
5	<210> 10 <211> 17 <212> ADN <213> Secuencia Artificial	
10	<220> <223> Sitio de Unión a GAL4 Sintético	
	<400> 10 cggatgacac tcctccg	17
15	<210> 11 <211> 17 <212> ADN <213> Secuencia Artificial	
20	<220> <223> Sitio de Unión a GAL4 Sintético	
	<400> 11 cgggccactg tcgtccg	17
25	<210> 12 <211> 17 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Sitio de Unión a GAL4 Sintético	
35	<400> 12 ggtagaggcc atccccg	17

5	<210> 13 <211> 17 <212> ADN <213> <i>Saccharomyces cerevisiae</i>
10	<400> 13 cggaacgactg tggatcg 17 <210> 14 <211> 17 <212> ADN <213> <i>Saccharomyces cerevisiae</i>
15	<400> 14 cgggcggcac tctccg 17
20	<210> 15 <211> 17 <212> ADN <213> Secuencia Artificial
25	<220> <223> Sitio de Unión a GAG4 Sintético <400> 15 aggtcgaggc catccg 17
30	<210> 16 <211> 13 <212> ADN <213> <i>Saccharomyces cerevisiae</i>
35	<400> 16 tattctggg gta 13

5	<210> 17 <211> 13 <212> ADN <213> <i>Saccharomyces cerevisiae</i>
10	<400> 17 ggtttggtggg gcc 13 <210> 18 <211> 13 <212> ADN <213> Secuencia Artificial
15	<220> <223> Sitio de Unión a MIG1 Sintético
20	<400> 18 gcataccggg gcc 13 <210> 19 <211> 13 <212> ADN <213> Secuencia Artificial
25	<220> <223> Sitio de Unión a MIG1 Sintético
30	<400> 19 attatgtggg gta 13 <210> 20 <211> 13 <212> ADN <213> Secuencia Artificial
35	

<220>

<223> Sitio de Unión a MIG1 Sintético

<400> 20

5

attatgtggg gta 13

<210> 21

<211> 13

<212> ADN

10

<213> Secuencia Artificial

<220>

<223> Región Reguladora de Genes Sintética

15

<400> 21

aaaatctggg gaa 13

<210> 22

<211> 872

20

<212> ADN

<213> Secuencia Artificial

<220>

<223> Región Reguladora de Genes Sintética

25

<400> 22

gagacagccc	tgggagcact	ctcctccgag	cgtgagcctc	aggtcgaggc	cactgtcggg	60
agcctcgtga	gaggaggctc	gcactcggag	tccagctccg	gtgatcgcg	tgcacttaca	120
tcaggtgcgc	tcgtgcgtta	ctaatacgga	ggacagcgct	acgtgaatgt	agtccacgcg	180
agcacgcaat	gattacgcct	cctgtgtcct	ccgatctatt	gacgtattta	cgtatactag	240
gagaataaat	cgctacagga	ggctagataa	ctgcataaat	gcatatgac	ctcttattta	300
gcgamggata	gctcgccccg	gtatgcaatc	attctacata	cgttaaaatc	gaactactat	360
cgagcggggc	catacgttag	taagatgtat	gcaattttag	cttgatcgaa	tataggatcg	420
tttttcaatt	tacggatatt	ctgggtggtat	tattatgctt	atatcctagc	aaaaagttaa	480
atgcctataa	gaccaccata	ataatamgm	tatgtggggt	gttaatttct	agggatattc	540
gttttagttag	cccccttact	atacacccca	caattaaaga	tccctataag	caaatcaatc	600
ggggagatga	atgacattac	ttctcctggg	atgggtgacta	cgatcttgta	taagtatcag	660
tactgtaatg	aagaggaccc	taccactgat	gctagaacat	attcatagtc	ggttgatccc	720
tccaaatgat	agctttcttc	aacatcgact	ctcataatcg	ccaactaggg	aggtttacta	780
tcgaaagaag	ttgtagctga	gagtattagc	atatacttga	aaaccaaca	catcactcaa	840
atatatgaac	ttttgggttg	tgtagtgagt	tt			872

<210> 23

<211> 912

<212> ADN

5

<213> Secuencia Artificial

<220>

<223> Región Reguladora de Genes Sintética

10

<400> 23

```

gagagacagc cctcggagca ctctcctccg tccgtgaccg tcagggtcgag gccactgtcg 60
ggagcctcgt gagaggaggc aggcaactggc agtccagctc cgggtgatccc gatgcactta 120
catcaggtcc gctcgtgcct tactaatccg gaggacaggg ctacgtgaat gtagtccagg 180
cgagcacgga atgattaggc ctctgtgtc ctccgatcta ttgacgtatt tacgtatact 240
aggagaataa atcgctacag gaggctagat aactgcataa atgcatatga tcctcttatt 300
tagcgamgga tagctcatca ttctatgcaa tcattcccca gattttgcta tcgaactact 360
atcgagtagt aagatacgtt agtaaggggt ctaaaacgat agcttgatcg aatataggat 420
cgtttaacaa ttacggttt ttctgggtgtt attagcaagc ttatatccta gcaaattgtt 480
aatgccaaa aagaccacca taatcgttmg aatctgggga attaattgtc cacgagcagg 540
atTTTTtGtc aggatattcg ttagaccct taattaacag gtgctcgtcc taaaaaacag 600
tcctataagc mgtttagtta gccctctac tatgacatta cttctcctgg gatggtgact 660
acaaatcaat cggggagatg atactgtaat gaagaggacc ctaccactga tggatcttgt 720
ataagtatca gggttgatcc ctccaaatga tagctttctt cactagaaca tattcatagt 780
cccaactagg gaggtttact atcgaaagaa gtacatcgac tctcataatc gatatacttg 840
aaaacccaac acatcactca aatgtagctg agagtattag ctatatgaac ttttgggttg 900
tgtagtgagt tt                                     . 912

```

<210> 24

<211> 872

<212> ADN

15

<213> Secuencia Artificial

<220>

<223> Región Reguladora de Genes Sintética

20

<400> 24

```

gagacagccc tcggatgaca ctccctccgag cgtgagcctc aggtagaggc cactgtcggg 60
agcctactgt gaggaggctc gcactcggag tccatctccg gtgatcgca tgcacttaca 120
tcagggtgcgc tcgtgcggtta ctaatgcggg ccacagcgct acgtgaatgt agtccacgcg 180
agcacgcaat gattacgccc ggtgtgtcgt ccgatctatt gacgtattta cgtatactag 240
gagaataaat ccgtacagca ggctagataa ctgcataaat gcatatgac ctcttattta 300
ggcamggata gctcgccccg gtatgcaatc attctacata cgttaaaatc gaactactat 360
cgagcgggggc catacgttag taagatgtat gcaatttttag cttgatcgaa tataggatcg 420
tttttcaatt tacggatatt ctggtggtat tattatgctt atatcctagc aaaaagttaa 480
atgcctataa gaccaccata ataatamgmg tatgtggggt gttaatttct agggatattc 540
gtttagttag cccctctact atacacccca caattaaaga tccctataag caaatcaatc 600
ggggagatga atgacattac ttctcctggg atggtgacta cgatcttgta taagtatcag 660
tactgtaatg aagaggaccc taccactgat gctagaacat attcatagtc ggttgatccc 720
tccaaatgat agctttcttc aacatcgact ctcataatcg ccaactaggg aggtttacta 780
tcgaaagaag ttgtagctga gagtattagc atatacttga aaaccaaca catcactcaa 840
atatatgaac ttttggggtg tgtagttagt tt 872

```

<210> 25

<211> 908

<212> ADN

<213> Secuencia Artificial

<220>

<223> Región Reguladora de Genes Sintética

<400> 25

```

gagacagccc tcggatgaca ctccctccgag cgtgagcctc aggtcgaggc cactgtcggg 60
agcctactgt gaggaggctc gcactcggag tccagctccg gtgatcgca tgcacttaca 120
tcagggtgcgc tcgtgcggtta ctaatgcggg ccacagcgct acgtgaatgt agtccacgcg 180
agcacgcaat gattacgccc ggtgtgtcgt ccgatctatt gacgtattta cgtatactag 240
gagaataaat cactacagca ggctagataa ctgcataaat gcatatgac ctcttattta 300
gtgamggata gctcatcatt ctatgcaatc attccccaga ttttgctcaa atgaacctat 360
cgagtagtaa gatacgttag taaggggtct aaaacgagtt tacttgcgat caaattaacg 420
tttaacaatt tacgggtttt ctggtggtat tagcaagcta gtttaattgc aaattgttaa 480
atgccaaaaa gaccaccata atcgttmgaa tctggggaat taattgtcac gagcaggatt 540
ttttgtcagg atattcgttt agacccctta attaacagtg ctgcgcctaa aaaacagtcc 600
tataagcamg ttagtttagc cctctactat gacattactt ctccctgggag ggtgactacg 660
aatcaatcgg ggagatgata ctgtaatgaa gaggacccta ccactgatgc atcttgata 720
agtatcaggg ttgatccctc caaatgatag ctttcttcaa tagaacatat tcatagtccc 780
aactagggag gtttactatc gaaagaagtt catcgactct cataatcgat atacttgaaa 840
acccaacaca tcaactcaaag tagctgagag tattagctat atgaactttt gggttgtgta 900
gtgagttt 908

```

<210> 26

<211> 870

<212> ADN

<213> Secuencia Artificial

<220>

<223> Región Reguladora de Genes Sintética

<400> 26

```

gagacagccc tcggatgaca ctccctccgag cgtgagcctc aggtcgaggc cactgtcggg 60
agcctactgt gaggaggctc gcactcggag tccagctccg gtgatcgcg tgcacttaca 120
tcaggtgcct cgtgcgttac taatgcgggc cacagcgcta cgtgaatgta gtccacggag 180
cacgcaatga ttacgcccgg tgtgtcgtcc gatctattga cgtattttac tatactagga 240
gaataaatcg ctacagcagg ctagataact gcataaatgc atatgaccc cttattttagc 300
gamggatagc tcatcattct atgcaatcat tccccagatt ttgctcaa atgaacctatcg 360
agtagtaaga tacgttagta aggggtctaa aacgagttta cttgcgatga aattaacggt 420
taacaattta cggtttttct ggtggtatta gtaagctact ttaattgcaa attgttaa at 480
gccaaaaaga ccaccataat cattmgmgaa tctggggaat taatttctag ggatattcgt 540
ttagttagcc cctctacttt agacccttta attaaagatc cctataagca aatcaatcgg 600
ggagatgaat gacattactt ctccctgggat ggtgactacg atcttgata agtatcagta 660
ctgtaatgaa gaggacccta ccactgatgc tagaacatat tcatagtcgg ttgatccctc 720
caaatgatag ctttcttcaa catcgactct cataatcgcc aactagggag gtttactatc 780
gaaagaagtt gtagctgaga gtattagcat atacttgaaa acccaacaca tcactcaa at 840
atatgaactt ttgggttggt tagtgagttt                                     870

```

REIVINDICACIONES

1. Una región de control de genes sintética que comprende una secuencia reguladora de genes sintética de una cepa de levadura y un promotor de una cepa de hongo filamentoso localizado cadena abajo de la secuencia reguladora de genes sintética; en la que el promotor puede reconocerse por los factores de transcripción generales y la ARN polimerasa de la cepa de levadura;

en la que la secuencia reguladora de genes sintética es capaz de regular la transcripción iniciada por un promotor de hongo filamentoso en la cepa de levadura.

2. La región de control de la reivindicación 1, en la que la secuencia reguladora de genes sintética comprende un sitio de unión para una proteína reguladora de genes, sin que dicho sitio de unión sea una secuencia en las regiones reguladoras de genes naturales de la cepa de levadura.

3. La región de control de la reivindicación 1 ó 2, en la que la cepa de levadura es *Saccharomyces cerevisiae*, *Pichia pastoris*, *Hansenula polymorpha* o *Yarrowia lipolytica*.

4. La región de control de la reivindicación 1, 2 ó 3, en la que la cepa de hongo filamentoso se selecciona entre el grupo que consiste en *Ustilago maydis*, *Aspergillus nidulans*, y *Penicillium purpurogenum*.

5. La región de control de la reivindicación 4, en la que el promotor se selecciona entre el grupo que consiste en promotores del gen XynA de *Penicillium purpurogenum*, el gen de *DPMJ* de *Ustilago maydis* y el gen *ArgB* de *Aspergillus nidulans*.

6. La región de control de la reivindicación 5, en la que el promotor es la SEC ID N°: 1.

7. La región de control de cualquier reivindicación anterior, en la que la secuencia reguladora de genes comprende un sitio de unión para un activador de genes de la cepa de levadura.

8. La región de control de la reivindicación 7, en la que el sitio de unión es un sitio de unión para la proteína GAL4 de *S. cerevisiae*.

9. La región de control de la reivindicación 8, en la que el sitio de unión comprende una secuencia de ADN seleccionada entre el grupo que consiste en las SEC ID N°: 10, 11, 12, y 15.
- 5 10. La región de control de una cualquiera de las reivindicaciones 1-6, en la que la secuencia reguladora de genes comprende un sitio de unión para un represor de genes de la cepa de levadura.
- 10 11. La región de control de la reivindicación 7, en la que la secuencia reguladora de genes comprende además un sitio de unión para un represor de genes.
12. La región de control de la reivindicación 11, en la que la secuencia reguladora de genes comprende además un sitio de unión para la proteína MIG1 de *S. cerevisiae*.
- 15 13. La región de control de la reivindicación 12, en la que el sitio de unión para la proteína MIG1 comprende una secuencia seleccionada del grupo que consiste en las SEC ID N°: 18, 19, y 20.
- 20 14. La región de control de la reivindicación 1, en la que la región de control de genes sintética comprende una secuencia seleccionada entre el grupo que consiste en las SEC ID N°: 21, 22, 23, 24, 25, y 26.
- 25 15. Un vector de expresión de ADN que comprende una región de control de genes sintética de una cualquiera de las reivindicaciones 1-14; una secuencia codificante que codifica una proteína, polipéptido o péptido bajo el control de la región de control, y un marcador de selección de levadura.
- 30 16. El vector de expresión de ADN de la reivindicación 15, que comprende además una secuencia señal de poliadenilación localizada cadena abajo de la secuencia codificante.
17. El vector de expresión de ADN de la reivindicación 15, que comprende además un terminador de la transcripción localizado cadena abajo de la secuencia codificante.
- 35 18. El vector de expresión de ADN de la reivindicación 15, 16 ó 17, en el que el marcador de selección de levadura es un marcador de selección de *S. cerevisiae*.

19. El vector de expresión de ADN de la reivindicación 18, en el que el marcador de selección de *S. cerevisiae* está seleccionado entre el grupo que consiste en *LEU2*, *TRP1*, *URA3* y *HIS3*.
- 5
20. El vector de expresión de ADN de la reivindicación 1, que comprende además un origen de replicación de *S. cerevisiae*.
21. El vector de expresión de ADN de la reivindicación 20, en el que el origen de replicación de *S. cerevisiae* se basa en la secuencia de ADN de 2 micrómetros de *S. cerevisiae*.
- 10
22. El vector de expresión de ADN de la reivindicación 15, que comprende además un origen de replicación bacteriano.
- 15
23. Una cepa de levadura que comprende un vector de expresión de ADN de una cualquiera de las reivindicaciones 15-22.
24. La cepa de levadura de la reivindicación 23, en la que la cepa de levadura está seleccionada entre el grupo que consiste en *Saccharomyces cerevisiae*, *Pichia pastoris*, *Hansenula polymorpha* y *Yarrowia lipolytica*.
- 20
25. La cepa de levadura de la reivindicación 23, en la que la cepa de levadura es *Saccharomyces cerevisiae*.
- 25
26. Un procedimiento para producir una proteína, polipéptido o péptido recombinante que comprende expresar la secuencia codificante de la cepa de levadura de la reivindicación 23.

```
1  GGATATTCGT  TTAGTTAGCC CCTCTACTAT GACATTACTT CTCCTGGGAT
   CCTATAAGCA AATCAATCGG GGAGATGATA CTGTAATGAA GAGGACCCTA
51  GGTGACTACG ATCTTGATA AGTATCAGGG TTGATCCCTC CAAATGATAG
   CCACTGATGC TAGAACATAT TCATAGTCCC AACTAGGGAG GTTTACTATC
101 CTTTCTTCAA CATCGACTCT CATAATCGAT ATACTTGAAA ACCCAACACA
   GAAAGAAGTT GTAGCTGAGA GTATTAGCTA TATGAACTTT TGGGTTGTGT
151 TCACTCAAA
   AGTGAGTTT
```

Figura 1A

```
1  TTGACGACAC AACCACGGAG CGGGTGTTAC TTCTTCTCAC TCACGATTCT
   AACTGCTGTG TTGGTGCCTC GCCCACAATG AAGAAGAGTG AGTGCTAAGA
51  CGATTCTCGA TTCACGGTTC TCGATTCTCG CCGTCTTGGT CTTTGGGAAG
   GCTAAGAGCT AAGTGCCAAG AGCTAAGAGC GGCAGAACCA GAAACCCTTC
101 GATCAGCTGC CGAGCTGGAC CACCGACCAC CATCGGTCTT AGCTACAACA
   CTAGTCGACG GCTCGACCTG GTGGCTGGTG GTAGCCAGAA TCGATGTTGT
151 ACCGCTCCAA ATCACTCATA CACGTCTAGA ATCTCA
   TGGCGAGGTT TAGTGAGTAT GTGCAGATCT TAGAGT
```

Figura 1B

```

1  AAGCTTTATT TCGCGGTTTT TTGGGGTAGT CATCTAATGA AACAGACCCG
   TTCGAAATAA AGCGCCAAAA AACCCCATCA GTAGATTACT TTGTCTGGGC
51  GACGCAGCAG AGGAAGCCCC GCGATGACTC TATACCACCG TACGCCGATA
   CTGCGTCGTC TCCTTCGGGG CGCTACTGAG ATATGGTGGC ATGCGGCTAT
101 TATCATCATC GCGGCGATGG AGAAGTGGGG TTGACTCCGA AGACACTTCA
   ATAGTAGTAG CGCCGCTACC TCTTCACCCC AACTGAGGCT TCTGTGAAGT
151 AAGGAGCGAC GCTGTTGATT TGTAGACGAC GCTTGATAGG GAGAAGCATT
   TTCCTCGCTG CGACAATAA ACATCTGCTG CGAACTATCC CTCTTCGTAA
201 ATTGTCGTGA TGCTCGCCCA ACAGAGGCCG ACTCGCCTCA TCCGTCATAA
   TAACAGCACT ACGAGCGGGT TGTCTCCGGC TGAGCGGAGT AGGCAGTATT
251 CGAACGCTGT GTAAAGCGGA GTGGGGGGGA AAGTGTGGAT TGTGGAGAGT
   GCTTGCGACA CATTTGCGCT CACCCCCCCT TTCACACCTA ACACCTCTCA
301 ATGCGATAGT GTTGAGGCTG ATCAGACGGC GAATCGGGCC AGATATGACC
   TACGCTATCA CAACTCCGAC TAGTCTGCCG CTTAGCCCGG TCTATACTGG
351 AGTTTAGAGG CCTCATTTGA CTATAATTTA CATAAATTAG ATAAATAGAG
   TCAAATCTCC GGAGTAAACT GATATTAAAT GTATTTAATC TATTTATCTC
401 ATGAACGCAT GCAATAATTG CAGCAAATAT TGATGAAGCG AGAGGTAGGA
   TACTTGCGTA CGTTATTAAC GTCGTTTATA ACTACTTCGC TCTCCATCCT
451 CGATGAAGGA CTGTGAGCAG TTCAAGGTAT CAGCAGAGTC AAGGGCCTGA
   GCTACTTCCT GACACTCGTC AAGTTCCATA GTCGTCTCAG TTCCCGGACT
501 TGCAATGGCG GTGATCCGTG ATCAGCGAAC GGAAGGGGCG CTAActCTGT
   ACGTTACCGC CACTAGGCAC TAGTCGCTTG CCTTCCCCGC GATTGAGACA
551 TTCTTTACCA ATGATCGGAA GCTCCTGCTG GCGGACTTAT GAGTCATTCA
   AAGAAATGGT TACTAGCCTT CGAGGACGAC CGCCTGAATA CTCAGTAAGT
601 CGAATCATTT CTCAGTTATT TGTGGATGCC CTCGTTCTGT CCACAATTTT
   GCTTAGTAAA GAGTCAATAA ACACCTACGG GAGCAAGACA GGTGTTAAAG
651 TTTCCGCCCC AAGTCTTTTA AGTTCTTTAA CATCTATATT CTTGCACTTC
   AAAGGCGGGG TTCAGAAAAT TCAAGAAATT GTAGATATAA GAACGTGAAG
701 CA
   GT

```

Figura 1C

	GAL4			GAL4	
1	<u>GACAGCCCTC</u>	<u>GGAGCACTCT</u>	<u>CCTCCGTCCG</u>	<u>TGACCGTCAG</u>	<u>GTCGAGGCCA</u>
	<u>CTGTCGGGAG</u>	<u>CCTCGTGAGA</u>	<u>GGAGGCAGGC</u>	<u>ACTGGCAGTC</u>	<u>CAGCTCCGGT</u>
	GAL4				
51	<u>TCCCGATGCA</u>	<u>CTTACATCAG</u>	<u>GTCCGCTCGT</u>	<u>GCCTTACTAA</u>	<u>TCCGGAGGAC</u>
	<u>AGGGCTACGT</u>	<u>GAATGTAGTC</u>	<u>CAGGCGAGCA</u>	<u>CGGAATGATT</u>	<u>AGGCCTCCTG</u>
101	<u>TGTCCTCCGA</u>	<u>TCTATTGACG</u>	<u>TATTTACGTA</u>	<u>TACTAGGAGA</u>	<u>ATAAATCCGT</u>
	<u>ACAGGAGGCT</u>	<u>AGATAACTGC</u>	<u>ATAAATGCAT</u>	<u>ATGATCCTCT</u>	<u>TATTTAGGCA</u>
	MIG1				
151	<u>GATAGCTCGC</u>	<u>CCCGGTATGC</u>	<u>AATCATTCTA</u>	<u>CATACGTTAA</u>	<u>AATCGAACTA</u>
	<u>CTATCGAGCG</u>	<u>GGGCCATACG</u>	<u>TTAGTAAGAT</u>	<u>GTATGCAATT</u>	<u>TTAGCTTGAT</u>
201	<u>CGAATATAGG</u>	<u>ATCGTTTTTC</u>	<u>AATTTACGGA</u>	<u>TATTCTGGTG</u>	<u>GTATTATTAT</u>
	<u>GCTTATATCC</u>	<u>TAGCAAAAAG</u>	<u>TTAAATGCCT</u>	<u>ATAAGACCAC</u>	<u>CATAATAATA</u>
	MIG1				
251	<u>TATGTGGGGT</u>	<u>GTTAATTTCT</u>	<u>AGGGATATTC</u>	<u>GTTTAGTTAG</u>	<u>CCCCTCTACT</u>
	<u>ATACACCCCA</u>	<u>CAATTAAAGA</u>	<u>TCCCTATAAG</u>	<u>CAAATCAATC</u>	<u>GGGGAGATGA</u>
301	<u>ATGACATTAC</u>	<u>TTCTCCTGGG</u>	<u>ATGGTGACTA</u>	<u>CGATCTTGTA</u>	<u>TAAGTATCAG</u>
	<u>TACTGTAATG</u>	<u>AAGAGGACCC</u>	<u>TACCACTGAT</u>	<u>GCTAGAACAT</u>	<u>ATTCATAGTC</u>
351	<u>GGTTGATCCC</u>	<u>TCCAAATGAT</u>	<u>AGCTTTCTTC</u>	<u>AACATCGACT</u>	<u>CTCATAATCG</u>
	<u>CCAAC TAGGG</u>	<u>AGGTTTACTA</u>	<u>TCGAAAGAAG</u>	<u>TTGTAGCTGA</u>	<u>GAGTATTAGC</u>
401	<u>ATATACTTGA</u>	<u>AAACCCAACA</u>	<u>CATCACTCAA</u>	A	
	<u>TATATGAACT</u>	<u>TTTGGGTTGT</u>	<u>GTAGTGAGTT</u>	T	

Figura 2A

GAL4

1 GACAGCCCTC GGAGCACTCT CCTCCGAGCG TGAGCCTCAG GTCGAGGCCA
 CTGTCGGGAG CCTCGTGAGA GGAGGCTCGC ACTCGGAGTC CAGCTCCGGT

GAL4

51 TCGCGATGCA CTTACATCAG GTGCGCTCGT GCGTTACTAA TGCGGAGGAC
 AGCGCTACGT GAATGTAGTC CACGCGAGCA CGCAATGATT ACGCCTCCTG

101 TGTCCTCCGA TCTATTGACG TATTTACGTA TACTAGGAGA ATAAATCGCT
ACAGGAGGCT AGATAACTGC ATAAATGCAT ATGATCCTCT TATTTAGCGA

MIG1

151 GATAGCTCGC CCCGGTATGC AATCATTTCTA CATACGTTAA AATCGAACTA
 CTATCGAGCG GGGCCATACG TTAGTAAGAT GTATGCAATT TTAGCTTGAT

201 CGAATATAGG ATCGTTTTTC AATTTACGGA TATTCTGGTG GTATTATTAT
 GCTTATATCC TAGCAAAAAG TTAAATGCCT ATAAGACCAC CATAATAATA

MIG1

251 TATGTGGGGT GTTAATTTCT AGGGATATTC GTTTAGTTAG CCCCTCTACT
ATACACCCCA CAATTAAAGA TCCCTATAAG CAAATCAATC GGGGAGATGA

301 ATGACATTAC TTCTCCTGGG ATGGTGAATA CGATCTTGTA TAAGTATCAG
 TACTGTAATG AAGAGGACCC TACCACTGAT GCTAGAACAT ATTCATAGTC

351 GGTTGATCCC TCCAAATGAT AGCTTTCTTC AACATCGACT CTCATAATCG
 CCAACTAGGG AGGTTTACTA TCGAAAGAAG TTGTAGCTGA GAGTATTAGC

401 ATATACTTGA AAACCCAACA CATCACTCAA A
 TATATGAACT TTTGGGTTGT GTAGTGAGTT T

Figura 2B

	GAL4			GAL4	
1	<u>GACAGCCCTC</u>	<u>GGAGCACTCT</u>	<u>CCTCCGTCCG</u>	<u>TGACCGTCAG</u>	<u>GTCGAGGCCA</u>
	<u>CTGTCGGGAG</u>	<u>CCTCGTGAGA</u>	<u>GGAGGCAGGC</u>	<u>ACTGGCAGTC</u>	<u>CAGCTCCGGT</u>
				GAL4	
51	<u>TCCCGATGCA</u>	<u>CTTACATCAG</u>	<u>GTCCGCTCGT</u>	<u>GCCTTACTAA</u>	<u>TCCGGAGGAC</u>
	<u>AGGGCTACGT</u>	<u>GAATGTAGTC</u>	<u>CAGGCGAGCA</u>	<u>CGGAATGATT</u>	<u>AGGCCTCCTG</u>
101	<u>TGTCCTCCGA</u>	<u>TCTATTGACG</u>	<u>TATTTACGTA</u>	<u>TACTAGGAGA</u>	<u>ATAAATCGCT</u>
	<u>ACAGGAGGCT</u>	<u>AGATAACTGC</u>	<u>ATAAATGCAT</u>	<u>ATGATCCTCT</u>	<u>TATTTAGCGA</u>
				MIG1	
151	<u>GATAGCTCAT</u>	<u>CATTCTATGC</u>	<u>AATCATTCCC</u>	<u>CAGATTTTGC</u>	<u>TATCGAACTA</u>
	<u>CTATCGAGTA</u>	<u>GTAAGATACG</u>	<u>TTAGTAAGGG</u>	<u>GTCTAAAACG</u>	<u>ATAGCTTGAT</u>
201	<u>CGAATATAGG</u>	<u>ATCGTTTAAC</u>	<u>AATTTACGGT</u>	<u>TTTTCTGGTG</u>	<u>GTATTAGCAA</u>
	<u>GCTTATATCC</u>	<u>TAGCAAATTG</u>	<u>TTAAATGCCA</u>	<u>AAAAGACCAC</u>	<u>CATAATCGTT</u>
	MIG1				
251	<u>AATCTGGGGA</u>	<u>ATTAATTGTC</u>	<u>CACGAGCAGG</u>	<u>ATTTTTTGTC</u>	<u>AGGATATTCG</u>
	<u>TTAGACCCCT</u>	<u>TAATTAACAG</u>	<u>GTGCTCGTCC</u>	<u>TAAAAAACAG</u>	<u>TCCTATAAGC</u>
	MIG1				
301	<u>TTTAGTTAGC</u>	<u>CCCTCTACTA</u>	<u>TGACATTACT</u>	<u>TCTCCTGGGA</u>	<u>TGGTGACTAC</u>
	<u>AAATCAATCG</u>	<u>GGGAGATGAT</u>	<u>ACTGTAATGA</u>	<u>AGAGGACCCT</u>	<u>ACCACTGATG</u>
351	<u>GATCTTGAT</u>	<u>AAGTATCAGG</u>	<u>GTTGATCCCT</u>	<u>CCAAATGATA</u>	<u>GCTTTCCTCA</u>
	<u>CTAGAACATA</u>	<u>TTCATAGTCC</u>	<u>CAACTAGGGA</u>	<u>GGTTTACTAT</u>	<u>CGAAAGAAGT</u>
401	<u>ACATCGACTC</u>	<u>TCATAATCGA</u>	<u>TATACTTGAA</u>	<u>AACCCAACAC</u>	<u>ATCACTCAAA</u>
	<u>TGTAGCTGAG</u>	<u>AGTATTAGCT</u>	<u>ATATGAACTT</u>	<u>TTGGGTTGTG</u>	<u>TAGTGAGTTT</u>

Figura 2C

GAL4

1 GACAGCCCTC GGATGACACT CCTCCGAGCG TGAGCCTCAG GTAGAGGCCA
CTGTCTGGGAG CCTACTGTGA GGAGGCTCGC ACTCGGAGTC CATCTCCGGT

GAL4

51 TCGCGATGCA CTTACATCAG GTGCGCTCGT GCGTTACTAA TCGGGGCCAC
AGCGCTACGT GAATGTAGTC CACGCGAGCA CGCAATGATT ACGCCCCGGTG

101 TGTCGTCCGA TCTATTGACG TATTTACGTA TACTAGGAGA ATAAATCCGT
ACAGCAGGCT AGATAACTGC ATAAATGCAT ATGATCCTCT TATTTAGGCA

MIG1

151 GATAGCTCGC CCCGGTATGC AATCATTCCTA CATACGTAA AATCGAATA
CTATCGAGCG GGGCCATACG TTAGTAAGAT GTATGCAATT TTAGCTTGAT

201 CGAATATAGG ATCGTTTTTC AATTTACGGA TATTCTGGTG GTATTATTAT
GCTTATATCC TAGCAAAAAG TTAAATGCCT ATAAGACCAC CATAATAATA

MIG1

251 TATGTGGGGT GTTAATTTCT AGGGATATTC GTTTAGTTAG CCCCTCTACT
ATACACCCCA CAATTAAAGA TCCCTATAAG CAAATCAATC GGGGAGATGA

301 ATGACATTAC TTCTCCTGGG ATGGTGACTA CGATCTTGTA TAAGTATCAG
TACTGTAATG AAGAGGACCC TACCACTGAT GCTAGAACAT ATTCATAGTC

351 GGTGATCCC TCCAAATGAT AGCTTTCTTC AACATCGACT CTCATAATCG
CCAAC TAGGG AGGTTTACTA TCGAAAGAAG TTGTAGCTGA GAGTATTAGC

401 ATATACTTGA AAACCCAACA CATCACTCAA A
TATATGAACT TTTGGGTGT GTAGTGAGTT T

Figura 2D

GAL4

1 GACAGCCCTC GGATGACACT CCTCCGAGCG TGAGCCTCAG GTCGAGGCCA
 CTGTCGGGAG CCTACTGTGA GGAGGCTCGC ACTCGGAGTC CAGCTCCGGT

GAL4

51 TCGCGATGCA CTTACATCAG GTGCGCTCGT GCGTTACTAA TGCGGGCCAC
 AGCGCTACGT GAATGTAGTC CACGCGAGCA CGCAATGATT ACGCCCCGGT
 101 TGTCGTCCGA TCTATTGACG TATTTACGTA TACTAGGAGA ATAAATCACT
 ACAGCAGGCT AGATAACTGC ATAAATGCAT ATGATCCTCT TATTTAGTGA

MIG1

151 GATAGCTCAT CATTCATATGC AATCATTCCC CAGATTTTGC TCAAATGAAC
 CTATCGAGTA GTAAGATACG TTAGTAAGGG GTCTAAAACG AGTTTACTTG
 201 CGATCAAATT AACGTTTAAC AATTTACGGT TTTTCTGGTG GTATTAGCAA
 GCTAGTTTAA TTGCAAATTG TTAAATGCCA AAAAGACCAC CATAATCGTT

MIG1

251 AATCTGGGGA ATTAATTGTC ACGAGCAGGA TTTTTTGTC GGATATTCGT
 TTAGACCCCT TAATTAACAG TGCTCGTCCT AAAAAACAGT CCTATAAGCA

MIG1

301 TTAGTTAGCC CCTCTACTAT GACATTACTT CTCCTGGGAT GGTGACTACG
 AATCAATCGG GGAGATGATA CTGTAATGAA GAGGACCCTA CCACTGATGC
 351 ATCTTGATATA AGTATCAGGG TTGATCCCTC CAAATGATAG CTTTCTTCAA
 TAGAACATAT TCATAGTCCC AACTAGGGAG GTTTACTATC GAAAGAAGTT
 401 CATCGACTCT CATAATCGAT ATACTTGAAA ACCCAACACA TCACTCAA
 GTAGCTGAGA GTATTAGCTA TATGAACTTT TGGGTTGTGT AGTGAGTTT

Figura 2E

GAL4

1 GACAGCCCTC GGATGACACT CCTCCGAGCG TGAGCCTCAG GTCGAGGCCA
 CTGTCGGGAG CCTACTGTGA GGAGGCTCGC ACTCGGAGTC CAGCTCCGGT

GAL4

51 TCGCGATGCA CTTACATCAG GTGCCTCGT GCGTTACTAA TGCGGGCCAC
 AGCGCTACGT GAATGTAGTC CACGGAGCA CGCAATGATT ACGCCCGGTG

100 TGTCGTCCGA TCTATTGACG TATTTACGTA TACTAGGAGA ATAAATCGCT
 ACAGCAGGCT AGATAACTGC ATAAATGCAT ATGATCCTCT TATTTAGCGA

MIG1

150 GATAGCTCAT CATTCTATGC AATCATTTCC CAGATTTTGC TCAAATGAAC
 CTATCGAGTA GTAAGATACG TTAGTAAGGG GTCTAAAACG AGTTTACTTG

200 CGATGAAATT AACGTTTAAC AATTTACGGT TTTTCTGGTG GTATTAGTAA
 GCTACTTTAA TTGCAAATTG TTAAATGCCA AAAAGACCAC CATAATCATT

MIG1

250 AATCTGGGGA ATTAATTTCT AGGGATATTC GTTTAGTTAG CCCCTCTACT
 TTAGACCCCT TAATTAAAGA TCCCTATAAG CAAATCAATC GGGGAGATGA

300 ATGACATTAC TTCTCCTGGG ATGGTGACTA CGATCTTGTA TAAGTATCAG
 TACTGTAATG AAGAGGACCC TACCACTGAT GCTAGAACAT ATTCATAGTC

350 GGTTGATCCC TCCAAATGAT AGCTTTCTTC AACATCGACT CTCATAATCG
 CCAACTAGGG AGGTTTACTA TCGAAAGAAG TTGTAGCTGA GAGTATTAGC

400 ATATACTTGA AAACCCAACA CATCACTCAA A
 TATATGAACT TTTGGGTTGT GTAGTGAGTT T

Figura 2F

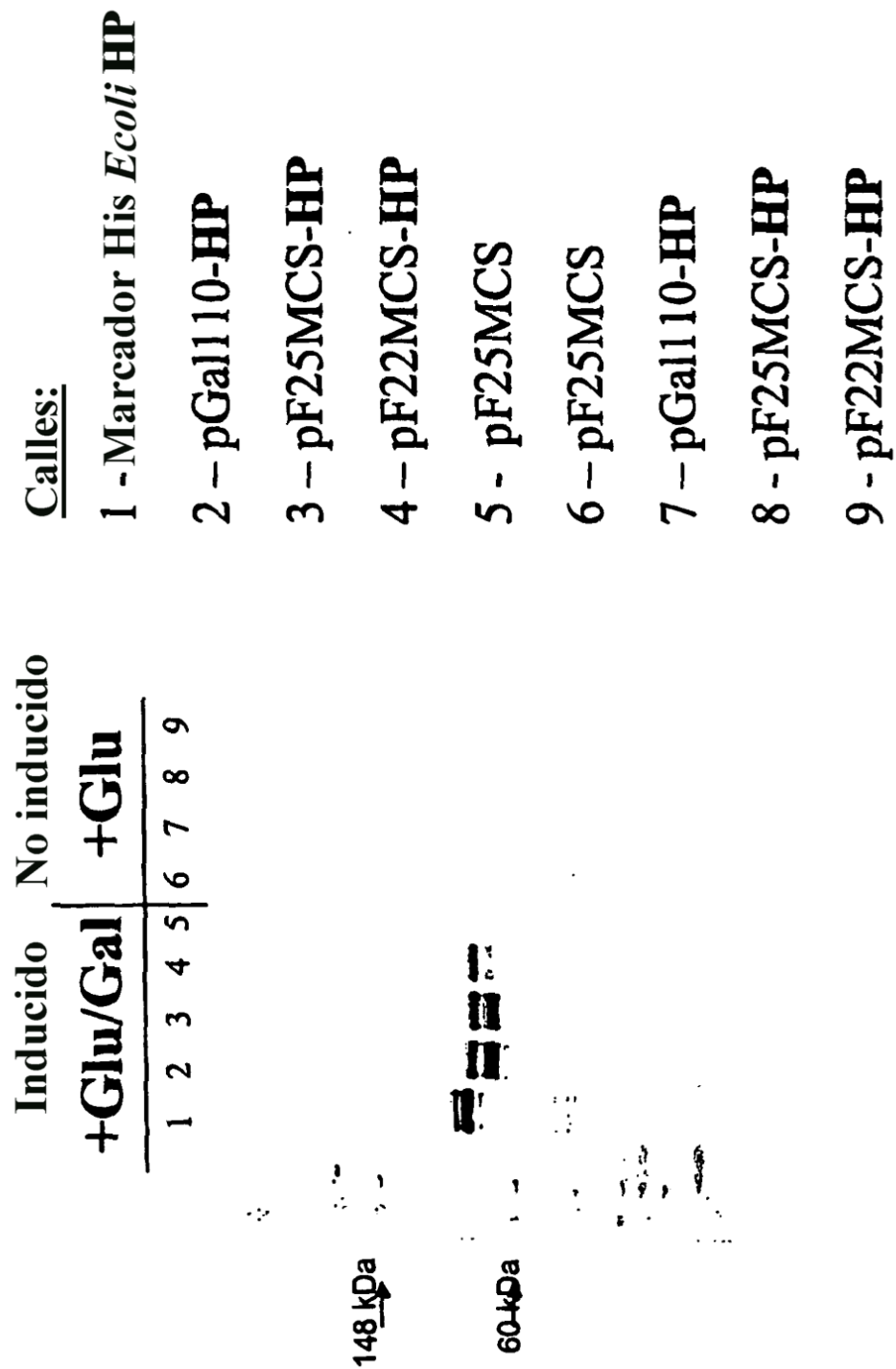


Figura 3