

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 381 384**

51 Int. Cl.:
C12N 15/63 (2006.01)
C07H 21/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **05806571 .5**
96 Fecha de presentación: **31.08.2005**
97 Número de publicación de la solicitud: **1791958**
97 Fecha de publicación de la solicitud: **06.06.2007**

54 Título: **Sistema bihíbrido de levaduras/bacterias y procedimientos para su utilización**

30 Prioridad:
31.08.2004 US 606266 P

45 Fecha de publicación de la mención BOPI:
25.05.2012

45 Fecha de la publicación del folleto de la patente:
25.05.2012

73 Titular/es:
**FOX CHASE CANCER CENTER
333 COTTMAN AVENUE
PHILADELPHIA, PA 19111-2497, US y
THE GENERAL HOSPITAL CORPORATION**

72 Inventor/es:
**SEREBRIISKII, Ilya;
GOLEMIS, Erica, A. y
JOUNG, Keith, J.**

74 Agente/Representante:
Ponti Sales, Adelaida

ES 2 381 384 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

Sistema bihíbrido de levaduras/bacterias y procedimientos para su utilización

5 CAMPO DE LA INVENCION

[0001] Esta invención está relacionada con el campo de la biología molecular. Más específicamente, la invención proporciona nuevas composiciones y procedimientos para facilitar el aislamiento y la caracterización de nuevas interacciones proteína-proteína involucradas en la regulación del crecimiento celular y el metabolismo.

10

CAMPO DE INVENCION

[0002] A lo largo de esta solicitud se citan varias publicaciones y documentos de patente para definir mejor el estado de la técnica a la que pertenece esta invención. Cada una de las siguientes citas se incorpora en el presente documento por referencia.

15

[0003] Los sistemas bihíbridos de levaduras (Chien y col. 1991; Fields y Song 1989; Gyuris y col. 1993; Vojtek y col. 1993) son herramientas habituales utilizadas para identificar nuevas interacciones proteína-proteína y para realizar análisis de estructura-función sobre interacciones proteína-proteína previamente definidas. Dichos sistemas son efectivos con una parte sustancial de proteínas eucariotas y han desempeñado un papel importante en análisis proteómicos de alto rendimiento dirigidos a establecer grupos de proteínas que interactúan (por ejemplo, Giot y col. 2003; Ito y col. 2000; Li y col. 2004; Uetz y col., 2000). Con el fin de incrementar la versatilidad de un enfoque bihíbrido para identificar y analizar interacciones de proteínas en aplicaciones de alto rendimiento, uno de los enfoques ha sido trasladar los componentes básicos del sistema bihíbrido de levaduras a un organismo hospedador bacteriano (Dove y col. 1997; Joung y col. 2000). Hasta la fecha, no se ha comparado directamente la eficacia relativa de la detección de las interacciones entre proteínas en bacterias y levaduras. No obstante, existen una serie de razones para anticipar que es posible que se observen diferencias. Puesto que las levaduras son eucariotas, es más probable que las proteínas eucariotas utilizadas como "cebos" en selecciones de bihíbridos se plieguen y se modifiquen post-traduccionalmente de manera apropiada en levaduras que en bacterias, incrementando así sus oportunidades de identificar pares fisiológicos. No obstante, ciertas proteínas pueden ser problemáticas como cebos en sistemas bihíbridos de levaduras; por ejemplo, proteínas que normalmente se excluyen del núcleo en eucariotas, que se pueden secuestrar mediante la interacción con un par abundante evolutivamente conservado en levaduras, o que estimulan la transcripción en levaduras (es decir, que "se autoactivan"). Se espera que todos estos posibles problemas sean menos problemáticos en el sistema bihíbrido de bacterias. Para maximizar las oportunidades de obtener todos los interactores relevantes para una proteína de interés, sería deseable tener la capacidad de someter a ensayo rápidamente un cebo dado tanto en entornos de levaduras como de bacterias.

RESUMEN DE LA INVENCION

[0004] Según la presente invención, se proporcionan plásmidos y cepas adecuados para su uso tanto en sistemas de interacción entre proteínas de levaduras como de bacterias. Se describen una nueva serie de vectores en los que un solo plásmido que contiene un promotor modificado dirige la expresión eficiente de una proteína cebo en levaduras o en bacterias, permitiendo así el estudio en paralelo en ambos organismos. Además, se proporcionan cepas de soporte informadoras de levaduras y bacterias optimizadas.

45

[0005] Así, en un aspecto de la invención, se proporciona un ácido nucleico aislado que comprende una secuencia promotora procedente del nucleótido en posición 4 al nucleótido en posición 545 de la Figura 1C que dirige la expresión de una secuencia codificante unida de manera operable tanto en levaduras como en bacterias. También se proporcionan plásmidos que comprenden este promotor seleccionado del grupo constituido por (SEQ ID NO: 2) (pGLS20), (SEQ ID NO: 3) (pGLS22) y pGLS23 en el que pGLS23 es idéntico a pGLS22 excepto por un cambio en la posición 3267 de gaattC a gaattA y pBR-AMP- α LPL (SEQ ID NO: 5). También se describen células hospedadoras que comprenden estos plásmidos. Dichas células hospedadoras son preferentemente células de *E. coli* y *S. cerevisiae*.

[0006] En el presente documento también se describen nuevas cepas de *E. coli*. Se proporcionan *E. coli* KJI 567 y *E. coli* AG58A (RP28). También se describen nuevas cepas de levadura, por ejemplo, *S. cerevisiae* PRT50 y *S. cerevisiae* PRT475.

[0007] En una forma de realización preferida de la invención, se proporciona un procedimiento para comparar las interacciones de unión entre una primera proteína y una segunda proteína tanto en entornos de organismos bacterianos como de levaduras utilizando una construcción que funciona en ambos organismos, que comprende:

- a) el suministro de células hospedadoras de levaduras y bacterianas, comprendiendo cada una,
- i) un gen informador unido de manera operable a una secuencia de ADN que comprende un sitio de unión a proteínas;
 - ii) un primer gen de fusión que expresa una primera proteína de fusión, dicha primera proteína de

65

fusión comprendiendo dicha primera proteína unida covalentemente a un resto de unión que es capaz de unirse específicamente a dicho sitio de unión a proteínas que está dirigido por un elemento promotor según la reivindicación 1 y

5 iii) un segundo gen de fusión que expresa una segunda proteína de fusión, dicha segunda proteína de fusión comprendiendo dicha segunda proteína unida covalentemente al resto de activación génico;

b) permitir que dichas primera y segunda proteínas interactúen; y

10 c) medir la expresión de dicho gen informador como medida de dicha interacción entre dichas primera y segunda proteínas tanto en *E. coli* como en levaduras.

[0008] En la presente invención también se proporcionan kits para poner en práctica el procedimiento descrito anteriormente. Un kit de ejemplo comprende:

15 a) un plásmido seleccionado del grupo constituido por la SEQ ID NO: 2 (pGLS20), SEQ ID NO: 3 (pGLS22) y pGLS23 en el que pGLS23 es idéntico a pGLS22 excepto por un cambio en la posición 3667 de gaattC a gaatta;

b) al menos uno de pAC-AW- α LPL, y/o pBR-AMP- α LPL;

20 c) una cepa de *E. coli* seleccionada del grupo constituido por *E. coli* KJI 567 o *E. coli* AG58A (RP28); y

d) una cepa de levaduras seleccionada del grupo constituido por PRT50, y PRT475.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

25

[0009] Figura 1. Esquema de los sistemas bihíbridos de levaduras y bacterias. A. En el sistema bihíbrido de levaduras mostrado, una proteína híbrida dimérica λ cl-cebo interactúa con un dominio de activación (AD)-proteína híbrida presa estimulando así la transcripción a partir de un promotor adyacente que dirige la expresión de un gen informador GusA cuantitativo o LYS2 seleccionable. **B.** En el sistema bihíbrido bacteriano mostrado, una proteína híbrida dimérica λ cl-cebo interactúa con una subunidad α de la ARN polimerasa de *E. coli* (ARNP)-proteína híbrida presa, enganchando así la ARNP a un promotor adyacente que dirige la expresión de un gen informador lacZ cuantitativo o HIS3 seleccionable. Nótese que ambos sistemas utilizan una proteína híbrida λ cl-cebo procedente de un único plásmido eficaz en cualquiera de los organismos. **C-E.** Secuencia en pGLS20 (SEQ ID NO: 2). **F-H.** Secuencia en pGLS22 (SEQ ID NO: 3). **I-J.** Secuencia en pBR_AMP alphaLPL (SEQ ID NO: 5).

35

Figura 2. Expresión del cebo a partir de un plásmido de expresión combinada en bacterias/levaduras. A. Los plásmidos pGLS20 y pGLS23 utilizan un promotor TEF1/uvLac combinado para expresar cebos λ cl fusionados en levaduras o bacterias. Los plásmidos se seleccionan en levaduras mediante selección para resistencia a G418 (pGLS20) o complementación HISS (pGLS23), y en bacterias mediante selección por resistencia a kanamicina (pGLS20) o por resistencia a cloranfenicol (pGLS23). La expresión relativa de cebos cl a partir de estos plásmidos, frente a la descrita previamente para los vectores pGBS10 (bihíbrido en levaduras, (Serebriiskii y col. 2002)) o pBT (bihíbrido en bacterias, Stratagene) se muestra en bacterias (panel central). **B.** Para demostrar los niveles relativos de cebo se confirmó la concentración de proteína total igual mediante tinción de Coomassie de un gel PAGE cargado con cantidades equivalentes de lisado celular para bacterias que expresan cada plásmido (no mostrados). A continuación, se cargaron volúmenes iguales de diluciones de 1:40 (para pGLS20) o 1:100 (para pGT) de extractos en tampón de muestra en paralelo con el mismo volumen de extracto sin diluir procedente de células portadoras de pG13S10. Se muestran transferencias de Western utilizando anticuerpos dirigidos contra cl. **C.** pGBS10 y pGLS20 expresan niveles comparables de cebos λ cl en levaduras, basándose en el análisis de Western con anticuerpos para λ cl. 1,2 denota dos transformantes independientes en bacterias o levaduras; -, denota levaduras que no contienen plásmido cebo.

40

45

50

Figura 3. Activación de informadores colorimétricos y auxótrofos mediante interacción de cremallera en levaduras. Los números de las bandas debajo de la gráfica de barras representan pares de muestras definidas en la Tabla 1. La gráfica de barras refleja la actividad informadora relativa medida mediante el ensayo de la beta-glucuronidasa utilizando PNP-gluc como sustrato. El recuadro indica muestras vueltas a analizar utilizando MU-gluc como sustrato. Para dar algo de contexto, los valores obtenidos para la combinación 6 (con una K_d de 1×10^{-8} M) fueron más de 10 veces superiores a los de la combinación 5 con el sustrato MU-gluc, lo que indica una función discriminadora significativa del sistema bihíbrido de levaduras en este intervalo de afinidad (no mostrado). Debajo de la gráfica de barras se muestra el crecimiento de dos puntos representativos de colonias 2 días después del cultivo en placa con un medio selectivo.

55

60

Figura 4. Activación de informadores colorimétricos y auxótrofos mediante interacción de cremallera en bacterias. Los números de las bandas debajo de la gráfica de barras representan pares de muestras definidas en la Tabla 1. La gráfica de barras refleja la actividad informadora relativa medida mediante el ensayo de la β -galactosidasa utilizando ONPG como sustrato. Debajo de la gráfica de barras se muestra el crecimiento de dos puntos representativos de colonias 24 días después del cultivo en placa con un medio selectivo.

65

Figura 5. Mapa del plásmido presa bacteriano de pBR_AMP_alphaLP_IS_B.gb. Este plásmido difiere de pAC-AMP-aLPL en que contiene un origen de replicación distinto, pBR en lugar de pAC.

5 **Figura 6.** La secuencia del promotor híbrido de la SEQ ID NO: 1.

DESCRIPCIÓN DETALLADA DE LA INVENCION

10 **[0010]** La selección de bihíbridos es una metodología habitual para identificar y caracterizar interacciones proteína-proteína y se ha convertido en un componente integral de muchas investigaciones proteómicas. El sistema bihíbrido inicialmente se desarrolló utilizando levadura como organismo hospedador. No obstante, los sistemas bihíbridos bacterianos también se han convertido en herramientas de laboratorio habituales y se prefieren en algunas circunstancias, a pesar de que los sistemas bihíbridos de levaduras y bacterias nunca se han comparado directamente. De acuerdo con la presente invención, se proporciona un sistema bihíbrido de levaduras y bacterias purificado en el que se usa un único plásmido de expresión cebo en los entornos de ambos organismos. Además, se generó una amplia serie de proteínas de fusión de cremalleras de leucina de afinidades conocidas para comparar la eficacia de la detección de la interacción utilizando ambos sistemas. A pesar de que ambos sistemas bihíbridos detectaron las interacciones producidas con un rango de afinidades de interacción comparable, cada uno de ellos demostró ventajas únicas. El sistema de levaduras produjo una lectura cuantitativa a lo largo de un rango dinámico mayor que el observado con bacterias. No obstante, el fenómeno de "autoactivación" mediante cebos fue mucho menos problemático en el sistema bacteriano que en el de levaduras. La capacidad de cambiar rápidamente entre sistemas de levaduras y bacterias suministrada con estos nuevos reactivos proporciona una ventaja notable para investigaciones con bihíbridos. Además, los vectores de expresión modificados deberían ser útiles para cualquier aplicación que requiera la expresión sencilla de una proteína de interés tanto en levaduras como en bacterias. Los sistemas bihíbridos convencionales han sido descritos en las patentes de EE.UU. 5.580.736 y 6.326.150, cuyos contenidos se incorporan en el presente documento por referencia. Además, en la presente invención se proporcionan kits útiles para llevar a cabo los procedimientos descritos en el presente documento.

30 **[0011]** Las siguientes definiciones se proporcionan para facilitar la comprensión de la presente invención.

[0012] Como se usa en el presente documento, "gen informador" se refiere a un gen cuya expresión se puede someter a ensayo; dichos genes incluyen, sin limitación, lacZ, β -glucuronidasa (GUS), genes de la biosíntesis de aminoácidos, por ejemplo, los genes de levadura LEU2, HIS3, LYS2 o URA3, genes de la biosíntesis de ácidos nucleicos, el gen de la cloranfenicol-transacetilasa de mamífero (CAT), el gen de la proteína fluorescente verde (GFP) o cualquier gen de un antígeno de superficie para el que haya disponibles anticuerpos específicos.

[0013] Un "promotor" es una secuencia de ADN situada próxima al inicio de la transcripción en el extremo 5' de una secuencia transcrita unida de manera operable. El promotor puede contener uno o más elementos reguladores o módulos que interaccionan en la modulación de la transcripción del gen unido de manera operable.

40 **[0014]** "Unido de manera operable" describe dos elementos macromoleculares dispuestos de manera que la modulación de la actividad del primer elemento induce un efecto sobre el segundo elemento. De esta forma, la modulación de la actividad de un elemento promotor se puede utilizar para alterar y/o regular la expresión de una secuencia codificante unida de manera operable. Por ejemplo, la transcripción de una secuencia codificante que está unida de manera operable a un elemento promotor se induce mediante factores que "activan" la actividad del promotor; la transcripción de una secuencia codificante que está unida de manera operable a un elemento promotor se inhibe mediante factores que "reprimen" la actividad del promotor. Así, una región promotora está unida de manera operable a la secuencia codificante de una proteína si la actividad de transcripción de dicha secuencia codificante se ve influida por la actividad del promotor.

50 **[0015]** "Construcción de fusión" se refiere de manera general a genes recombinantes que codifican proteínas de fusión.

[0016] Una "proteína de fusión" es una proteína híbrida, es decir, una proteína que se ha construido para que contenga dominios procedentes de al menos dos proteínas diferentes. Como se usa en el presente documento, una proteína de fusión es una proteína híbrida que posee (a) un dominio regulador de la transcripción procedente de una proteína reguladora de la transcripción, o (b) un dominio de unión al ADN procedente de una proteína de unión al ADN unido a una proteína heteróloga para comprobar la interacción. La estructura de la proteína de fusión es tal que el dominio regulador de la transcripción y el dominio de unión al ADN están dispuestos de manera que permita que ambos dominios sean biológicamente activos. La proteína de la que procede el dominio regulador de la transcripción es diferente de la proteína de la que procede el dominio de unión al ADN. En otras palabras, los dos dominios son heterólogos entre sí.

65 **[0017]** El dominio regulador de la transcripción de la proteína de fusión puede activar o reprimir la transcripción de los genes diana, dependiendo de la actividad biológica nativa del dominio. Las proteínas cebo de la invención también son proteínas de fusión codificadas mediante un gen de fusión que comprende una proteína de interés

unida de manera operable a un resto de unión al ADN.

[0018] El término "gen de una proteína de fusión" se refiere a una secuencia de ADN que codifica una proteína de fusión. Un gen de una proteína de fusión además puede proporcionar elementos reguladores de la transcripción y de la traducción para el control de la transcripción y la traducción del mismo.

[0019] La "expresión" es el proceso mediante el cual se revela la información codificada en un gen. Si el gen codifica una proteína, la expresión supone tanto la transcripción del ADN en ARNm, y el procesamiento del ARNm (si fuera necesario) en un producto de ARNm maduro, como la traducción del ARNm maduro en una proteína.

[0020] Se dice que una molécula de ácidos nucleicos, tal como un ADN o un gen, es "capaz de expresar" un polipéptido si la molécula contiene las secuencias codificantes para el polipéptido y las secuencias para el control de la expresión que, en el entorno apropiado dentro del hospedador, proporcionan la capacidad para transcribir, procesar y traducir la información genética contenida dentro del ADN en un producto proteico, y si dichas secuencias para el control de la expresión están unidas de manera operable a la secuencia de nucleótidos que codifica el polipéptido.

[0021] Como se usa en el presente documento, un "vehículo de clonación" es cualquier entidad que sea capaz de introducir una secuencia de ácidos nucleicos en una célula hospedadora para su clonación. Ejemplos de vehículos de clonación incluyen plásmidos o genomas fágicos. Se desea especialmente un plásmido que se pueda replicar de manera autónoma en la célula hospedadora. Alternativamente, es útil una molécula de ácidos nucleicos que se pueda insertar (integrar) en el ADN cromosómico de la célula hospedadora, especialmente una molécula que se inserte en el ADN cromosómico de la célula hospedadora de una manera estable, es decir, de forma que permita que dicha molécula sea heredada por las células hijas.

[0022] Los vehículos de clonación a menudo están caracterizados por un o por un pequeño número de sitios de reconocimiento de endonucleasas en los que dichas secuencias de ADN se pueden cortar de una forma determinada sin la pérdida de una función biológica esencial del vehículo, y en los que el ADN se puede empalmar con el fin de conseguir su replicación y clonación.

[0023] El vehículo de clonación además puede contener un marcador adecuado para su uso en la identificación de células transformadas con el vehículo de clonación. Por ejemplo, un "gen marcador" puede ser un gen que confiera resistencia a un antibiótico específico a la célula hospedadora.

[0024] La palabra "vector" a veces se usa indistintamente con "vehículo de clonación".

[0025] Como se usa en el presente documento, un "vehículo de expresión" es un vehículo o vector similar al vehículo de clonación pero que está diseñado especialmente para proporcionar un entorno que permita la expresión del gen clonado después de la transformación en el hospedador. Una forma de proporcionar dicho entorno es incluir secuencias reguladoras de la transcripción y de la traducción en dicho vehículo de expresión, dichas secuencias reguladoras de la transcripción y de la traducción que se pueden unir de manera operable al gen clonado. Otra forma de proporcionar dicho entorno es suministrar un sitio o sitios de clonación a dicho vehículo, en el que se pueden clonar un gen clonado deseado y elementos reguladores de la expresión deseados. En un vehículo de expresión, el gen a clonar normalmente está unido de manera operable a ciertas secuencias control tales como secuencias promotoras. Las secuencias para el control de la expresión variarán dependiendo de si el vector está diseñado para expresar el gen unido de manera operable en un hospedador procarionta o eucariota y adicionalmente pueden contener elementos transcripcionales tales como elementos potenciadores, secuencias de terminación, elementos para la especificidad de tejido, y/o sitios para la iniciación y terminación de la traducción.

[0026] Un "hospedador" se refiere a cualquier organismo no humano que sea el receptor de un vehículo de clonación o expresión. En formas de realización preferidas, el hospedador de la invención es una célula de levaduras o una célula animal cultivada tal como una célula de mamífero o de insecto. En una forma de realización especialmente preferida, la célula de levadura es *Saccharomyces cerevisiae*.

[0027] Un "resto de unión" es un segmento de aminoácidos que es capaz de dirigir la unión del polipéptido específico a una secuencia de ADN particular (es decir, un "sitio de unión a la proteína"). También denominadas en el presente documento como dominios de unión al ADN, estas proteínas pueden ser homodímeros o monómeros que se unen al ADN de una manera específica de secuencia. Ejemplos de dominios de unión al ADN de la invención incluyen LexA, cl, dominios de unión al receptor de glucocorticoides y el dominio Ume6.

[0028] Un "resto que activa un gen" es un segmento de aminoácidos que es capaz de inducir débilmente la expresión de un gen a cuya región control está unido. Como se usa en el presente documento, "débilmente" quiere decir por debajo del nivel de activación llevado a cabo por la región II de activación de GAL4 (Ma y Ptashne, Cell, 48: 347, 1987) y preferentemente se encuentra al o por debajo del nivel de activación producido por el dominio de activación B42 de Ma y Ptashne (Cell, 51: 413, 1987). Los niveles de activación se pueden medir utilizando cualquier sistema de genes informadores aguas abajo y comparando, en ensayos en paralelo, el nivel de expresión

estimulado por la región II de GAL4-polipéptido con el nivel de expresión estimulado por el polipéptido a probar.

[0029] "ADN purificado" es ADN que no se encuentra inmediatamente contiguo a las dos secuencias codificantes a las cuales está inmediatamente contiguo (una del extremo 5' y otra del extremo 3') en el genoma en estado natural del organismo del que procede. Por tanto el término incluye, por ejemplo, un ADN recombinante que se haya incorporado a un vector; a un plásmido o virus que se replica de manera autónoma; o al ADN genómico de un procarionte o una eucariota, o que existe como molécula distinta (por ejemplo, un ADNc o un fragmento de ADN genómico producido mediante PCR o tratamiento con endonucleasas de restricción) independiente de otras secuencias. También incluye ADN recombinante que sea parte de un gen híbrido que codifica secuencias polipeptídicas adicionales.

[0030] "Sustancialmente idéntico", en referencia a una secuencia de aminoácidos, significa una secuencia de aminoácidos que difiere solamente en sustituciones conservativas de aminoácidos, por ejemplo, la sustitución de un aminoácido por otro del mismo tipo (por ejemplo, valina por glicina, arginina por lisina, etc.) o por una o más sustituciones no conservativas, eliminaciones, o inserciones localizadas en posiciones de la secuencia de aminoácidos que no destruya la función de la proteína (sometida a ensayo, por ejemplo, como se describe en el presente documento). Una secuencia de ácidos nucleicos "sustancialmente idéntica" codifica una secuencia de aminoácidos sustancialmente idéntica como se ha definido anteriormente.

[0031] Una "célula transformada" es una célula de levaduras o de bacterias en la que (o en uno de cuyos ancestros) se ha introducido ADN exógeno por medio de técnicas de ADN recombinante.

[0032] La frase "situado para la expresión" se refiere a una molécula que codifica ADN que está situado adyacente a una secuencia de ADN que dirige la transcripción y traducción de la secuencia.

[0033] Un "anticuerpo purificado" es un anticuerpo en el que al menos el 60 % en peso está exento de las proteínas y moléculas orgánicas de origen natural a las que está asociado en estado natural. Preferentemente, la preparación comprende anticuerpo en una cantidad de al menos el 75 % en peso, más preferentemente al menos el 90 % en peso, y lo más preferentemente al menos el 99 % en peso.

[0034] Los siguientes ejemplos se proporcionan para facilitar la comprensión de la presente invención. No se pretende que limiten la invención de ningún modo.

EJEMPLO I

[0035] Los siguientes materiales y procedimientos se proporcionan para facilitar la puesta en práctica de este ejemplo.

Manipulación molecular y microbiológica. La clonación de nuevas construcciones se llevó a cabo utilizando protocolos convencionales. Los detalles de las secuencias y los sitios de clonación albergados en los plásmidos se describen en la sección de resultados y a continuación se proporcionan otras caracterizaciones básicas de las propiedades de expresión de estos plásmidos.

[0036] Resumiendo, el plásmido pGLS20 se construyó reemplazando el promotor ADH1 de pGKS9 con una combinación del promotor TEF1 (procedente del plásmido pLexZeo, Invitrogen) y un promotor lacUV5 (procedente del plásmido pBT, Stratagene). Para producir pGLS23, se construyó un casete HIS5Cm^R en el vector pCR2.1 mediante la combinación de un casete HIS5 procedente de pJFK (R. Hopkins, no publicado) y un casete Cm^R procedente de pMW108. A continuación este casete se utilizó para sustituir el casete G418R en pGLS20. En consecuencia, las diferencias entre pGLS22 y pGLS23 son mínimas -un cambio en la posición 3267 del pGLS22 (gaattC --> gaattA) lo convierte en pGLS23 con la destrucción de un sitio EcoRI. La Figura 4 muestra los resultados utilizando el vector pGLS23 (sin sitio EcoRI).

[0037] El plásmido presa bihíbrido bacteriano pAc-AMP- α LPL se construyó reemplazando el gen de resistencia a cloranfenicol presente en el plásmido pKJ1267 (J.K.J., no publicado) con el gen de resistencia a ampicilina procedente del plásmido pACYC177. Las secuencias de la cremallera de leucinas se seleccionaron entre los péptidos descritos en (Krylov y col. 1998; Krylov y col. 1994; Moll y col. 2001): se sintetizó artificialmente ADN para que codificase las secuencias peptídicas descritas. Para fusionar las diferentes cremalleras de leucina al dominio amino-terminal y al enlazador inter-dominios de la subunidad α de la ARN polimerasa de *E. coli*, en el plásmido se insertaron fragmentos de ADN que codifican variantes de las cremalleras utilizando los sitios de restricción únicos Not I y Xho I.

[0038] *Expresión las proteínas cebo y presa.* La expresión de las proteínas cebo y presa (a excepción de fusiones de la subunidad α de la ARN polimerasa bacteriana, para la cual no se disponía de anticuerpo) se confirmó mediante análisis de Western, con anticuerpo primario contra cl para los cebos, o hemaglutinina para las presas expresadas en levadura. Aunque cl es el cebo puesto de ejemplo en el presente documento, se conocen muchos cebos diferentes y están disponibles para la persona experta. Para comparar los niveles de expresión de proteína cl en *E.*

coli, se transformaron los plásmidos correspondientes en la cepa DH5a y se prepararon extractos de proteínas a partir de cultivos en crecimiento exponencial. La concentración de proteína igual se confirmó mediante tinción con Coomassie de un gel PAGE, y a continuación se cargaron en paralelo volúmenes iguales de diluciones de 1:40 (para pGLS20) o 1:100 (para pBT) de los extractos en tampón de muestras con el mismo volumen de extractos sin diluir
5 procedentes de células portadoras de pGLS10. Las proteínas se resolvieron en un gel PAGE, y se llevó a cabo un análisis de transferencia de Western, utilizando anticuerpos dirigidos contra cl. Para comparar los niveles de expresión de proteína cl en la levadura, los plásmidos correspondientes se transformaron en la cepa SKY191 y se prepararon extractos de proteínas a partir de cultivos en crecimiento exponencial. La concentración de proteína igual se confirmó mediante tinción con Coomassie de un gel PAGE (no mostrado). A continuación se cargaron en el gel
10 volúmenes iguales de los extractos en tampón de muestras, y se llevó a cabo un análisis de transferencia de Western.

[0039] *Ensayos informadores.* Para las levaduras, la actividad de informadores cuantitativos se determinó en un lector de placas utilizando una técnica modificada de Serebriiskii y col. 2000. Brevemente, a 50 μ l de cultivos en crecimiento exponencial en los pocillos de las células en 96 pocillos se le añadió un volumen igual de tampón Z 2x
15 que contiene 2 mg/ml del sustrato correspondiente y el 50 % de Y-PER (Pierce), para la levadura. La actividad se calcula como $(DO_{420f} - DO_{420i})$ dividido por la DO_{600} , donde la diferencia entre la DO_{420i} y la DO_{420f} (lecturas inicial y final) refleja la conversión del sustrato incoloro (PNPgluc) en producto amarillo durante un período de tiempo de ~10-30 minutos, y la DO_{600} es una medida de la densidad celular en una muestra dada. Para cada punto de los
20 datos en cada experimento con levaduras, se midieron y se promediaron las actividades de 5 a 8 clones. Todas las lecturas se tomaron en un lector de placas; se demostró previamente (Serebriiskii y col. 2000) que las mediciones del lector de placas y las unidades derivadas están correlacionadas de manera proporcional con las unidades de DO tomadas en un espectrofotómetro.

[0040] Para mediciones del gen informador de la β -galactosidasa bacteriana, se realizaron ensayos esencialmente como se ha descrito (Thibodeau y col. 2004). Brevemente, cultivos inoculados procedentes de una única colonia fresca se crecieron hasta la mitad de la fase logarítmica y se lisaron mediante la adición de un volumen de 1/10 de PopCulture™ (Novagen). En una placa de microtitulación de 96 pocillos, se añadieron 15 μ l de lisado celular a una mezcla de 135 μ l de tampón Z y 30 μ l de ONPG a una concentración de 4 mg/ml para iniciar la reacción. Se llevaron
30 a cabo ensayos cinéticos mediante el control de la DO_{415} durante 0-30 minutos utilizando un lector de placas. Todos los ensayos con la β -galactosidasa bacteriana se realizaron por triplicado.

[0041] Se sometieron a ensayo informadores auxótrofos como se describe en (Serebriiskii y Joung 2002). Los plásmidos cebo y presa se transformaron en la cepa de selección correspondiente, *S. cerevisiae* SKY191 o *E. coli* KJ1567. El crecimiento en placas de selección se midió durante 5 días (levaduras; nota, todas las colonias que se desarrollaron eran prominentes a los 2 días) o 1 día (bacterias).

RESULTADOS

[0042] Se han desarrollado plásmidos que facilitan la expresión y selección en paralelo de una sola proteína cebo tanto en sistemas bihíbridos de levaduras como bacterianos utilizando un único plásmido de expresión (Figura 1). Como se muestra en la Figura 1, las proteínas cebo se expresan como fusiones a la proteína λ cl tanto en sistemas bihíbridos de levaduras como bacterianos. Para conseguir esto, se realizaron varias modificaciones al plásmido pGBS9 (Serebriiskii y col. 2002), desarrollado originalmente para expresar las proteínas cebo como fusiones al
45 represor λ cl en un sistema bihíbrido de levaduras (Figura 2A). El promotor ADH1 procedente de este plásmido se reemplazó con un promotor en tándem, en el que tanto el promotor TEF1 (Nagashima y col. 1986) de *S. cerevisiae* como el promotor lac UV5 de *E. coli* extremadamente potentes dirigen la expresión de una secuencia codificante λ cl y un sitio de clonación de un polienlazador. El plásmido resultante, pGLS20, se puede mantener en levaduras o en bacterias mediante G418 o por resistencia a kanamicina, respectivamente (Figura 2A). Otros derivados del plásmido
50 estrechamente relacionados (pGLS22, pGLS23) albergan el gen HIS5 que confiere la selección en levaduras y resistencia a cloranfenicol para la selección en bacterias (Figura 2A). Como se muestra en la Figura 2B, la expresión del represor λ cl utilizando el plásmido pGLS20 en bacterias es comparable a la obtenida con el plásmido pBT (un vector optimizado para el sistema bihíbrido bacteriano, Stratagene) y es más de 40 veces superior a la proporcionada por el plásmido de expresión bihíbrido habitual en levaduras pGBS10 (Serebriiskii y col., 2002). En
55 levaduras, la expresión de fusiones del represor λ cl procedente de pGLS20 y sus derivados es comparable o superior a la de pGBS10 (Figura 2C).

[0043] Se han utilizado estos plásmidos pGLS bifuncionales para determinar si los sistemas bihíbridos de levaduras y bacterianos mostraban alguna diferencia en sus capacidades para detectar una serie de interacciones
60 con afinidades diferentes. Para ello, se ha creado una serie de proteínas de fusión cebo y presa utilizando un grupo de variantes de cremalleras de leucina previamente caracterizadas (Krylov y col. 1998; Krylov y col. 1994; Moll y col. 2001) con afinidades de interacción definidas que abarcan de $K_d > 10^{-4}$ a 10^{-15} M como se determina *in vitro* (Tabla 1). Para el análisis en el sistema bihíbrido bacteriano, se utilizó el plásmido pAC-AMP- α LPL (Tabla 2) para expresar las proteínas presa a partir de un promotor en tándem lpp/acUV5 fuerte e inducible como fusiones al dominio amino-
65 terminal de la subunidad α de la ARN polimerasa. Para el sistema bihíbrido de levaduras, se utilizó pJG4-5 (Gyuris y col. 1993) para expresar las proteínas presa a partir de un promotor GAL 1 inducible como fusiones al dominio

sintético de activación de la transcripción B42 (Figura 1). La capacidad de cada par de cremalleras para activar la transcripción de un informador cuantitativo y un informador auxótrofo se evaluó entonces en bacterias y en levaduras.

5 Tabla 1. Propiedades de las cremalleras de leucina utilizadas en este ejemplo.

Los cálculos de pI se realizaron utilizando el sitio http://us.expasy.org/tools/pi_tool.html. Las cremalleras de leucina para muchos de los cebos habían sido descritas originalmente, y las propiedades de interacción habían sido caracterizadas *in vitro* en (Krylov y col. 1994).

10

Combinación	Cebo	pl	Presa	K _d para el cebo-presa (en M)
1	EE ₁₂₃₄₅ L	4,2	EE ₁₂₃₄₅ L	No detectable
2	RR ₁₂ EE ₃₄₅ L	6,5	RR ₁₂ EE ₃₄₅ L	
3	EE ₃₄	5,3	EE ₃₄	-4 8,1 x 10 ⁻⁴
4	RR ₃₄	10,5	RR ₃₄	3,9 x 10 ⁻⁵
5	RR ₁₂₃₄ L	11,8	RR ₁₂₃₄ L	2,5 x 10 ⁻⁷
6	RR ₃₄	10,5	EE ₃₄	1,0 x 10 ⁻⁸
7	EE ₃₄	5,3	RR ₃₄	1,0 x 10 ⁻⁸
8	RR ₁₂ EE ₃₄₅ L	6,5	EE ₁₂ RR ₃₄₅ L	1,3 x 10 ⁻¹¹
9	EE ₁₂ RR ₃₄₅ L	10,4	RR ₁₂ EE ₃₄₅ L	1,3 x 10 ⁻¹¹
10	RR ₁₂₃₄₅ ^L	12,2	EE ₁₂₃₄₅ L	1,1 x 10 ⁻¹¹
11	RR ₁₂₃₄ L	11,8	EE ₁₂₃₄ L	1,0 x 10 ⁻¹⁵
12	EE ₁₂₃₄ L	4,3	RR ₁₂₃₄ L	1,0 x 10 ⁻¹⁵

Tabla 2. Cepas y plásmidos utilizados en este estudio.

Plásmidos	Selección en levaduras/en <u><i>E. coli</i></u>	Comentario/Descripción
Cebos		
pGLS20*	G418 ^R	Km ^R El promotor TEF garantiza la expresión de cl en levaduras, mientras que el promotor <i>lpp/lacUV5</i> permite la expresión en <u><i>E. coli</i></u>
pGLS22/23*	<i>HIS5</i>	Cm ^R Similar a pGLS20, véase el texto para más detalles
Informadores		
pRG61	URA3*	(Km ^R) Los operadores <i>lcl</i> dirigen la transcripción del gen <i>gusA</i> ; pRG61 es un informador menos sensible y con menos historial que pDR8
pDR8		
Fusiones del dominio de activación		
pJG4-5	<i>TRP1</i>	Ap ^R El promotor <i>GAL1</i> proporciona una expresión eficaz en levaduras de un gen fusionado a un casete que consta de secuencia de localización nuclear, dominio de activación de la transcripción, y marcador del epítipo HA
pAC-AMP- α LPL*	N/A	Ap ^R En cada uno de los dos plásmidos presa, los promotores en tándem <i>lpp/lacUV5</i> proporcionan una expresión eficaz en <u><i>E. coli</i></u> de un gen fusionado a los restos 1-248 de la subunidad α de la ARNP de <u><i>E. coli</i></u> . pAC-AMP- α LPL tiene el origen de replicación de pACYC, mientras que pBR-AMP- α LPL tiene el origen de replicación de pBR322. Así, se pueden regular el número de copias de los plásmidos cebo/presa, y por tanto los niveles de expresión de cebo/presa
pBR-AMP- α LPL*	N/A	Ap ^R
Cepas		
<u><i>S. cerevisiae</i></u> SKY191	MAT α <i>trp1, his3, ura3, clop-LYS2</i>	Comentario/Descripción
<u><i>S. cerevisiae</i></u> PRT50*	MAT α <i>trp1, his3, HIS5, ura3, clop-LYS2</i>	Cepas informadoras en las que la expresión del gen informador <i>LYS2</i> está dirigida por un promotor débil que porta un sitio de unión al ADN λ cl
<u><i>S. cerevisiae</i></u> PRT475	MAT α <i>trp1, his3, ura3, clop-LYS2</i>	Igual que el anterior

<i>E. coli</i> KJ1567*	$\Delta hisB463, \Delta(GPT-$ <i>proAB-Arg-lac) XIII zaj::</i> <i>Tn10 [F'lacI^q HIS3</i> <i>aadA Kan^R]</i>	Cepa informadora en la que la expresión de los genes informadores HIS3 y <i>aadA</i> está dirigida por un promotor débil que porta un sitio de unión al ADN λ cl
<i>E. coli</i> AG58A (RP28)*	$\Delta hisB463, \Delta(GPT-$ <i>proAB-Arg-lac) XIII zaj::</i> <i>Tn10 [F'lacI^q lacZ Kan^R]</i>	Cepa informadora en la que la expresión del gen informador <i>lacZ</i> está dirigida por un promotor débil que porta un sitio de unión al ADN λ cl

* Reactivo producido en este estudio. Se han descrito pRG61 (Serebriiskii y col. 2002), pDR8 (Serebriiskii y col. 2002), SKY191 (Serebriiskii y col. 1999) y pJG4-5 (Gyuris y col. 1993).

[0044] Los resultados obtenidos con el sistema basado en levaduras demuestran que las combinaciones de cremalleras cebo-presa activan la transcripción de un informador β -glucuronidasa (GusA) cuantificable a lo largo de un intervalo sustancial de afinidades (Figura 3, gráfico de barras). En este ensayo, los pares de cremalleras con interacción reportada y constantes de disociación de 1×10^{-8} o inferior (bandas 6-12) activaron fuertemente la expresión del gen informador, detectado utilizando un sustrato colorimétrico (PNP-gluc). Aquellos con valores de K_d de $2,5 \times 10^{-7}$ M o superior (con una excepción -véase más abajo) no activaron fuertemente el gen informador (Figura 3A, bandas 1-5). En general, la actividad β -glucuronidasa se indujo del orden de ~30-180 veces sobre los valores basales con los pares de cremalleras de leucina con mayor afinidad. Las pruebas adicionales de los pares que interaccionan con una menor afinidad utilizando un sustrato fluorescente más sensible para la β -glucuronidasa, el MU-gluc (Figura 3A, recuadro) indicaron que también es posible detectar convincentemente interacciones en el intervalo de 10^{-7} M, aunque la estimulación de la expresión del gen GusA observado en estas muestras es notablemente menos fuerte que las obtenidas con interacciones en el intervalo de 10^{-8} M. Con la cepa informadora auxótrofa (Figura 3, paneles bajo la gráfica de barras), las células crecieron en condiciones selectivas sólo si las cremalleras que interaccionan poseían constantes de disociación de 1×10^{-8} M o inferior, replicando los resultados obtenidos con el gen informador GusA cuantitativo. El sistema no presenta una capacidad significativa para discriminar interacciones que interactúan con constantes de disociación de 10^{-8} M o inferior, lo que sugiere que la expresión del gen informador estaba saturada. Es importante destacar que, para algunos de los cebos examinados, la expresión del cebo solo en ausencia de la presa era suficiente para activar fuertemente la transcripción de los informadores, lo que hace difícil demostrar convincentemente la interacción con proteínas (véase Figura 3, muestras 1, 10 y 12).

[0045] A continuación se examinaron las capacidades de las mismas combinaciones de cremalleras cebo-presa para activar la transcripción en el sistema bihíbrido bacteriano (Figura 1) utilizando el gen informador *lacZ* cuantificable (Figura 4). De acuerdo con nuestros resultados en el sistema basado en levaduras, los pares de cremalleras de leucina con constantes de disociación reportadas inferiores a 10^{-8} M estimulaban claramente la expresión del gen informador *lacZ* (Figura 3, muestras 6-12) mientras que los pares de interacción con constantes de disociación de $2,5 \times 10^{-7}$ M o superior no conseguían estimular la expresión de *lacZ* (Figura 3, muestras 1-5). También se analizó la activación basada en las cremalleras del gen informador auxótrofo HIS3 (Figura 4, paneles bajo la gráfica de barras). Los resultados obtenidos utilizando el gen informador auxótrofo HIS3 replicaron con poca desviación los resultados obtenidos con el informador *lacZ*: sólo las células que albergan pares de cremalleras con constantes de disociación de 1×10^{-8} M o inferior presentaban crecimiento después de 24 horas sobre placas selectivas. En contraste con los resultados obtenidos en el sistema basado en levaduras, ninguno de los cebos probado mostró autoactivación en ausencia de los pares de presa (comparar las muestras 1, 10 y 12 en las Figuras 3 y 4).

[0046] Los resultados obtenidos utilizando un grupo estrechamente relacionado de fusiones cebo y presa de cremalleras de leucina pequeñas sugieren ventajas diferenciales para la detección de interacciones proteína-proteína en sistemas bihíbridos de levaduras y bacterianos. En primer lugar, los resultados obtenidos utilizando informadores cuantificables sugieren que el sistema basado en levaduras posee un rango dinámico más amplio para la detección de interacciones (contrastar las Figuras 3 y 4). En el sistema de levaduras, las interacciones caracterizadas por constantes de disociación de hasta 10^{-8} M se pueden detectar como un incremento en la expresión del gen informador GusA (o de hasta 10^{-7} M si se utiliza un sustrato más sensible para la detección de GusA). En contraste, en el sistema bacteriano, sólo las interacciones caracterizadas por constantes de disociación de 10^{-8} M o inferior se pueden detectar como un incremento en la expresión de *lacZ*. En segundo lugar, se observó que los experimentos realizados con sistemas bihíbridos bacterianos producen colonias en medio selectivo un poco más rápidamente que las obtenidas en el sistema de levaduras (un día frente a dos). En tercer lugar, los resultados también sugieren que la autoactivación por las proteínas cebo es probable que sea menos problemática en bacterias que en levaduras (comparar las bandas 1, 10, 12 en las Figuras 3 y 4). Este hallazgo no es del todo sorprendente teniendo en cuenta las diferencias fundamentales en los mecanismos de activación de los genes y la distancia evolutiva entre procariontes y eucariontes. La menor frecuencia de auto-activación por cebos en el sistema bihíbrido bacteriano es una ventaja potencialmente importante de este sistema en comparación con su homólogo en levaduras.

- [0047]** Los datos también sugieren que la fuerza de interacción umbral necesaria para una activación consistente de la transcripción es similar en ambos organismos. Tanto en sistemas de levaduras como bacterianos, la activación completa parece requerir una afinidad de interacción entre las proteínas de fusión cebo y presa definida por una constante de disociación en el intervalo de 10^{-7} a 10^{-8} M. A pesar de que los resultados demuestran una transición brusca entre la no activación y la activación completa de los genes informadores, estudios previos en ambos sistemas han demostrado que la magnitud de la activación de la transcripción observada puede estar relacionada con la afinidad de la interacción cebo-presa (Dove y col. 1997; Estojak y col. 1995). Aunque se desconoce el motivo exacto de esta diferencia en los resultados en comparación con estudios anteriores, se observa que Estojak y sus colaboradores evaluaron las interacciones utilizando una serie de informadores de estrictencia variable (es decir, que contienen un número diferente de sitios de unión para los cebos) para expandir el intervalo de detección: no hay ninguna limitación técnica para utilizar una estrategia similar con este nuevo sistema. En general, los resultados sugieren que el uso del sistema actual como herramienta de selección funcionará mejor para la detección de interacciones con constantes de disociación en el intervalo nanomolar de medio a alto.
- 15 **[0048]** Los inventores indican que, a su entender, esta es la primera descripción de una combinación de promotores que es potente tanto en entornos de levaduras como bacterianos. De hecho, se ha encontrado que los plásmidos pGLS expresan niveles suficientes de proteínas de fusión cebo para su actividad en sistemas bihíbridos bacterianos incluso sin inducir el promotor bacteriano fuerte. Los niveles de expresión del cebo se pueden regular adicionalmente con la selección de los plásmidos presa: pBR-AMP- α LPL comparte el mismo origen de replicación de pBR con los plásmidos cebo pGLS. Por lo tanto, mediante la co-transformación en *E. coli*, el número total de copias de cebo es menor, lo que podría ser una ventaja si la sobreexpresión de la proteína cebo fuera perjudicial para la célula *E. coli*. La utilización de pAC-AMP- α LPL, que tiene un origen de replicación (pACYC) diferente, permite el número de copias completo del plásmido cebo y por tanto posibilita niveles de expresión más elevados. Por último, si bien este trabajo se centra en el uso de los plásmidos pGLS en un contexto bihíbrido, los inventores anticipan que su diseño general del promotor también podría ser útil en otros estudios de caracterización funcional.
- 25

[0049]

- Chien, C.T., P.L. Bartel, R. Stennglanz, and S. Fields. 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Nat. Acad. Sci. USA* 88: 9578-9582.
- Dove, S.L., J.K. Joung, and A. Hochschild. 1997. Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* 386: 627-630.
- Estojak, J., R. Brent, and E.A. Golemis. 1995. Correlation of two-hybrid affinity data with in vitro measurements. *Mol. Cell. Biol.* 15: 5820-5829.
- 35 Fields, S. and O. Song. 1989. A novel genetic system to detect protein-protein interaction. *Nature* 340: 245-246. Giot, L., J.S. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y.L. Hao, C.E. Ooi, B. Godwin, E. Vitols, G. Vijayadamar, P. Pochart, H. Machineni, M. Welsh, Y. Kong, B. Zerhusen, R. Malcolm, Z. Varrone, A. Collis, M. Minto, S. Burgess, L. McDaniel, E. Stimpson, F. Spriggs, J. Williams, K. Neurath, N. Ioime, M. Agee, E. Voss, K. Furtak, R. Renzulli, N. Aanensen, S. Carrolla, E. Bickelhaupt, Y. Lazovatsky, A. DaSilva, J. Zhong, C.A. Stanyon, R.L. Finley, Jr., K.P. White, M. Braverman, T. Jarvie, S. Gold, M. Leach, J. Knight, R.A. Shimkets, M.P. McKenna, J. Chant, and J.M. Rothberg. 2003. A protein interaction map of *Drosophila melanogaster*. *Science* 302: 1727-1736.
- 40 Gyuris, J., E.A. Golemis, H. Chertkov, and R. Brent. 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75: 791-803.
- Ito, T., K. Tashiro, S. Muta, R. Ozawa, T. Chiba, M. Nishizawa, K. Yamamoto, S. Kuhara, and Y. Sakaki. 2000. Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc Natl Acad Sci U S A* 97: 1143-1147.
- Joung, J.K., E.I. Ramm, and C.O. Pabo. 2000. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci USA* 97: 7382-387.
- Krylov, D., J. Barchi, and C. Vinson. 1998. Inter-helical interactions in the leucine zipper coiled coil dimer: pH and salt dependence of coupling energy between charged amino acids. *J Mol Biol* 279: 959-972.
- 50 Krylov, D., I. Mikhailenko, and C. Vinson. 1994. A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. *Embo J* 13: 2849-2861.
- Li, S., C.M. Armstrong, N. Bertin, H. Ge, S. Milstein, M. Boxem, P.O. Vidalain, J.D. Han, A. Chesneau, T. Hao, D.S. Goldberg, N. Li, M. Martinez, J.F. Rual, P. Lamesch, L. Xu, M. Tewari, S.L. Wong, L.V. Zhang, G.F. Berriz, L. Jacotot, P. Vaglio, J. Reboul, T. Hirozane-Kishikawa, Q. Li, H.W. Gabel, A. Elewa, B. Baumgartner, D.J. Rose, H. Yu, S. Bosak, R. Sequerra, A. Fraser, S.E. Mango, W.M. Saxton, S. Strome, S. Van Den Heuvel, F. Piano, J. Vandenhaute, C. Sardet, M. Gerstein, L. Doucette-Stamm, K.C. Gunsalus, J.W. Harper, M.E. Cusick, F.P. Roth, D.E. Hill, and M. Vidal. 2004. A map of the interactome network of the metazoan *C. elegans*. *Science* 303: 540-543.
- Moll, J.R., S.B. Ruvinov, I. Pastan, and C. Vinson. 2001. Designed heterodimerizing leucine zippers with a ranger of pIs and stabilities up to $10^{(-15)}$ M. *Protein Sci.* 10: 649-655.
- 60 Nagashima, K., M. Kasai, S. Nagata, and Y. Kaziro. 1986. Structure of the two genes coding for polypeptide chain elongation factor 1 alpha (EF-1 alpha) from *Saccharomyces cerevisiae*. *Gene* 45: 265-273.
- Serebriiskii, I. and E.A. Golemis. <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>.
- Serebriiskii, I., V. Khazak, and E.A. Golemis. 1999. A two-hybrid dual bait system to discriminate specificity of protein interactions. *J. Biol. Chem.* 274: 17080-17087.
- 65

- Serebriiskii, I.G. and J.K. Joung. 2002. Yeast and bacterial two-hybrid selection systems for studying protein-protein interactions. In *Protein-Protein Interactions: A Molecular Cloning Manual* (ed. E. Golemis), pp. 93-142. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 5 Serebriiskii, I.G., O. Mitina, E. Pugacheva, E. Benevolenskaya, E. Kotova, G.G. Toby, V. Khazak, W.G. Kaelin, J. Chernoff, and E.A. Golemis. 2002. Detection of peptides, proteins, and drugs that selectively interact with protein targets. *Genome Res.* 12: 1785-1791.
- Serebriiskii, I.G., G.G. Toby, and G. E.A. 2000. Streamlined yeast colorimetric reporter assays, using scanners and plate readers. *Biotechniques* 29: 278-279, 282-274, 286-278.
- 10 Thibodeau, S.A., R. Fang, and J.K. Joung. 2004. High-throughput beta-galactosidase assay for bacterial cell-based reporter systems. *Biotechniques* 36: 410-415.
- Uetz, P., L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qweshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleish, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, and J.M. Rothberg. 2000. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403: 623-627.
- 15 Vojtek, A.B., S.M. Hollenberg, and J.A. Cooper. 1993. Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* 74: 205-214.

[0050] Aunque esta invención se ha descrito con referencia a formas de realización específicas, los expertos en la materia pueden llevar a cabo otras formas de realización y variaciones de esta invención sin apartarse del verdadero alcance de la invención. Se pretende que las reivindicaciones adjuntas incluyan todas estas formas de realización y variaciones equivalentes.

20

REIVINDICACIONES

1. Un ácido nucleico aislado que comprende una secuencia promotora procedente del nucleótido en posición 4 al nucleótido en posición 545 de la Figura 1C que dirige la expresión de una secuencia codificante unida de manera operable tanto en levaduras como en bacterias.
2. Un plásmido que comprende el promotor de la reivindicación 1 seleccionado del grupo constituido por la SEQ ID NO: 2 (pGLS20), SEQ ID NO: 3 (pGLS22) y pGLS23 en el que pGLS23 es idéntico a pGLS22 excepto por un cambio en la posición 3267 de gaattC a gaattA.
3. Una célula hospedadora que expresa un plásmido de la reivindicación 2.
4. La célula hospedadora de la reivindicación 3, seleccionada del grupo constituido por células eucariotas y células bacterianas.
5. La célula hospedadora de la reivindicación 4 seleccionada del grupo constituido por *S. cerevisiae* PRT50, y cepas diploides que resultan del apareamiento de las cepas SKY191 y PRT50 con las cepas homólogas adecuadas.
6. La célula hospedadora de la reivindicación 4, que es *E. coli* KJ1567 o *E. coli* AG58A (RP28).
7. Un procedimiento para comparar las interacciones de unión entre una primera proteína y una segunda proteína tanto en entornos de organismos bacterianos como de levaduras utilizando una construcción que funciona en ambos organismos, que comprende:
- el suministro de células hospedadoras de levaduras y bacterianas, comprendiendo cada una,
 - un gen informador unido de manera operable a una secuencia de ADN que comprende un sitio de unión a proteínas;
 - un primer gen de fusión que expresa una primera proteína de fusión, dicha primera proteína de fusión comprendiendo dicha primera proteína unida covalentemente a un resto de unión que es capaz de unirse específicamente a dicho sitio de unión a proteínas que está dirigido por una secuencia promotora entre el nucleótido en posición 4 y el nucleótido en posición 545 de la Figura 1C, y
 - un segundo gen de fusión que expresa una segunda proteína de fusión, dicha segunda proteína de fusión comprendiendo dicha segunda proteína unida covalentemente al resto de activación génico;
 - permitir que dichas primera y segunda proteínas interactúen; y
 - medir la expresión de dicho gen informador como medida de dicha interacción entre dichas primera y segunda proteínas tanto en *E. coli* como en levaduras.
8. Un sistema bihíbrido unificado de levaduras y bacterias para la puesta en práctica del procedimiento de la reivindicación 7, que comprende:
- un plásmido seleccionado del grupo constituido por la SEQ ID NO: 2 (pGLS20), SEQ ID NO: 3 (pGLS22) y pGLS23 en el que pGLS23 es idéntico a pGLS22 excepto por un cambio en la posición 3267 de gaattC a gaattA;
 - al menos un plásmido presa seleccionado del grupo constituido por pBR-AMP- α LPL (SEQ ID NO: 5) y pAC-AMP- α LPL;
 - una cepa de *E. coli* seleccionada del grupo constituido por *E. coli* KJ1567 o *E. coli* AG58A (RP28); y
 - una cepa de levaduras seleccionada del grupo constituido por PRT50, SKY191 y PRT475.

Figura 1

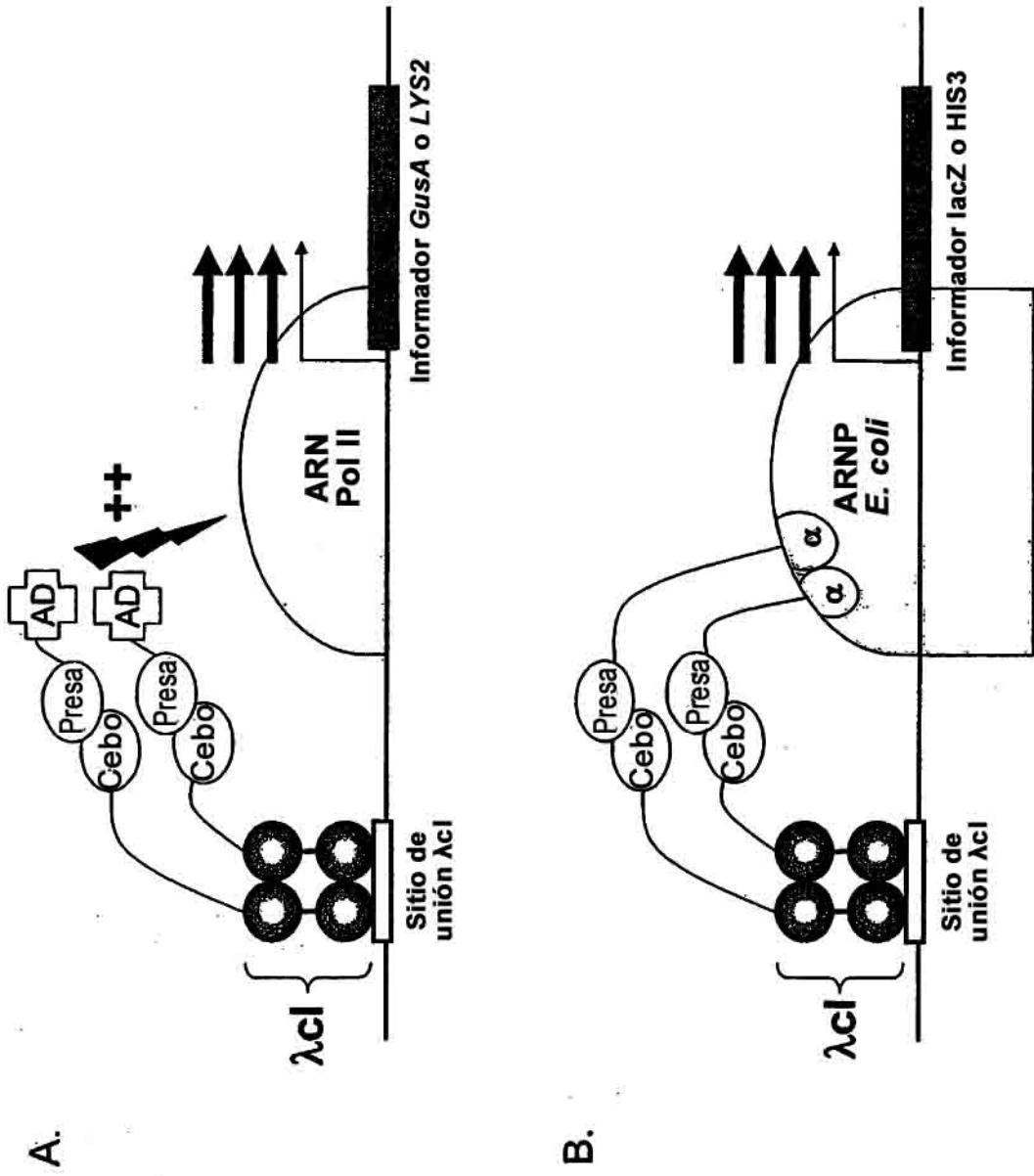


Figura 1C

LOCUS	PGLS20 .GB	5296 BP DS-DNA	CIRCULAR	SYN	15-JAN-1998	
DEFINICIÓN -						
Nº DE ACCESO -						
PALABRAS CLAVE-						
FUENTE -						
CARACTERÍSTICAS	Localización/Clasificación					
CDS	join(543..>1252,<1254..1256)					
	/note="cI gene [Split]"					
terminador	1336..1528					
	/note="ADH Terminator"					
origen de replicación	1663..2820					
	/note="2µ origin"					
fragmento	3078..4280					
	/note="71 to 1273 of pFA6a-kanMX4"					
fragmento	3078..4280					
	/note="1 to 1203 of Kan/G418 (Bgl/Sca)-kanMX4 (C++ cut with BamH1/StuI)"					
promotor	3122..3465					
	/note="TEF promoter (AG)"					
CDS	3466..4275					
	/note="Kan R gene from Tn903"					
terminador	4298..4570					
	/note="cytC terminator"					
origen de replicación	4605..4874					
	/note="ColE1 ori"					
fragmento	1..542					
	/note="4 to 545 of TEF + UVlac promoter"					
promotor	10..422					
	/note="TEF promoter from S.cerevisiae see M15666."					
fragmento	467..542					
	/note="1555 to 1630 of pBT.mv"					
promotor	468..526					
	/note="lacUV5-lacO"					
RECUENTO BASAL	1414 A	1177 C	1160 G	1545 T	0 OTHER	
ORIGEN						
1	TCATGAGGGG	ATCCCCACA	CACCATAGCT	TCAAATGTT	TCTACTCCTT	TTTTACTCTT
61	CCAGATTTTC	TCGGACTCCG	CGCATCGCG	TACCACCTCA	AAACACCCAA	GCACAGCATA
121	CTAAATTTCC	CCTCTTTCTT	CCTCTAGGGT	GTCGTTAATT	ACCCGTACTA	AAGGTTTGGG
181	AAAGAAAAAA	GAGACCGCCT	CGTTTCTTTT	TCTTCGTCGA	AAAAGGCAAT	AAAAATTTTT
241	ATCACGTTTC	TTTTTCTTGA	AAATTTTTTT	TTTGATTTTT	TTCTCTTTTCG	ATGACCTCCC
301	ATTGATATTT	AAGTTAATAA	ACGGTCTTCA	ATTTCTCAAG	TTTCAGTTTC	ATTTTCTTGT
361	TTCTATTACA	ACTTTTTTTA	CTTCTTGCTC	ATTAGAAAGA	AAGCATAGCA	ATCTAATCTA
421	AGGGCGGTGT	TGACAATTA	TCATCGGCAT	AGTATATCGG	CCTAGGCTTT	ACACTTTATG
481	CTTCCGGCTC	GTATAATGTG	TGGAATGTG	AGCGGATAAC	AATTTACAC	AGGAAACAGC
541	GTATGAGCAC	AAAAAAGAAA	CCATTAACAC	AAGAGCAGCT	TGAGGACGCA	CGTCGCCTTA
601	AAGCAATTTA	TGAAAAAAG	AAAAATGAAC	TTGGCTTATC	CCAGGAATCT	GTCGCAGACA
661	AGATGGGGAT	GGGGCAGTCA	GGCGTTGGTG	CTTTATTTAA	TGGCATCAAT	GCATTAATG
721	CTTATAACGC	CGCATTGCTT	GCAAAAATTC	TCAAAGTTAG	CGTTGAAGAA	TTTAGCCCTT
781	CAATCGCCAG	AGAAATCTAC	GAGATGTATG	AAGCGGTTAG	TATGCAGCCG	TCACTTAGAA
841	GTGAGTATGA	GTACCCTGTT	TTTTCTCATG	TTCAGGCAGG	GATGTTCTCA	CCTGAGCTTA
901	GAACCTTTAC	CAAAGGTGAT	GCGGAGAGAT	GGGTAAGCAC	AACCAAAAAA	GCCAGTGATT
961	CTGCATTCTG	GCTTGAGGTT	GAAGGTAATT	CCATGACCGC	ACCAACAGGC	TCCAAGCCAA
1021	GCTTTCCTGA	CGGAATGTTA	ATTCTCGTTG	ACCCTGAGCA	GGCTGTTGAG	CCAGGTGATT
1081	TCTGCATAGC	CAGACTTGGG	GGTGATGAGT	TTACCTTCAA	GAAACTGATC	AGGGATAGCG
1141	GTCAGGTGTT	TTTACAACCA	CTAAACCCAC	AGTACCCCAAT	GATCCCATGC	AATGAGAGTT
1201	GTTCCGTTGT	GGGGAAAGTT	ATCGCTAGTC	AGTGGCCTGA	AGAGACGTTT	GGCGAATTCA

Figura 1D

1261 AGCTTGAGCT CAGATCTCAG CTGGGCCCGG TACCGCGGCC GCTCGAGTCG ACCTGCAGCC
1321 AAGCTAATTC CGGGCGAATT TCTTATGATT TATGATTTTT ATTATTAAAT AAGTTATAAA
1381 AAAAATAAGT GTATACAAAT TTTAAAGTGA CTCTTAGGTT TTA AACGAA AATTCCTGTT
1441 CTTGAGTAAC TCTTTCCTGT AGGTCAGGTT GCTTCTCAG GTATAGCATG AGGTCGCTCT
1501 TATTGACCAC ACCTCTACCG GCATGCCGAG CAAATGCCTG CAAATCGCTC CCCATTTTAC
1561 CCAATTGTAG ATATGCTAAC TCCAGCAATG AGTTGATGAA TCTCGGTGTG TATTTTATGT
1621 CCTCAGAGGA CAATACCTGT TGTAATCCGT CCCAAGCTAA CGAAGCATCT GTGCTTCATT
1681 TTGTAGAACA AAAATGCAAC GCGAGAGCGC TAATTTTTCA AACAAAGAAT CTGAGCTGCA
1741 TTTTACAGA ACAGAAATGC AACGCGAAAG CGCTATTTTA CCAACGAAGA ATCTGTGCTT
1801 CATTTTGTGA AAACAAAAT GCAACGCGAG AGCGCTAATT TTTCAAACAA AGAATCTGAG
1861 CTGCATTTTT ACAGAACAGA AATGCAACGC GAGAGCGCTA TTTTACCAAC AAAGAATCTA
1921 TACTTCTTTT TTGTTCTACA AAAATGCATC CCGAGAGCGC TATTTTCTA ACAAAGCATC
1981 TTAGATTACT TTTTTTCTCC TTTGTGCGCT CTATAATGCA GTCTCTTGAT AACTTTTTGC
2041 ACTGTAGGTC CGTTAAGGTT AGAAGAAGGC TACTTTGGTG TCTATTTTCT CTTCCATAAA
2101 AAAAGCCTGA CTCCACTTCC CGCGTTTACT GATTACTAGC GAAGCTGCGG GTGCATTTTT
2161 TCAAGATAAA GGCATCCCGG ATTATATTCT ATACCGATGT GGATTGCGCA GATTGCTGA
2221 ACAGAAAGT ATAGCGTTGA TGATTCTTCA TTGGTCAGAA AATTATGAAC GGTTCCTTCT
2281 ATTTGTCTC TATATACTAC GTATAGGAAA TGTTTACATT TTCGTATTGT TTTTCGATTCA
2341 CTCTATGAAT AGTCTTACT ACAATTTTTT TGCTAAAGA GTAATACTAG AGATAAACAT
2401 AAAAAATGTA GAGGTCGAGT TTAGATGCAA GTTCAAGGAG CGAAAGGTGG ATGGGTAGGT
2461 TATATAGGGA TATAGCACAG AGATATATAG CAAAGAGATA CTTTTGAGCA ATGTTTGTGG
2521 AAGCGGTATT CGCAATATTT TAGTAGCTCG TTACAGTCCG GTGCGTTTTT GGTTTTTTGA
2581 AAGTGCGTCT TCAGAGCGCT TTTGGTTTTT AAAAGCGCTC TGAAGTCCCT ATACTTTCTA
2641 GCTAGAGAAT AGGAACTTCG GAATAGGAAC TTCAAAGCGT TTCCGAAAC GAGCGCTTCC
2701 GAAAAATGCAA CGCGAGCTGC GCACATACAG CTCACTGTTT ACCTGCGACC TATATCTGCG
2761 TGTTGCCTGT ATATATATAT ACATGAGAAG AACGGCATAG TGCGTGTTTA TGCTTAAATG
2821 CGTTATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC CAGCCCCGAC
2881 ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGCTT GCTCCGGCA TCCGCTTACA
2941 GACAAGCTGT GACCGTCTCC GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA
3001 AACGCGCGAG ACGAAAGGGC CTCGTGATAC GCCTATTTTT ATAGGTTAAT GTCATGATAA
3061 TAATGGTTTC TTAGGGGGAT CTGTTTAGCT TGCTCGTCC CCGCGGGC GAGCGCTTCC
3121 CGACATGGAG GCCCAGAATA CCTCCTTGA CAGTCTTGAC GTGCGCAGCT CAGGGGCATG
3181 ATGTGACTGT CGCCCGTACA TTTAGCCCAT ACATCCCAT GTATAATCAT TTGCATCCAT
3241 ACATTTTATGAT GGCCGCACGG CGCGAAGCAA AAATTACGGC TCCTCGCTGC AGACCTGCGA
3301 GCAGGAAAC GCTCCCTCA CAGACCGGTT GAATTGTCCC CACGCGCGC CCCTGTAGAG
3361 AAATATAAAA GGTTAGGATT TGCCACTGAG GTTCTTCTTT CATATACTTC CTTTTAAAT
3421 CTTGCTAGGA TACAGTCTC ACATCACATC CGAACATAAA CAACCATGGG TAAGGAAAAG
3481 AATCAGTTTT CGAGGCGGCG ATTAATTTCC AACATGGATG CTGATTTATA TGGGTATAAA
3541 TGGGCTCGCG ATAATGTCGG GCAATCAGGT GCGACAATCT ATCGATTGTA TGGGAAGCCC
3601 GATGCGCCAG AGTTGTTTCT GAAACATGGC AAAGGTAGCG TTGCCAATGA TGTTACAGAT
3661 GAGATGGTCA GACTAAACTG GCTGACGGAA TTTATGCCTC TTCCGACCAT CAAGCATTTT
3721 ATCCGFACTC CTGATGATGC ATGGTTACTC ACCACTGCGA TCCCGGCAA AACAGCATTC
3781 CAGGTATTAG AAGAATATCC TGATTCAGGT GAAAATATTG TTGATGCGCT GGCAGTGTTC
3841 CTGCGCCGGT TGCATTGAT TCCGTGTTGT AATTGTCCTT TTAACAGCGA TCGCGTATTT
3901 CGTCTCGCTC AGGCGCAATC ACGAATGAAT AACGGTTTGG TTGATGCGAG TGATTTTATG
3961 GACGAGCGTA ATGGCTGGCC TGTTGAACAA GTCTGGAAAG AAATGCATAA GCTTTTGCCA
4021 TTCTACCGG ATTTCAGTCGT CACTCATGGT GATTTCTCAC TTGATAACCT TATTTTTGAC
4081 GAGGGGAAAT TAATAGGTTG TATTGATGTT GGACGAGTCG GAATCGCAGA CCGATACCAG
4141 GATCTTGCCA TCCTATGGAA CTGCCTCGGT GAGTTTTCTC CTTTATTACA GAAACGGCTT
4201 TTTCAAATAT ATGGTATTGA TAATCCTGAT ATGAATAAAT TGAGTTTCA TTTGATGCTC
4261 GATGAGTTTT TCTAATCAGT CCTCGGAGAT CCGTCCCTT TTTCTTTGT CGATATCATG
4321 TAATTAGTTA TGTCACGCTT ACATTCACGC CCTCCCCCA CATCCGCTCT AACCGAAAAG
4381 GAAGGAGTTA GACAACCTGA AGTCTAGGTC CCTATTTATT TTTTATAGT TATGTTAGTA
4441 TTAAGAACGT TATTTATATT TCAAATTTTT CTTTTTTTTC TGTACAGACG CGGTACGCA
4501 TGTAACATTA TACTGAAAC CTTGCTTGAG AAGGTTTTGG GACGCTCGAA GGCTTTAATT

Figura 1E

4561 TGCAAGCTGG AGACCAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG
 4621 GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCTG ACGAGCATCA CAAAAATCGA
 4681 CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGAATAAAA GATACCAGGC GTTCCCCCT
 4741 GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC
 4801 TTTCTCCCTT CGGGAAGCGT GGCGCTTTCT CAATGCTCAC GCTGTAGGTA TCTCAGTTCG
 4861 GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTCA GCCCGACCGC
 4921 TGCGCCTTAT CCGGTAAC TA TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA
 4981 CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG
 5041 TTCTTGAAGT GGTGGCCTAA CTACGGCTAC .ACTAGAAGGA CAGTATTTGG TATCTGCGCT
 5101 CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC
 5161 ACCGCTGGTA GCGGTGGTTT TTTTGTGTTGC AAGCAGCAGA TTACGCGCAG AAAAAAAGGA
 5221 TCTCAAGAAG ATCCTTTGAT CTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA
 5281 CGTTAAGGGA TTTTGG

//

Figura 1F

LOCUS PGLS20HC 6187 BP DS-DNA CIRCULAR SYN 15-JAN-1998

DEFINICIÓN -

Nº DE ACCESO -

PALABRAS CLAVE -

FUENTE -

CARACTERÍSTICAS

CDS

Localización/Clasificación

join(543..1252,<1259..1258)

/note="cI gene [Split]"

terminador

1336..1528

/note="ADH Terminator"

origen de replicación

1663..2820

/note="2µ origin"

origen de replicación

5496..5765

/note="ColE1 ori"

fragmento

1..542

/note="4 to 545 of TEF + UVlac promoter"

promotor

10..422

/note="TEF promoter"

fragmento

467..542

/note="1555 to 1630 of pBT.mv"

promotor

468..526

/note="lacUV5-lacO"

fragmento

2834..5472

/note="1 to 2639 of Untitled1"

fragmento

complement(>2834..5472)

/note="67 to 2706 of CamR-HIS5 cassette"

CDS

complement(3718..4884)

/note="HIS5 yeast"

promotor

complement(4885..5470)

/note="HIS5 prom"

fragmento

complement(>2834..3710)

/note="95 to 1017 of CamR cassette [Split]"

CDS

2941..3708

/note="CamR"

promotor

2835..2940

/note="Cam R promoter"

RECUENTO BASAL 1773 A 1297 C 1281 G 1836 T 0 OTHER

ORIGEN

```

1 TCATGAGGGG ATCCCCACA CACCATAGCT TCAAAATGTT TCTACTCCTT TTTTACTCTT
61 CCAGATTTTC TCGGACTCCG CGCATCGCCG TACCACTTCA AAACACCCAA GCACAGCATA
121 CTAATTTTCC CCTCTTTCTT CCTCTAGGGT GTCGTTAATT ACCCGTACTA AAGGTTGGA
181 AAAGAAAAAA GAGACCGCCT CGTTTCTTTT TCTTCGTCGA AAAAGGCAAT AAAAATTTTT
241 ATCACGTTTC TTTTCTTGA AAATTTTTTT TTTGATTTTT TTCTCTTTCG ATGACCTCCC
301 ATTGATATTT AAGTTAATAA ACGGTCTTCA ATTTCTCAAG TTTCAGTTTC ATTTTTCTTG
361 TTCTATTACA ACTTTTTTTA CTTCTTGCTC ATTAGAAAGA AAGCATAGCA ATCTAATCTA
421 AGGGCGGTGT TGACAATTAA TCATCGGCAT AGTATATCGG CCTAGGCTTT AACTTTATG
481 CTCCCGGCTC GTATAATGTG TGAATTTGTG AGCGGATAAC AATTTACAC AGGAAACAGC
541 GTATGAGCAC AAAAAAGAAA CCATTAACAC AAGAGCAGCT TGAGGACGCA CGTCGCCTTA
601 AAGCAATTTA TGAATAAAG AAAAATGAAC TTGGCTTATC CCAGGAATCT GTCGCAGACA
661 AGATGGGGAT GGGGCAGTCA GCGTTGGTG CTTTATTTAA TGGCATCAAT GCATTAAATG
721 CTTATAACGC CGCATTGCTT GCAAAAATTC TCAAAGTTAG CGTTGAAGAA TTTAGCCCTT
781 CAATCGCCAG AGAAATCTAC GAGATGTATG AAGCGGTTAG TATGCAGCCG TCACTTAGAA
841 GTGAGTATGA GTACCCTGTT TTTTCTCATG TTCAGGCAGG GATGTTCTCA CCTGAGCTTA
901 GAACCTTTAC CAAAGGTGAT GCGGAGAGAT GGGTAAGCAC AACCAAAAAA GCCAGTGATT
961 CTGCATTCTG GCTTGAGGTT GAAGGTAATT CCATGACCGC ACCAACAGGC TCCAAGCCAA
1021 GCTTTCCTGA CGGAATGTTA ATTCTCGTTG ACCCTGAGCA GGCTGTTGAG CCAGGTGATT
    
```

Figure 1G

1081 TCTGCATAGC CAGACTTGGG GGTGATGAGT TTACCTTCAA GAAACTGATC AGGGATAGCG
 1141 GTCAGGTGTT TTTACAACCA CTAAACCCAC AGTACCCAAT GATCCCATGC AATGAGAGTT
 1201 GTTCCGTTGT GGGGAAAGTT ATCGCTAGTC AGTGGCCTGA AGAGACGTTT GGCGAATTCA
 1261 AGCTTGAGCT CAGATCTCAG CTGGGCCCGG TACCGCGGCC GCTCGAGTCG ACCTGCAGCC
 1321 AAGCTAATTC CGGGGAATTT TCTTATGATT TATGATTTTT ATTATTAAAT AAGTTATAAA
 1381 AAAAATAAGT GTATACAAAT TTTAAAGTGA CTCTTAGGTT TTA AACGAA AATTC TTGTT
 1441 CTTGAGTAAC TCTTCTCTGT AGGTCAGGTT GCTTCTCAG GTATAGCATG AGGTCGCTCT
 1501 TATTGACCAC ACCTCTACCG GCATGCCGAG CAAATGCCTG CAAATCGCTC CCCATTTAC
 1561 CCAATGTAG ATATGCTAAC TCCAGCAATG AGTTGATGAA TCTCGGTGTG TATTTTATGT
 1621 CCTCAGAGGA CAATACCTGT TGTAAATCCG CCCAAGCTAA CGAAGCATCT GTGCTTCATT
 1681 TTGTAGAACA AAAATGCAAC GCGAGAGCGC TAATTTTTCA AACAAAGAAT CTGAGCTGCA
 1741 TTTTACAGAC ACAGAAATGC AACGCGAAAG CGCTATTTTA CCAACGAAAG ATCTGTGCTT
 1801 CATTTTTGTA AAACAAAAAT GCAACGCGAG AGCGCTAATT TTTCAAACA AGAATCTGAG
 1861 CTGCATTTTT ACAGAACAGA AATGCAACGC GAGAGCGCTA TTTTACCAAC AAAGAATCTA
 1921 TACTTCTTTT TTGTTCTACA AAAATGCATC CCGAGAGCGC TATTTTCTA ACAAAGCATC
 1981 TTAGATTACT TTTTTCTCC TTTGTGCGCT CTATAATGCA GTCTCTTGAT AACTTTTTGC
 2041 ACTGTAGGTC CGTTAAGGTT AGAAGAAGGC TACTTTGGTG TCTATTTCT CTCCATAAAA
 2101 AAAAGCCTGA CTCCACTTCC CGCGTTTACT GATTACTAGC GAAGCTGCGG TACGATTTTT
 2161 TCAAGATAAA GGCATCCCGG ATTATATTCT ATACCGATGT GGATTGCGCA CTCTTTGTTGA
 2221 ACAGAAAGTG ATAGCGTTGA TGATTCTTCA TTGGTCAGAA AATTATGAAC GGTTCCTTCT
 2281 ATTTTGTCTC TATATACTAC GTATAGGAAA TGTTTACATT TTCGTATTGT TTTGATTCA
 2341 CTCTATGAAT AGTTCCTACT ACAATTTTTT TGCTTAAAGA GTAATACTAG AGATAAACAT
 2401 AAAAATGTA GAGGTCGAGT TTAGATGCAA GTTCAAGGAG CGAAAGGTGG ATGGGTAGGT
 2461 TATATAGGGA TATAGCACAG AGATATATAG CAAAGAGATA CTTTGTAGCA ATGTTTGTGG
 2521 AAGCGGTATT CGCAATATTT TAGTAGCTCG TTACAGTCCG GTGCGTTTTT GGTTTTTGA
 2581 AAGTGCGTCT TCAGAGCGCT TTTGGTTTTC AAAAGCGCTC TGAAGTTCT ATACTTTCTA
 2641 GCTAGAGAA AGGAACTTCG GAATAGAAC TTCAAAGCGT TTCGAAAAC GAGCGCTTCC
 2701 GAAAATGCAA CGCGAGCTGC GCACATACAG CTCACTGTTC ACGTCGCACC TATATCTGCG
 2761 TGTTGCTGT ATATATATAT ACATGAGAAG AACGGCATAG TGCGTGTTTA TGCTTAAATG
 2821 CGTTATGGTG CACTTTGCGC CGAATAAATA CCTGTGACGG AAGATCACTT CGCAGAATAA
 2881 ATAAATCCTG GTGTCCTGT TGATACGGG AAGCCCTGGG CCAACTTTTG GCGAAAATGA
 2941 GACGTTGATC GGCACGTAAG AGGTTCCAAC TTTCAACATA ATGAAATAAG ATCACTACCG
 3001 GCGGTATTTT TTGAGTTATC GAGATTTTCA GGAGCTAAGG AAGCTAAAT GGAGAAAAA
 3061 ATCACTGGAT ATACCACCGT TGATATATCC CAATGGCATC GTAAAGAACA TTTGTAGGCA
 3121 TTTCAGTCAG TTGCTCAATG TACCTATAAC CAGACCGTTC AGCTGGATAT TACGGCCTTT
 3181 TTAAGACCG TAAAGAAAAA TAAGCACAAG TTTTATCCGG CCTTTATTCA CATTCTTGCC
 3241 CGCTGATGA ATGCTCATCC GGAATCCGT ATGGCAATGA AAGACGCTGA CTGGTGATA
 3301 TGGGATAGTG TTCACCTTG TTACACCGTT TTCCATGAGC AAACGAAAC GTTTTCATCG
 3361 CTCTGGAGTG AATACCACGA CGATTTCCGG CAGTTCTAC ACATATATTC GCAAGATGTG
 3421 GCGTGTACG GTGAAAACCT GGCCTATTTT CCTAAAGGGT TTATTGAGAA TATGTTTTTCT
 3481 GTCTAGCCA ATCCCTGGGT GAGTTTACC AGTTTTGATT TAAACGTGGC CAATATGGAC
 3541 AACTTCTTCG CCCCCGTTTT CACCATGGGC AAATATTATA CGCAAGGCGA CAAGGTGCTG
 3601 ATGCCGCTGG CGATTCAGGT TCATCATGCC GTTTGTGATG GCTTCCATGT CGGCAGAATG
 3661 CTTAATGAAT TACAACAGTA CTGCGATGAG TGGCAGGGCG GGGCGTAAGA CGTCTATTTA
 3721 TTCATTGGCC AGCTTATATA ACGTCTCTT GAAGTACTTT ATCAAATGTG TGTTCTCTC
 3781 ATGGGTCCA ACGGTAATTC TCAAACATCC GGAACAGCCT AATTCGTTAC CTCTAAATCT
 3841 GACGACAACC CCAGATTGAG TAGCCAATTG GTAGTATAAC TTCTTTGCCA AGACATTGTC
 3901 ACCCCGTTG ATCCGTATTA AAAGAAAATT AGCATCTAAT CCACCAACAT ATTGGTCATC
 3961 AACGTAATCC AAAGCAGTTA ATTCCTTTAA GAGGCGCATT TTCTCTTCAT TGATTATTTT
 4021 CGAAGTGGCT TCCATCTTCT TTAGATTACT GTCCTGAACA GCTTTTAGTG CATATTCAGA
 4081 GGCTAGGGAG GAAATATTAT AAGGCGCCTT CATTGCATTT AAAATCTGG CCAACTCTGC
 4141 TGTTGCATAT GTCATACCCA ACCTAATCCC GGCTAAACCG AATGACTTGG ATAGAGTTTG
 4201 CAAAGTAACC AAGTTAGGAT ACTTGGTGAC TAGTGGAGCT GTAGAGCCAC CACAAAAATC
 4261 TACGTAAGCT TCATCAACAA CACTAACC ATTTGCCAA TTCTGTAAGA CCTTTTCGAT
 4321 TAAACTGGTC TTAATTTTGG CTCCTGTTGG ATTACCTGGT GAAGTAACGA ACATCAACTT
 4381 AATTAGCGGG TCGTTTTTCA AAATGGTTAA TACAGCTTCG GTATCCATTT GAAAAGAACC

Figure 1H

```

4441 GTCGGAAACA GTTAAAGGAC ATTGGACGAC TTCTATATCA TTAATGTTTG CACAAACAGA
4501 ATACATAGAA TATGTTGGTG GAAGAACCAG AATCTTTTCT TTCCCGGGAA CACAGCATGC
4561 TCTAATAATA GCATCAATAC TCTCATCAGA TCCCACACCT AGGCACAGAT TGTCAGCAGT
4621 TAAAGGTTTT ACCTCTGGGT CATTGGCATA ACTGCTTGTT TTGTTCTGTG ATTTTCGTCAT
4681 TGCGGTCTTG AACTCCAATT GGTGAGGATC TGGGTAACGA TGTA AATTGG TCTTGCTCAA
4741 TTCAACTGGA GTAGGTCCAT GGGCGTTTTC ATTGGCGTCT AGCAATATAC CCTCGGTGAA
4801 ATCATCTCTT GCACAGCGAT AAGGTTCCAA GTTATAAATT TTTGGTCTAA CAATTCTTTT
4861 CAAATCAAAA ACCATAGTGT AATTTTAATA TATACGACAC ACACGTCCTG CTGGTTTATC
4921 AGGAAACAAA ATAAGAGTAG TCAATGGAAA AACTGTTTTT ACATATTAGA GGGTATATTA
4981 AACAGAACTG TGTGCATCCT TTTCAAGTTA TATAACGTGA GAGATAAAAT ATCAAGTATG
5041 TCATGTCAGG GTAAGAAAACA TCAATTGAAG TGAGTCAACA GATCCAAGAA AAAAAAGCAC
5101 TAACTACGTC ACTACACCAT GAACTATTGA AAATTGGTAG TTTAGTCATC TCAGATTCCA
5161 TTCATTGGAA AAAACAATTG ATTCATAACA ATTA AACTTCG GATTAGTCAT TAATTATTTT
5221 AATGCTTGAC TCCTTTTTGA ATAGTATCAC CCGGATCGTG GTCACATGAT CAAATAAATT
5281 ATTGCATTAC CAATGGCTTC TGTATTAGTT ACTCTCCAGG AAATGTCTCA ACATAACCGG
5341 TCACCATATT TATGATAACA ATTTTTAACC ATTTACCCTT TATTTTTGCA AAGTTATGAC
5401 CTTTGG AATG CAGCAGAAGA AAAAAATTGA TGAAGTAGTC ATCAAACAGG TTTGGCGGAA
5461 AGACAGATCA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGPTG
5521 CTGGCGTTTT TCCATAGGCT CCGCCCCCTT GACGAGCATC ACAAAAATCG ACGCTCAAGT
5581 CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGG AAGCTCC
5641 CTCGTGCGCT CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
5701 TCGGGAAGCG TGGCGCTTTC TCAATGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC
5761 GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCGTTT AGCCCGACCG CTGCGCCTTA
5821 TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
5881 GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG
5941 TGGTGGCCTA ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC TCTGCTGAAG
6001 CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT
6061 AGCGGTGGTT TTTTTGTTTG CAAGCAGCAG ATTACGCGCA GAAAAAAGG ATCTCAAGAA
6121 GATCCTTTGA TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAA AACTC ACGTTAAGGG
6181 ATTTTGG
//

```

FIGURA 1I

LOCUS	pBR-AMP- α LPL	4320	BP	DS-DNA	CIRCULAR	SYN	18-DEC-2002
DEFINICIÓN	-						
Nº DE ACCESO	-						
PALABRAS CLAVE	-						
RECUENTO BASAL	1088	A	1061	C	1085	G	1086
ORIGEN	-						O OTHER
1	GAATTCGAAC	CCCTCCTACG	ACGTCTAAGA	AACCATTATT	ATCATGACAT	TAACCTATAA	
61	AAATAGCGGT	ATCACGAGGC	CCTTTGGATA	ACCAGAAGCA	ATAAAAAATC	AAATCGGATT	
121	TCACTATATA	ATCTCACTTT	ATCTAAGATG	AATCCGATGG	AAGCATCCTG	TTTTCTCTCA	
181	ATTTTTTTTAT	CTAAAACCCA	GCGTTCGATG	CTTCTTTGAG	CGAACGATCA	AAAATAAGTG	
241	CCTTCCCATC	AAAAAATAT	TGACAACATA	AAAAACTTTG	TGTTATACTT	GTAACGCTAC	
301	ATGGAGATTA	ACTCAATCTA	GCTAGAGAGG	CTTTACACTT	TATGCTTCCG	GCTCGTATAA	
361	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTT	ACACAGGAAA	CAGCTATGAC	CATGATTACG	
421	GATTCACTGG	AACTCTAGAC	CAAAGAGAGG	ACACAATGCA	GGGTTCTGTG	ACAGAGTTTC	
481	TAAAACCGCG	CCTGGTTGAT	ATCGAGCAAG	TGAGTTCGAC	GCACGCCAAG	GTGACCCTTG	
541	AGCCTTTAGA	GCGTGGCTTT	GGCCATACTC	TGGGTAACGC	ACTGCCCCGT	ATTCTGCCTC	
601	CATCGATGCC	GGGTTGCGCG	GTGACCGAGG	TTGAGATTGA	TGGTGTACTA	CATGAGTACA	
661	GCACCAAAGA	AGGCGTTCAG	GAAGATATCC	TGAAAATCCT	GCTCAACCTG	AAAGGGCTGG	
721	CGGTGAGAGT	TCAGGGCAAA	GATGAAGTTA	TTCCTACCTT	GAATAAATCT	GGCATTGGCC	
781	CTGTGACTGC	AGCCGATATC	ACCCACGACG	GTGATGTCGA	AATCGTCAAG	CCGCAGCAGC	
841	TGATCTGCCA	CCTGACCGAT	GAGAACGCGT	CTATTAGCAT	CGGTATCAAA	GTTACGCGCG	
901	GTCGTGGTTA	TGTGCCGGCT	TCTACCCGAA	TTCAATTCGGA	AGAAGATGAG	CGCCCAATCG	
961	GCCGTCTGCT	GGTCGACGCA	TGCTACAGCC	CTGTGGAGCG	TATTGCCCTAC	AATGTTGAAG	
1021	CAGCGCGTGT	AGAACAGCGT	ACCGACCTGG	ACAAGCTGGT	CATCGAAATG	GAAACCAACG	
1081	GCACAATCGA	TCCTGAAGAG	GCGATTTCGT	GTGCGGCAAC	CATTCTGGCT	GAACAACCTG	
1141	AAGCTTTTCGT	TGACTTACGT	GATGTACGTC	AGCCTGAAGT	GAAAGAAGAG	AAACCAGAGG	
1201	CGGCCGCGCA	ATTGGAGCTC	CTCGAGGGAT	CCTAAGTAAG	AAGACACAGG	CGAGAGCCGC	
1261	TAGTCTAGAG	ACTAGAAAAA	GGCCGACAAG	TCCCCTCCG	CTGAAGATCC	TGGCGTAATA	
1321	GCGAAGAGGC	CCGCACCGAT	CGCCCTTCCC	AACAGTTGCG	CAGCCTGAAT	GGCGAATGGG	
1381	ACGCGCCCTG	TAGCGGGCGA	TTAAGCGCGG	CGGGTGTGGT	GGTTACGCGC	AGCGTGACCG	
1441	CTACACTTGC	CAGCGCCCTA	GCGCCCGCTC	CTTTCGCTTT	CTTCCCTTCC	TTTCTCGCCA	
1501	CGTTCGCCGG	CTTTCCCGGT	CAAGCTCTAA	ATCGGGGGCT	CCCTTTAGGG	TTCCGATTTA	
1561	GTGCTTTTACG	GCACCTCGAC	CCCCAAAAAC	TTGATTAGGG	TGATGGTTCA	CGTAGTGGGC	
1621	CATCGCCCTG	ATAGACGGTT	TTTCGCCCTT	TGACGTTGGA	GTCCACGTTT	TTTAATAGTG	
1681	GACTCTTGTT	CCAAACTGGA	ACAACACTCA	ACCCATCTC	GGTCTATTCT	TTTGATTTAT	
1741	AAGGGATTTT	GCCGATTTCG	GCCTATTGGT	TAAAAAATGA	GCTGATTTAA	CAAAAATTTA	
1801	ACGCGAATTT	TAAACAAAATA	TTAACGCTTA	CAATTTAGGT	GGCACTTTTC	GGGGAAATGT	
1861	GCGCGGAACC	CCTATTTGTT	TATTTTTCTA	AATACATTCA	AATATGTATC	CGCTCATGAG	
1921	ACAATAACCC	TGATAAATGC	TCAATAATA	TTGAAAAAGG	AAGAGTATGA	GTATTCAACA	
1981	TTTCCGTGTC	GCCCTTATTC	CCTTTTTTGC	GGCATTTTGC	CTTCTGTTT	TTGCTCACCC	
2041	AGAAAACGCTG	GTGAAAGTAA	AAGATGCTGA	AGATCAGTTG	GGTGCACGAG	TGGGTTACAT	
2101	CGAACTGGAT	CTCAACAGCG	GTAAGATCCT	TGAGAGTTT	CGCCCGAAG	AACGTTTTTC	
2161	AATGATGAGC	ACTTTTAAAG	TTCTGCTATG	TGGCGGGTA	TTATCCCGTG	TTGACGCCGG	
2221	GCAAGAGCAA	CTCGGTCCGC	GCATACACTA	TTCTCAGAAT	GACTTGGTTG	AGTACTCACC	
2281	AGTCACAGAA	AAGCATCTTA	CGGATGGCAT	GACAGTAAGA	GAATTATGCA	GTGCTGCCAT	
2341	AACCATGAGT	GATAACACTG	CGGCCAACTT	ACTTCTGACA	ACGATCGGAG	GACCGAAGGA	
2401	GCTAACCGCT	TTTTTGCACA	ACATGGGGGA	TCATGTAAC	CGCCTTGATC	GTTGGGAACC	
2461	GGAGCTGAAT	GAAGCCATAC	CAAACGACGA	GCGTGACACC	ACGATGCCTG	CAGCAATGGC	
2521	AACAACGTTG	CGCAAACAT	TAACTGGCGA	ACTACTTACT	CTAGCTTCCC	GGCAACAATT	
2581	AATAGACTGG	ATGGAGGCGG	ATAAAGTTGC	AGGACCCTT	CTGCGCTCGG	CCCTTCCGGC	

Figura 1J

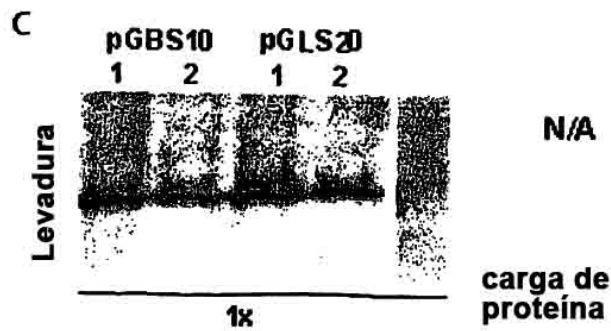
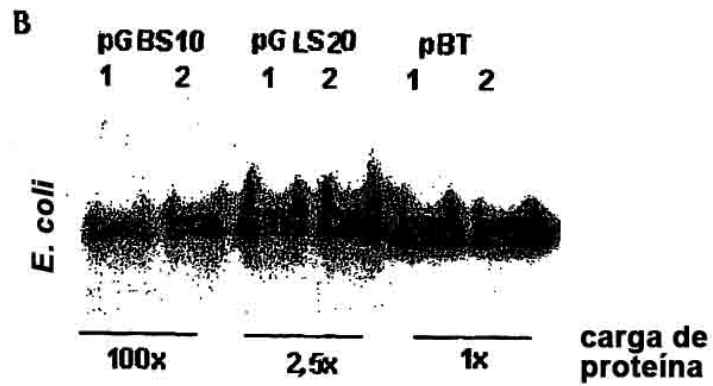
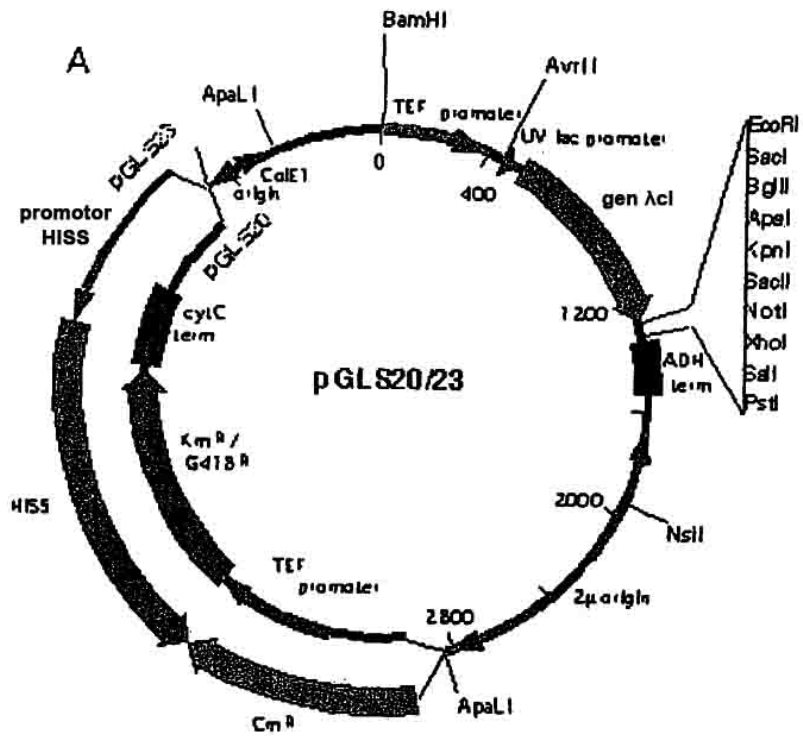
```

2641 TGGCTGGTTT ATTGCTGATA AATCTGGAGC CGGTGAGCGT GGGTCTCGCG GTATCATTGC
2701 AGCACTGGGG CCAGATGGTA AGCCCTCCCG TATCGTAGTT ATCTACACGA CGGGGAGTCA
2761 GGCAACTATG GATGAACGAA ATAGACAGAT CGCTGAGATA GGTGCCTCAC TGATTAAGCA
2821 TTGGTAACTG TCAGACCAAG TTTACTCATA TATACTTTAG ATTGATTTAA AACTTCATTT
2881 TTAATTTAAA AGGATCTAGG TGAAGATCCT TTTTGATAAT CTCATGACCA AAATCCCTTA
2941 ACGTGAGTTT TCGTTCCTACT GAGCGTCAGA CCCCCTAGAA AAGATCAAAG GATCTTCTTG
3001 AGATCCTTTT TTTCTGCGCG TAATCTGCTG CTTGCAAACA AAAAAACCAC CGTACCAGC
3061 GGTGGTTTGT TTGCCGGATC AAGAGCTACC AACTCTTTTT CCGAAGGTAA CTGGCTTCAG
3121 CAGAGCGCAG ATACCAAATA CTGTCCTTCT AGTGTAGCCG TAGTTAGGCC ACCACTTCAA
3181 GAACTCTGTA GCACCGCCTA CATACTCGC TCTGCTAATC CTGTTACCAG TGGCTGCTGC
3241 CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTIONAAGA CGATAGTTAC CGGATAAGGC
3301 GCAGCGGTTCG GGCTGAACGG GGGGTTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA
3361 CACCGAAGT GAGATACCTAC AGCGTGAGCT ATGAGAAAGC GCCACGCTTC CCGAAGGGAG
3421 AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTGCGAACA GGAGAGCGCA CGAGGGAGCT
3481 TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCTGGG TTTCCGCCACC TCTGACTTGA
3541 GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC
3601 GGCCTTTTTA CGGTTCTTGG CCTTTTGCTG GCCTTTTGCT CACATGTTCT TTCTGCGTGT
3661 ATCCCCTGAT TCTGTGGATA ACCGTATTAC CGCCTTTGAG TGAGCTGATA CCGCTCGCCG
3721 CAGCCGAAAC ACCGAGCGCA GCGAGTCAGT GAGCGAGGAA GCGGAAGAGC GCCTGATGCG
3781 GTATTTTCTC CTTACGCATC TGTGCGGTAT TTCACACCGC ATATGGTGCA CTCTCAGTAC
3841 AATCTGCTCT GATGCCGCAT AGTTAAGCCA GTATACTC CGCTATCGCT ACGTGACTGG
3901 GTCATGGCTG CGCCCCGACA CCCGCCAACA CCCGCTGACG CGCCCTGACG GGCTTGTCTG
3961 CTCCCGGCAT CCGCTTACAG ACAAGCTGTG ACCGTCTCCG GGAGCTGCAT GTGTGAGAGG
4021 TTTTCACCGT CATCACCGAA ACGCGCGAGG CAGCTGCGGT AAAGCTCATC AGCGTGGTCG
4081 TGAAGCGATT CACAGATGTC TGCTGTGTTCA TCCGCGTCCA GCTCGTTGAG TTTCTCCAGA
4141 AGCGTTAATG TCTGGCTTCT GATAAAGCGG GCCATGTTAA GGGCGGTTTT TTCTGTTTTG
4201 GTCAGTATG CCTCCGTGTA AGGGGGATTT CTGTTTCATG GGGTAATGAT ACCGATGAAA
4261 CGAGAGAGGA TGCTCACGAT ACGGGTACT GATGATGAAC ATGCCCGGTT ACTGGAACGG

```

//

Figura 2



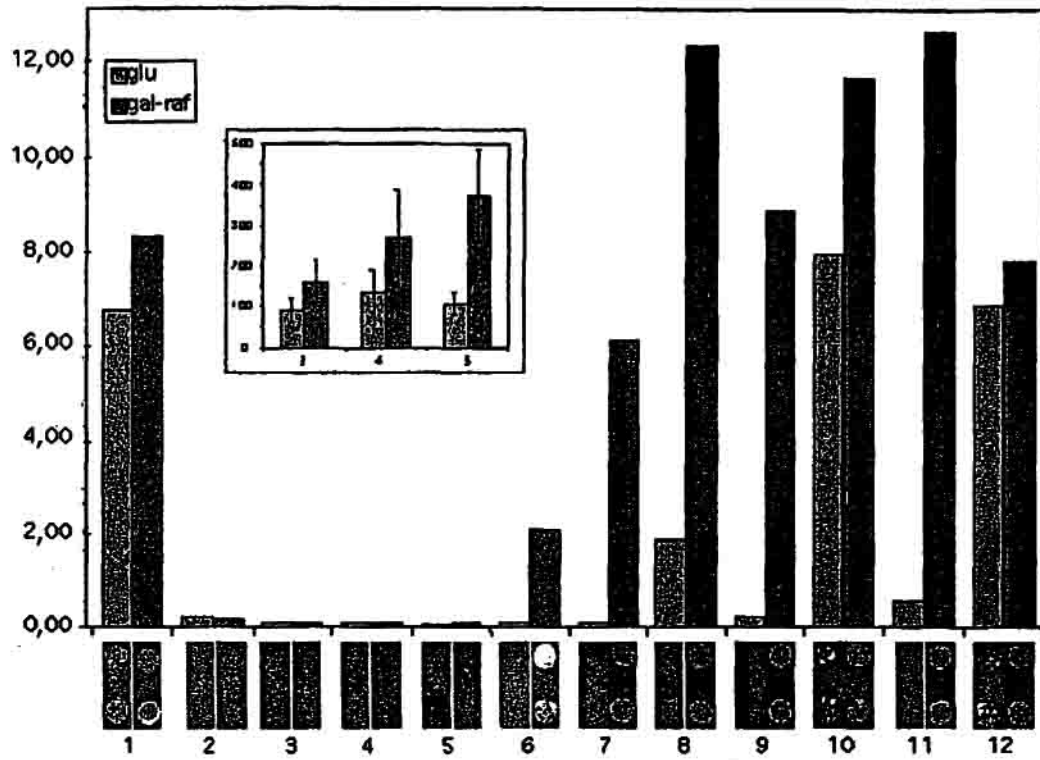


Figura 3

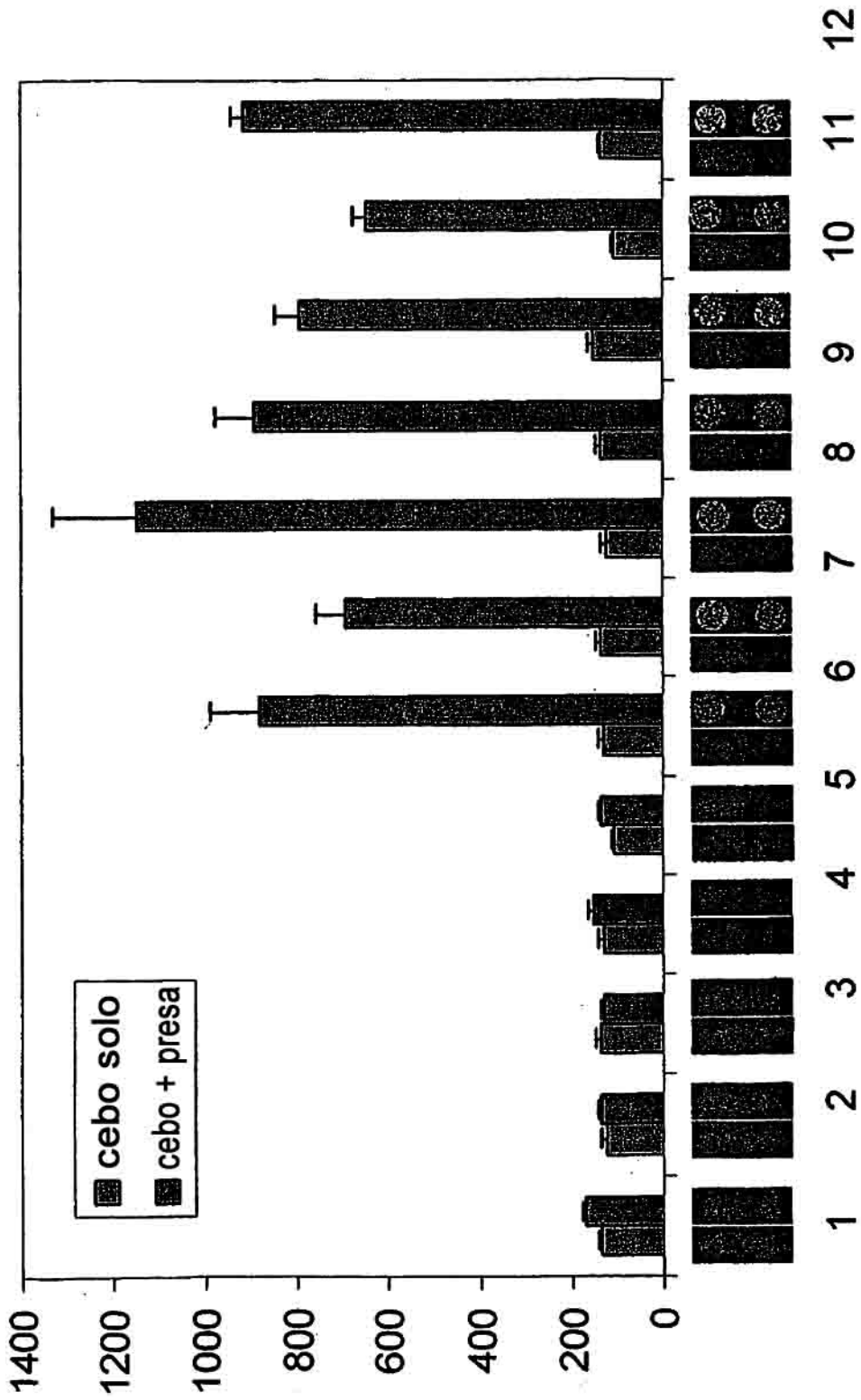


Figura 4

Figura 5

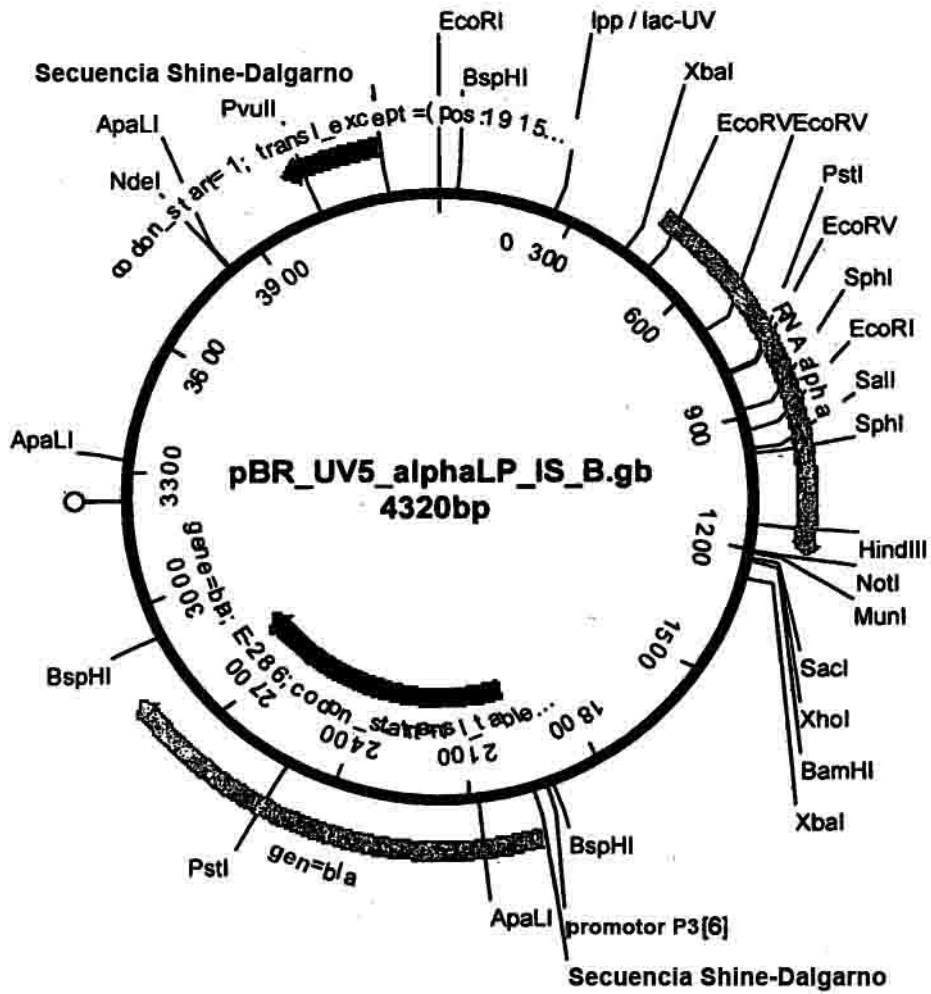


Figura 6

Un elemento promotor híbrido que consta del promotor TEF de levaduras (recuadrado) y del promotor lacUV5 bacteriano (E. coli) (en negrita) (SEQ IN NO: 1)

```

TCATGAGGGGATCCCCCACACACCATAGCTTCAAAATGTTTCTACTCCTTT
TTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCAAA
ACACCCAAGCACAGCATACTAAATTTCCCTCTTTCTTCTAGGGTGTC
GTTAATTACCCGTACTAAAGTTTGGAAAAGAAAAAGAGACCGCCTCGT
TTCTTTTTCTTCGTCGAAAAGGCAATAAAAATTTTATCACGTTTCTTTTT
CTTGAAAATTTTTTTTTGATTTTTTCTTTTCGATGACCTCCATTGATAT
TTAAGTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTG
TTCTATTACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAGCAAA
TCTAATCTAAGGGCGGTGTGACAATTAATCATCGGCATAGTATATCGGCCC
TAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGAATTGTG
AGCGGATAACAATTTCACACAGGAAACAGCGT
    
```