



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 284 245**

51 Int. Cl.:

**C12N 1/19** (2006.01)

**C12N 15/64** (2006.01)

**C12N 15/81** (2006.01)

**G01N 33/50** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **99915053 .5**

86 Fecha de presentación : **26.03.1999**

87 Número de publicación de la solicitud: **1315958**

87 Fecha de publicación de la solicitud: **04.06.2003**

54 Título: **Método para detectar interacción proteína-proteína.**

30 Prioridad: **26.03.1998 US 79480 P**  
**27.03.1998 US 49325**

45 Fecha de publicación de la mención BOPI:  
**01.11.2007**

45 Fecha de la publicación del folleto de la patente:  
**01.11.2007**

73 Titular/es: **GLAXO GROUP LIMITED**  
**Glaxo Wellcome House, Berkeley Avenue**  
**Greenford, Middlesex UB6 0NN, GB**

72 Inventor/es: **Weiner, Michael Phillip y**  
**Buckholz, Richard Gordon**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 284 245 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

Método para detectar interacción proteína-proteína.

5 **Campo de la invención**

La presente invención se refiere a métodos útiles para detectar interacciones proteína-proteína. Las interacciones proteína-proteína permiten la asociación de dos o más proteínas mediante la formación de enlaces no covalentes cuando dos superficies de proteínas concuerdan de forma precisa. Estos enlaces permiten explicar la especificidad del reconocimiento. Las interacciones proteína-proteína están implicadas, por ejemplo, en el ensamblaje de subunidades enzimáticas; en reacciones antígeno-anticuerpo; en la formación de estructuras supramoleculares de ribosomas, filamentos y virus; en el transporte; y en la interacción de los receptores en una célula con los factores de crecimiento y las hormonas. Los productos de los oncogenes pueden dar lugar a la transformación neoplásica a través de interacciones proteína-proteína.

15 **Antecedentes de la invención**

El ensayo de doble híbrido de levaduras (abreviado generalmente como Y2H por sus iniciales en inglés: yeast two-hybrid) es un método para detectar las interacciones proteína-proteína usando un sistema genético. La técnica se puede usar para elaborar un mapa de las interacciones entre proteínas y, por lo tanto, identificar las parejas potenciales en las vías genéticas. El ensayo es sensible y produce las secuencias de ADN que codifican las proteínas que interactúan. En un ensayo de doble híbrido típico, una proteína conocida que forma parte de un híbrido del dominio de unión al ADN se ensaya frente a una biblioteca de todas las proteínas posibles presentes como híbridos del dominio de activación transcripcional. Algunas de las aproximaciones del ensayo de doble híbrido se basan en la unión por interacción. En este método, la proteína fusionada con el dominio de unión al ADN y la proteína fusionada con el dominio de activación se expresan en dos cepas de levadura haploide diferentes con tipos de unión opuestos y las cepas se combinan para determinar si las dos proteínas interactúan. Cuando las cepas de levadura haploide con tipos de unión opuestos se ponen en contacto, se produce la combinación y da lugar a la fusión de los dos haploides para formar una cepa de levadura diploide. Por lo tanto, se puede determinar una interacción midiendo la activación de un gen indicador del doble híbrido en la cepa diploide.

Finley R. L. *et al.* (1997) describen métodos para realizar los ensayos de unión por interacción en conjuntos pequeños o grandes de proteínas usando la trampa de interacción en la que una placa de cepas cebo y una placa de cepas presa se presionan cada una en el mismo tampón de terciopelo para réplicas y la impresión se recoge en una placa que contiene un medio YPD (“Two-Hybrid Analysis of Genetic Regulatory Networks” *Advances in Molecular Biology* (1997) páginas 197-214). Finley R. L. *et al.* (1994) describen las interacciones entre las proteínas reguladoras del ciclo celular de la *Drosophila melanogaster* con una técnica de unión por interacción de levaduras que extiende la trampa de interacción (un método de doble híbrido de levaduras) e implica exponer el patrón de interacciones en matrices de interacción, redes bidimensionales de las cepas diploides en placas de indicador apropiadas (*Proceedings of the National Academy of Sciences of USA National Academy of Science*, Washington, 91:12980-12984).

El documento WO 94/10300 y la Patente Estadounidense N° 5.283.173 describen métodos para detectar la interacción entre proteínas usando la reconstitución de la actividad de un activador transcripcional. Esta reconstitución utiliza genes híbridos que expresan proteínas híbridas. El primer híbrido contiene el dominio de unión al ADN de un activador transcripcional fusionado con una proteína conocida (la proteína “cebo”), con el dominio de unión al ADN situado en dirección 5’ de un gen indicador. Las proteínas “presa” se clonan bien como secuencias aleatorias o bien como ADNc y se fusionan con el extremo amino o carboxi de un dominio de activación de la transcripción. Si las proteínas cebo y presa son capaces de interactuar, ponen en proximidad cercana los dos dominios del activador transcripcional. Esta proximidad es suficiente para producir la transcripción que puede detectarse por la actividad del gen indicador que contiene un sitio de unión para el dominio de unión al ADN.

La desventaja de estas técnicas es que se generan interacciones irrelevantes con las proteínas de la levadura. Estas incluyen interacciones falso positivo que es improbable que se encuentren en células vivas, e interacciones falso negativo, es decir aquellas interacciones que se podrían detectar de otra forma pero que no se detectan. Las técnicas descritas en el documento WO 94/10300 y en la Patente Estadounidense N° 5.283.173 precisan de la utilización de unión en un medio sólido, lo cual es difícil de manejar, requiere mucho trabajo y no preserva las células diploides para análisis posteriores.

Se ha desarrollado una estrategia de unión del ensayo de doble híbrido de levaduras en un formato automatizado que permite procesar muchas proteínas cebo. El formato usa un medio de ordenamiento, por ejemplo, placas de microvaloración, y unión masiva en medio líquido de un subconjunto de una biblioteca compleja grande. Mediante la búsqueda de las interacciones positivas en la biblioteca, también se ha desarrollado un método para crear una biblioteca funcionalmente sustraída, es decir, una biblioteca que puede desproveerse de un fenotipo marcable. Por ejemplo, nuestro método permite la determinación o detección de híbridos que reaccionan indiscriminadamente con muchas dianas, tales como proteínas de choque térmico, y su eliminación en cualquier consideración futura.

### Compendio de la invención

Según la presente invención, se proporciona un método para detectar interacciones proteína-proteína que comprende la unión masiva en medio líquido de subseries de una gran biblioteca compleja. El método proporciona un medio para sustraer interacciones proteína-proteína irrelevantes para producir un ensayo “funcionalmente sustraído”.

### Descripción detallada de la invención

Según un aspecto de la presente invención, se proporciona un método para detectar una interacción entre una primera proteína de ensayo y una segunda proteína de ensayo, que comprende:

- (a) proporcionar una célula huésped que contiene un gen indicador, en la que el gen indicador expresa una proteína detectable cuando el gen indicador es activado por una secuencia de aminoácidos que incluye un dominio de activación transcripcional cuando el dominio de activación transcripcional está suficientemente próximo al gen indicador;
- (b) proporcionar un primer gen híbrido que puede ser expresado en la célula huésped, comprendiendo el primer gen híbrido una secuencia de ADN que codifica una primera proteína híbrida, comprendiendo la primera proteína híbrida:
  - (i) un dominio de unión al ADN que reconoce un sitio de unión en el gen indicador en la célula huésped; y
  - (ii) una primera proteína de ensayo o uno de sus fragmentos, en la que se va a ensayar la interacción con al menos una segunda proteína de ensayo o uno de sus fragmentos;
- (c) proporcionar un segundo gen híbrido que puede ser expresado en la célula huésped, comprendiendo el segundo gen híbrido una secuencia de ADN que codifica una segunda proteína híbrida, comprendiendo la segunda proteína híbrida:
  - (i) el dominio de activación transcripcional; y
  - (ii) una segunda proteína de ensayo o uno de sus fragmentos, en la que se va a ensayar la interacción con la primera proteína de ensayo o uno de sus fragmentos; en la que la interacción entre la primera proteína de ensayo y la segunda proteína de ensayo en la célula huésped hace que el dominio de activación transcripcional active la transcripción del gen indicador;
- (d) introducir el segundo gen híbrido en la célula huésped y subsiguientemente introducir dichas células en un medio de ordenamiento, creando de este modo una placa de biblioteca maestra;
- (e) introducir las células de la placa de biblioteca maestra en un segundo medio de ordenamiento que es un medio de ordenamiento para medios líquidos, creando de este modo un conjunto de unión;
- (f) introducir el primer gen híbrido en la célula huésped e introducir subsiguientemente dicha célula en el conjunto de unión, permitiendo de este modo que se produzca la unión en el medio líquido;
- (g) elegir el producto de la interacción de los genes primero y segundo;
- (h) retirar una parte del conjunto de unión a un tercer medio de ordenamiento, creando de este modo un conjunto de liberación;
- (i) determinar si el gen indicador ha sido expresado en el conjunto de unión; y
- (j) analizar las células de la placa de liberación.

El término “gen indicador” o “gen marcador” como se usa en la presente memoria significa cualquier gen cuya expresión se puede analizar. Más de un gen indicador puede ser codificado por la célula huésped en la etapa (a) anterior.

El término “medio de ordenamiento” como se usa en la presente memoria significa cualquier método para mantener los clones en un medio líquido, suspensión o medio sólido, por ejemplo placas de microvaloración o tubos de ensayo.

El término “elegir el producto” como se usa en la presente memoria significa cualquier método que usa un medio elegible bien para amplificar o bien para aislar un conjunto de proteínas que interaccionan. Este medio elegible puede incluir el producto en un medio de crecimiento nutricionalmente deficiente en el que las proteínas que interaccionan producen la transcripción de un gen o vía biosintética. Los ejemplos de otros medios elegibles útiles incluyen genes de biosíntesis de aminoácidos, metabólicas, catabólicas y de ácidos nucleicos, tales como las levaduras HIS3, URA3 y LYS2, GAL1, *E. coli* galK y CAT, GUS, resistencia a los antibióticos y cualquier gen que codifique un antígeno de

## ES 2 284 245 T3

superficie celular para el que están disponibles los anticuerpos. Se puede permitir que el crecimiento continúe durante 5-10 días antes de elegir el producto.

5 El término “analizar” como se usa en la presente memoria significa cualquier método para obtener información que se refiera a las interacciones proteína-proteína, por ejemplo, elegir clones positivos, realizar una PCR, análisis de secuencia de ADN y comparación con bases de datos, tales como LifeSeq® (Incyte Pharmaceuticals) o Genbank.

10 El término “funcionalmente sustraído” significa desprovisto de un fenotipo detectable que representa una interacción proteína-proteína irrelevante.

15 En un aspecto adicional de la invención, la determinación de la expresión de un gen indicador y el análisis de las células se pueden realizar en una etapa, es decir las etapas (i) y (j) anteriores se pueden combinar. Alternativamente, las etapas (h), (i) y (j) se pueden eliminar.

20 Una cepa huésped de levadura se puede modificar genéticamente para expresar la proteína (el “cebo”) de interés terapéutico o diagnóstico como una proteína de fusión enlazada covalentemente con un dominio de enlace al ADN conocido de un activador transcripcional. La cepa huésped de levadura también contiene uno o más “genes indicadores”, es decir genes cuya transcripción se detecta como respuesta a la interacción cebo-presa. Las proteínas cebo, a través de su dominio de enlace al ADN, se enlazan a su sitio de ADN específico en dirección 5’ de un gen indicador; sin embargo, la transcripción del gen indicador no está estimulada porque la proteína cebo carece de su propio dominio de activación.

25 Para aislar los genes que codifican las nuevas proteínas interaccionantes, las células de esta cepa que contienen un gen indicador y expresan una proteína cebo se transforman con miembros individuales de una biblioteca de expresión de ADN (por ejemplo, un ADNc). Cada miembro de la biblioteca se dirige a la síntesis de una proteína interaccionante candidata fusionada con un marcador del dominio de activación genético débil e invariante. Las proteínas codificadas por la biblioteca (proteínas “presa”) que interaccionan físicamente con la proteína cebo enlazada al promotor activan de forma detectable la transcripción del gen indicador en dirección 3’ y proporcionan un ensayo disponible para identificar células particulares que hospedan un clon de ADN que codifica una proteína interaccionante de interés.

30 En un modo de realización, una biblioteca de ADNc, creada en *E. coli*, y que comprende ADNc fusionado con la secuencia de ADN que codifica el dominio de activación del activador transcripcional, proteína GAL4, se coloca en placas sobre agar 960 LB y con una densidad de 1.000 clones por placa. Las colonias de *E. coli* de cada placa se combinan, se aíslan los ADN plasmídicos y los ADN se usan para transformar la levadura. Las levaduras transformadas se ponen en placas sobre un medio sólido y las colonias de cada placa se combinan y se ponen alícuotas en pozos separados de una placa de microvaloración de 96 pozos para crear un conjunto ordenado de 10 placas de “biblioteca maestra”. Se vuelven a tomar alícuotas de cinco microlitros de cada pozo del conjunto de la biblioteca maestra para crear un “conjunto de unión” y entonces se añaden separadamente 5  $\mu$ L de la levadura que contiene el cebo en cada pozo. El “cebo” comprende un gen híbrido que expresa una proteína híbrida que contiene el dominio de unión al ADN de la GAL4 fusionado con una proteína conocida. La cepa de levadura huésped contiene el gen GAL1-lac-Zm que es capaz de unirse al dominio de unión del ADN de la GAL4. El gen GAL1-lacZ contiene el gen lacZ de la *E. Coli* que codifica la  $\beta$ -galactosidasa. La actividad de la  $\beta$ -galactosidasa es una medida de la función GAL4. El crecimiento de levadura sobre galactosa requiere la transcripción de genes regulados por la GAL4 y también es una medida de la función GAL4. Se deja que la unión masiva en medio líquido prosiga durante un periodo de tiempo y la mezcla de unión se diluye 100 veces con un medio de salida por goteo de leucina. Después de producir levaduras diploides unidas que interaccionan positivamente en el medio de salida por goteo, se retira una porción a un conjunto separado de placas de “liberación” y se realiza un análisis de la  $\beta$ Gal en el conjunto de unión. La activación transcripcional se puede determinar midiendo la actividad  $\beta$ -galactosidasa en un medio que contiene galactosa. Se identifican los pozos que presentan cualquier actividad  $\beta$ Gal y los clones del conjunto de pozos correspondiente de las placas de liberación se analizan mediante secuenciación por PCR.

35 40 45 50 Las interacciones proteína-proteína irrelevantes se pueden eliminar recombinando solo los clones productivos de la biblioteca maestra, eliminando de este modo los clones que producen proteínas que se sabe que interaccionan con muchas otras proteínas, por ejemplo las proteínas de choque térmico.

55 También se proporciona un método para una estrategia de clonación de marco de lectura abierto que implica la recodificación dinámica de los extremos de moléculas de ADN. Esta estrategia de clonación aumenta la eficacia del ensayo eliminando del análisis todos los clones que codifican proteínas que presentan un desplazamiento del marco de lectura con respecto al dominio de activación.

60 65 En la recodificación dinámica de un dominio de activación, el extremo 3’ del gen del dominio de activación puede volverse a codificar para incorporar un péptido híbrido de aminoácido que también codifique el ADN que controla los elementos necesarios para la expresión del gen de la *E. coli*. En un aspecto, estos elementos comprenden, en serie: i) una secuencia, por ejemplo, una secuencia -35 y -10, que actúa como un promotor de la *E. coli* para iniciar la transcripción del ARNm, ii) un sitio de unión al ribosoma y un codón ATG del fMet necesarios para iniciar la traducción de la proteína, iii) un sitio de clonación múltiple compuesto por uno o más sitios de restricción que son preferiblemente únicos para el vector de clonación, en los que clonar fragmentos de relleno de ADN que pueden codificar fusiones de proteína al dominio de activación, y iv) un gen indicador, por ejemplo el gen lacZ, clonado con

## ES 2 284 245 T3

desplazamiento del marco de lectura con respecto al codón ATG. En el sistema de clonación de marco de lectura abierto, el ATG puede no presentar modificaciones en el marco de lectura con respecto al dominio de activación, el ATG puede presentar un desplazamiento del marco de lectura con respecto al gen lacZ, hay una cantidad despreciable de proteína  $\beta$ Gal producida por la célula huésped en ausencia de un fragmento de relleno que restaura el marco de lectura del gen lacZ y hay ausencia de codones de terminación en el extremo del gen del dominio de activación y el codón ATG.

El término “fragmentos de relleno” significa cualquier fragmento de ADN generado sintéticamente, o mediante el uso de un método generalmente utilizado para generar moléculas de ADNc con cebado aleatorio o en el extremo 3', las cuales pueden ser clonadas en el sitio de clonación múltiple del sistema de clonación de marco de lectura abierto anterior.

En un aspecto, el ADNc con cebado aleatorio usado como fragmento de relleno puede ser elegido en tamaño por electroforesis sobre gel de agarosa o de poliacrilamida. El ADNc individual elegido por tamaño mediante electroforesis sobre gel u otros medios puede contener fragmentos que, cuando se clonan en el sistema vector descrito, puede estar en uno de los seis marcos de lectura (tres marcos de lectura en cada orientación, directa e inversa).

El dominio de activación recodificado se puede usar conjuntamente con el gen indicador con desplazamiento del marco de lectura para elegir los clones que restauran el marco de lectura del gen indicador. Por ejemplo, si el gen lacZ presenta inicialmente un desplazamiento del marco de lectura con respecto al inicio ATG de la porción recodificada del dominio de activación, entonces los clones que restauran el marco de lectura entre el ATG y el gen lacZ producirán fusiones de proteínas de este clon con el producto del gen lacZ. Las fusiones que restauran la actividad  $\beta$ gal se pueden elegir mediante ensayos cromogénicos usando colorantes bien conocidos (por ejemplo, Xgal) o en un medio de crecimiento selectivo que contiene lactosa como única fuente de carbono.

En el sistema de clonación de marco de lectura abierto, un codón de terminación suprimible de *E. coli* (por ejemplo, un codón de terminación ámbar TAG) puede ser codificado entre el fragmento de relleno y el gen indicador de forma que en cepas huésped de *E. coli* que lo suprimen fenotípicamente, el codón de terminación se suprime mediante una molécula de ARNt supresora que inserta un aminoácido específico. En células huésped no supresoras en las que se realiza el ensayo de interacción, la terminación de la traducción de la proteína ocurriría en el codón de terminación. La ventaja de tener este sistema suprimible es que la proteína indicadora de marco de lectura abierto no se fusionará en el extremo carboxílico de la proteína híbrida del dominio de activación con fragmento de relleno codificado.

La célula huésped es una célula de levadura, ventajosamente *Saccharomyces cerevisiae*.

La proteína cebo puede derivarse de una proteína bacteriana, una proteína vírica, una proteína codificada por un oncogén, un factor de crecimiento o una enzima. Las proteínas cebo se pueden elegir entre cualquier proteína de importancia diagnóstica o terapéutica conocida o sospechada. Las proteínas cebo preferidas incluyen las oncoproteínas (tales como myc, ras, src o fos) o cualquier otra proteína implicada en la regulación del ciclo celular (tales como las quinasas o las fosfatasa).

Las proteínas presa pueden ser codificadas por una biblioteca de plásmidos que contienen insertos de ADN que se derivan del ADN genómico, ADNc, o secuencias de ADN generadas sintéticamente fusionadas con la secuencia de ADN que codifica el segundo dominio de aminoácido. Los ADNc se pueden construir a partir de cualquier población de ARNm e insertar en un vector de expresión equivalente. Tal biblioteca de elección se puede construir *de novo* usando kits disponibles comercialmente (por ejemplo en Stratagene, La Jolla, CA) o usando procedimientos preparativos bien establecidos (por ejemplo *Current Protocols in Molecular Biology*, Nueva York, John Wiley & Sons, 1987). Alternativamente, se puede usar una biblioteca de ADNc disponible comercialmente. Una proteína presa puede ser codificada por una secuencia sintética o puede ser el producto de un marco de lectura abierto generado aleatoriamente o una de sus porciones.

Se puede usar cualquier gen indicador, por ejemplo el gen LEU2 o el gen lacZ. Los ejemplos de otros genes útiles cuya transcripción se puede detectar, incluyen genes de biosíntesis de aminoácidos y de ácidos nucleicos, tales como las levaduras HIS3, URA3 y LYS2, GAL1, *E. coli* galK, FP, CUP1 y CAT, GUS, resistencia a los antibióticos y cualquier gen que codifique un antígeno de superficie celular para el que están disponibles los anticuerpos.

Los expertos en la técnica también reconocerán que el gen indicador, el dominio de unión al ADN y los componentes del dominio de activación del gen se pueden derivar a partir de cualquier genoma o ADNc apropiado de células eucarióticas o procarióticas, así como de secuencias artificiales.

Los constructos plasmídicos, la transformación, transfección, cultivo celular y detección de la transcripción se pueden realizar por cualquier método conocido en la técnica, por ejemplo en la Patente Estadounidense N° 5.283.173 y el documento WO 94/10300.

Se puede usar cualquier método para introducir genes en células huésped, por ejemplo electroporación, transfección, transformación o unión.

## ES 2 284 245 T3

Las ventajas de la invención descrita incluye una eficacia aumentada mediante la eliminación de análisis adicionales de proteínas de interacción indiscriminada en bibliotecas ordenadas, la creación de un medio para sustraer funcionalmente clases de proteínas a partir de bibliotecas, la eliminación de análisis adicionales de clones que no presenten un marco de lectura especificado, un trabajo menor con respecto a métodos habituales, reutilización de bibliotecas primarias a partir de conjuntos de bibliotecas maestras ordenadas y el conocimiento acumulado a lo largo del tiempo de la composición de los clones ordenados.

La invención puede ilustrarse por los siguientes Ejemplos no limitantes.

### 10 Ejemplo 1

#### *Unión masiva en medio líquido, ensayo de doble híbrido de levaduras funcionalmente sustraídas*

Las enzimas de restricción y de modificación del ADN se obtuvieron de varios fabricantes y se usaron según sus recomendaciones.

Creación de bibliotecas de ADN ordenadas (Figura 1). Las bibliotecas de ADNc de *E. coli* se obtuvieron de Invitrogen y se colocaron con una densidad baja (aproximadamente 1.000 clones por placa) en placas con LB + Amp y se incubaron durante 1-2 días a 37°C. A continuación, se añadieron 3-4 mL de LB (que contenía 15% de glicerol) a cada placa, la placa se sacudió sobre una plataforma agitadora a baja velocidad y el LB se cultivó después de que fuera aparente que las colonias volvían a estar en suspensión en el LB. Se retiró una porción de 200  $\mu$ L de las células vueltas a poner en suspensión para aislar el ADN plasmídico y las células restantes se congelaron para su almacenamiento a largo plazo a -80°C.

El ADN plasmídico se aisló por medio de un kit obtenido en Qiagen. Se añadieron doscientos cincuenta (250)  $\mu$ L de disolución de P2 (Qiagen) a la porción de 200  $\mu$ L de células en un tubo de centrifugadora Eppendorf de 2 mL. Se mezclaron suavemente las dos disoluciones y luego se añadieron 250  $\mu$ L de disolución de P3 (Qiagen) y se agitó el tubo. Después se centrifugó la mezcla a alta velocidad (14.000 rpm) en una centrifugadora Eppendorf. El sobrenadante clarificado (500  $\mu$ L) se pipeteó a un nuevo tubo de centrifugadora Eppendorf de 1 mL y se añadió etanol para precipitar el ADN. El ADN precipitado se centrifugó a alta velocidad (14.000 rpm) durante 15 minutos, se decantó la disolución de etanol y el precipitado se secó a vacío. El precipitado se volvió a poner en suspensión en 50  $\mu$ L de H<sub>2</sub>O destilada y se usó directamente para transformar la levadura.

Las levaduras se transformaron usando el kit EZ Yeast Transformation (Zymo Research) según las recomendaciones del fabricante, usando 2,5  $\mu$ L de ADN, 25  $\mu$ L de cepa de levadura idónea EGY48 y 250  $\mu$ L de EZ3. Las levaduras transformadas se incubaron durante 1 hora a 30°C y el total se colocó en placas con agar SD - trp. Las placas se incubaron durante 3-4 días adicionales a 30°C y las células se cultivaron como para la *E. coli* usando 3-4 mL de SD - trp + 15% de glicerol. Se tomaron alícuotas de las células cultivadas de cada placa y se pusieron separadamente en pozos diferentes de placas de 96 pozos de formato hondo (las placas de la "biblioteca maestra") y se congelaron a -80°C para almacenamiento a largo plazo.

Unión de levaduras en medio líquido (Figura 2). Cinco  $\mu$ L de cada una de las levaduras de los pozos de la biblioteca maestra se inocularon en 100  $\mu$ L de SD - trp o SGal-trp en una placa de 96 pozos y se dejaron crecer durante la noche a 30°C. Se transfirieron cinco  $\mu$ L de cada pozo a una nueva placa de 96 pozos "de unión". Se añadió una alícuota de 5  $\mu$ L de un cultivo de cebo (OD<sub>600</sub> = 1,0) en cada pozo junto con 10  $\mu$ L de medio YPD. Las placas de unión se colocaron en una bolsa de plástico resellable y se incubaron durante 12-36 horas a 30°C. A continuación, cada pozo se diluyó 1:10, dos veces en serie, (dilución final 100 veces) usando S-min (- leu, - his, - trp, - ura, + gal, + raff) hasta un volumen final de 110  $\mu$ L. Alternativamente, cada pozo se diluyó entonces 1:10 en S-min, se incubó durante 30°C durante dos días y luego se diluyó 1:40 en S-min (dilución final 400 veces). Las uniones diluídas se incubaron durante 5-10 días adicionales a 30°C. 10  $\mu$ L de los pozos de unión se transfirieron entonces a un segundo conjunto de placas antes de realizar el análisis de la  $\beta$ Gal (más tarde estos 10  $\mu$ L de disoluciones madre de unión y de productos ("placas de liberación") se usaron para liberar los clones positivos).

Análisis de la  $\beta$ Gal. Se lisaron las células por adición de 100  $\mu$ L de una disolución tampón Z [Na<sub>2</sub>HPO<sub>4</sub>, (16,1 g L<sup>-1</sup>), NaH<sub>2</sub>PO<sub>4</sub>, (5,5 g L<sup>-1</sup>), KCl (0,75 g L<sup>-1</sup>) y MgSO<sub>4</sub>, (0,25 g L<sup>-1</sup>), se ajustó el pH a pH 7,0 y se esterilizó en autoclave] que contenía oxalícas (100 U mL<sup>-1</sup>), SDS (0,1%) y sustrato CPRG (2 mg mL<sup>-1</sup>). Las placas se incubaron a temperatura ambiente hasta que se reveló el sustrato cromogénico rojo de la  $\beta$ Gal (generalmente entre 10 minutos y 2 horas). Para analizar cuantitativamente los pozos fue necesario eliminar los residuos celulares, bien por centrifugación o bien por filtración. El sustrato CPRG puede medirse con una absorbancia de 575 angstrom. Alternativamente, la actividad  $\beta$ -gal se midió usando un sustrato quimioluminiscente. Con este objetivo se usó el kit Tropix Galacton Plus. Se transfirieron veinticinco microlitros de cada pozo de ensayo a los pozos correspondientes de placas de 96 pozos para luminiscencia. Se añadieron en cada pozo veinticinco microlitros de disolución tampón de reacción CL (disolución tampón Z que contenía 0,2% de Igepal CA-630, 100 U/mL de oxalícas y 1% de Galacton Plus) y las placas se incubaron durante la noche a temperatura ambiente. Se añadieron en cada pozo cincuenta microlitros de Accelerator II (diluido 1:1 en Na<sub>2</sub>CO<sub>3</sub> 0,1M/NaHCO<sub>3</sub>, pH 10,5) y las placas se incubaron durante 5 minutos. Entonces se midió la quimioluminiscencia en un luminómetro de 96 pozos.

## ES 2 284 245 T3

Análisis de la sensibilidad a la combinación. Se realizó un análisis de la estrategia de unión en medio líquido combinada usando los conocidos interactores fuertes del Y2H RPB4 (subunidad polII de levadura) y PB7 (subunidad polII de levadura) como controles. La subunidad RPB4 se subclonó en el vector del dominio de activación pJG4.5. La fusión del RPB4 recombinante se subclonó en el vector del dominio de unión al ADN pEG202, se transformó en la cepa presa y se mezcló a varios porcentajes (de 0 a 100%) con la misma cepa presa que contenía el vector pJG4.5 precursor.

Los resultados (mostrados en la Figura 3) demostraron que mediante esta interacción se fue capaz de recuperar la cepa presa, incluso cuando la presa inicialmente representaba aproximadamente 0,1% de la “mezcla de unión” presa. Los resultados sugirieron que podía ser necesaria la dilución en aproximadamente 100 veces del medio YPD complejo con el fin de observar el crecimiento diferencial de los pares que interactúan positivamente. La dilución de las muestras para disminuir la concentración de complejo YPD puede ser preferible con respecto a otros métodos, tales como centrifugación o filtración. Esto es debido a que la dilución es más barata, más rápida y más fácil de automatizar. El análisis de la  $\beta$ Gal en el ensayo de la activación del indicador en un formato de placa de microvaloración combinada no mostró diferencias significativas con la fusión recombinante entre 0,1 y 100% en el punto de dilución de 100 veces. A mayores diluciones se puede producir dispersión de la actividad  $\beta$ Gal. Puede ser que para mayores diluciones (o menor porcentaje de combinaciones) se pierda el muestreo de interactores positivos.

Análisis de las bibliotecas de ADNc ordenadas combinadas. En el primer análisis del experimento de bibliotecas de ADNc ordenadas, se analizaron los receptores nucleares RXR y LXRA frente a aproximadamente  $6 \times 10^5$  ADNc en 6 placas de microvaloración. La mayoría de los ADNc eran de bibliotecas de ADNc disponibles comercialmente obtenidas de hígado fetal humano (Invitrogen A202-01) y de cerebro fetal humano (Invitrogen A212-01).

Los clones que contenían el ADNc se sembraron con aproximadamente  $1 \times 10^3$  clones por pozo. En resumen, la cepa cebo (que contiene la proteína diana, en este caso bien RXR o bien LXRA) se añadió a los clones de la biblioteca de ADNc en los pozos y se permitió que se produjera la unión en un medio complejo. Las mezclas de unión se diluyeron en un medio mínimo (-leucina) y se permitió que se produjera el crecimiento de interactores (indicando el crecimiento una interacción exitosa) durante 5 o más días. A continuación se realizaron análisis de  $\beta$ Gal en los pozos (véase por ejemplo, la Figura 2), y los clones de 10 pozos que presentaban actividad  $\beta$ Gal esencial se volvieron a aislar rayando una alícuota del pozo de la biblioteca sobre un medio sólido mínimo (-leucina). Se aislaron los plásmidos de estos clones y se sometieron a secuenciación del ADN y análisis bioinformáticos. Los resultados se muestran a continuación en la Tabla I.

Algunos de los clones secuenciados se han encontrado mediante análisis Y2H tradicional. Estos incluyeron TRIP6 (del inglés: thyroid receptor interacting protein 6, proteína 6 que interactúa con el receptor del tiroides) que se ha descrito previamente en la bibliografía de otros experimentos de trampa de interacción estándar frente a otros receptores nucleares (todavía no se ha ensayado frente a LXRA) y TIF1. Se cree que estos representan positivos verdaderos. Los otros clones, que codifican ambos un homólogo del GCN5, se aislaron dos veces (en dos pozos diferentes). Aún no se sabe si el homólogo del GCN5 es un falso positivo o un positivo verdadero.

Se encontró que aproximadamente un tercio de los clones que interactúan tenían homología con los ADNc de las bases de datos Incyte o GenBank, pero no se les ha adscrito ninguna función.

Varios clones parecen ser positivos indiscriminados conocidos en los experimentos de trampa de interacción (principalmente, la cofilina y las proteínas de choque térmico). Ahora que se sabe cuáles de estos pozos están en el medio, se pueden eliminar para futuros análisis. Sin embargo, debe señalarse que cuando se excluyen estos pozos, se está perdiendo información de los otros clones en ese pozo. Por ejemplo, usando RXR como cebo, se encontró una interacción con la proteína relacionada con la timopoyetina en el pozo 1D9. En este mismo pozo, cuando se investigó con LXRA se encontró una interacción positiva con el HSP90 positivo indiscriminado. Se espera que se utilice opcionalmente una biblioteca de ADNc suficientemente grande para obtener redundancia en los análisis de la biblioteca.

55

60

65

# ES 2 284 245 T3

TABLA 1

*Resultados del análisis Y2H*

Cebo	Secuencia ID, pozo #	Homología representativa	Comentarios
RXR	5rxr, 1A10	TIF1	NR (+) conocido <sup>c</sup>
LXRa	13gor4, 3G4	TRIP6	NR (+) conocido <sup>c</sup> encontrado en un muestreo estándar <sup>c</sup>
RXR	1rxr, 1G6	anexina (IPP)	encontrado en un muestreo estándar <sup>c</sup>
LXRa	7gor4, 5A10	Homólogo del GCN5	Implicado en la transcripción <sup>d</sup>
LXRa	20gor4, 5C5	Homólogo del GCN5	Implicado en la transcripción
RXR	3rxr, 1D9	relacionado con la timopoitena	positivo verosímil
LXRa	12gor4, 2G1	KIAA0229	Genbank EST, (función no conocida)
LXRa	5gor4, 1H8	Incyte 3122030 <sup>b</sup>	sin anotación en el GenBank, véase también 21gor4
LXRa	9gor4, 2F8	Incyte 004215 <sup>b</sup>	sin anotación en el GenBank
LXRa	15gor4, 4C2	Incyte 1366945 <sup>b</sup>	sin anotación en el GenBank
LXRa	21gor4, 2B10	Incyte 3122030 <sup>b</sup>	sin anotación en el GenBank
RXR	7rxr, 3B8	gen de la nucleolina	falso (+) conocido <sup>a</sup>
LXRa	1gor4, 1A8	huHSP86	falso (+) conocido <sup>a</sup>
LXRa	3gor4, 1D9	huHSP90	falso (+) conocido <sup>a</sup>
LXRa	11gor4, 5E11	Cofilina	implicado en la estructura celular, interacciona con la actina

<sup>a</sup> Positivo común conocido en otros muestreos Y2H (E. Golemis).

<sup>b</sup> No se encontró anotación en la secuencia específica de la base de datos Genbank.

<sup>c</sup> Se sabe que esta proteína interacciona con varios otros receptores nucleares.

<sup>d</sup> GCN5 presenta actividad histona acetiltransferasa (HAT).

<sup>e</sup> La proteína también se aisló usando metodología de doble híbrido tradicional.

## Ejemplo 2

### *Estrategia de clonación de marco de lectura abierto*

*Clonación de marcos de lectura abiertos y utilización de la supresión para controlar la fusión del gen en 3'*

Se aislaron ADNc cortados aleatoriamente de aproximadamente 600 pares de bases y se clonaron en un gen de  $\beta$ gal con cambio del marco de lectura (Figura 5A). Las células de *E. coli* transformadas que se convirtieron en  $\beta$ gal<sup>+</sup> contenían un marco de lectura abierto. En un vector con un codón de terminación suprimible ámbar entre el extremo 3' del ADNc y el extremo 5' de la  $\beta$ gal, la fusión del ADNc con la  $\beta$ gal se controló por el fenotipo Sup de la cepa de *E. coli* (figura 5B). El mismo tipo de esquema de clonación se puede adaptar para fusionar el extremo 5' de la proteína

de exposición del fago M13 al ADNc, en este caso el fago viable indicará una clonación exitosa del marco de lectura abierto (Figura 5C).

#### Recodificación dinámica del extremo 3' del dominio de activación de la levadura

5 Se recodificó el extremo 3' del dominio de activación de la levadura para incorporar los elementos que controlan la expresión del gen de *E. coli*. La Figura 6 es un ejemplo de recodificación de los elementos de control necesarios para la expresión de la proteína del bacteriófago T7. El dominio de activación de la levadura recodificado se usó entonces conjuntamente con el sistema de clonación de marco de lectura abierto para fusionar el marco de lectura correcto al  
10 dominio de activación y, simultáneamente, a una proteína de fusión 3' separada (por ejemplo,  $\beta$ gal o M13gp3).

La selección fenotípica en *E. coli* de los marcos de lectura abiertos de ADNc y la fusión simultánea de ellos con el extremo 3' del dominio de activación de la levadura y con el extremo 5' del M13gpIII se puede realizar según la Figura 7: A. fragmentos de ADNc de 500 bp clonados aleatoriamente; transformación T7RNAP<sup>+</sup>, Sup<sup>+</sup> *E. coli*; placas  
15 de muestreo. B. Muestra del fago en *E. coli*. C. Doble híbrido en levadura: transformación T7RNAP<sup>+</sup>, Sup<sup>+</sup> *E. coli*; transformación M13 en levadura Sup<sup>-</sup>.

#### Breve descripción de los dibujos

20 Figura 1. Creación de las bibliotecas de ADNc ordenadas. Las bibliotecas de ADNc de *E. coli* se pusieron con baja densidad (aproximadamente 1.000 colonias por placa) sobre placas con LB + Amp. A continuación se añadieron 3-4 mL de LB (que contenía 15% de glicerol) a cada placa y se cultivó el LB después de que la resuspensión de las colonias en el LB fue evidente. El ADN plasmídico se aisló por medio de un kit obtenido de Qiagen (Valencia, CA) y se usó directamente para transformar la levadura. Las levaduras transformadas se pusieron sobre placas de agar SD - trp. Las  
25 placas se incubaron y las células se cultivaron como para la *E. coli* usando 3-4 mL de SD - trp + 15% de glicerol. Se tomaron alícuotas de las células cultivadas de cada placa y se pusieron separadamente en pozos diferentes de placas de 96 pozos de formato hondo (las placas de la "biblioteca maestra") y se congelaron a -80°C para almacenamiento a largo plazo.

30 Figura 2. Formato de Y2H automatizable. Para realizar el análisis de Y2H se usaron las siguientes etapas en una placa de microvaloración: i) Se añade la cepa cebo a la cepa de la biblioteca de ADNc en un pozo, ii) se permite que se produzca la unión en un medio complejo, iii) se diluye la mezcla de unión en un medio de salida por goteo mínimo (-leu), iv) se permite el crecimiento de las proteínas que interaccionan positivamente (cultivadas como lectura), v) se realiza el análisis de la  $\beta$ Gal (lectura cuantitativa), vi) clones con secuencia (+), investigación en las bases de datos.  
35 Leyenda: 1 = biblioteca maestra; 2 = descendiente; 3 = cebo añadido (unión durante 24-36 horas); 4 = producto selectivo (incubación > 5 días); 5 = archivo; 6 =  $\beta$ Gal.

Figuras 3A y 3B. Análisis de la activación del indicador en un pozo de la placa de microvaloración combinada. Los resultados se expresan como % de interacto en la combinación de no interactores. Se mezclaron interactores  
40 conocidos en una relación conocida y se estudió la fusión con el cebo en el formato de unión en medio líquido. Eje X: dilución de unión; Eje Y: A<sub>600</sub> o A<sub>574</sub>. Figura 3A: producto selectivo después de la dilución de la levadura unida en el medio de salida por goteo de leucina. Fusión con el dominio de activación: pJG-scrPB4 (subunidad polII de levadura); fusión con el dominio de unión al ADN: pEG-scrPB7 (subunidad polIII de la levadura). Figura 3B: análisis de la  $\beta$ gal en los pozos. Fusión con el dominio de activación: pJG-scrPB4 (subunidad polII de levadura); fusión con  
45 el dominio de unión al ADN: pEG-scrPB7 (subunidad polIII de la levadura).

Figura 4. Los análisis de Y2H del receptor nuclear humano RXR se muestrearon frente a (aproximadamente) 88 x 10<sup>4</sup> clones de ADNc. Cada pozo de las tres placas de 96 pozos mostradas representa un análisis de la  $\beta$ gal realizado en una cepa de levadura que contenía el plásmido cebo pEG202 RXR unida con aproximadamente 1.000 clones de  
50 levadura de una biblioteca pJG4.5 AD. Los ocho pozos más a la izquierda en cada placa son los controles positivo y negativo (de arriba a abajo de la placa); a) pEG202 x pJGRXR, b) pEG202, c) pEGKREV1 x pJGRAF, d) pEGKREV1 x pJGKRIT1, e) pEGRAS x pJGKRIT1, f) pEGRAS x pJGRAF, g) pEGRXR x pJGGOR4 y h) pEGGOR4 x pJGRXR.

Figura 5. Estrategia de clonación de marco de lectura abierto. Clonación de marcos de lectura abiertos. A. Fusión  
55 con la  $\beta$ gal<sup>+</sup> con modificación del marco de lectura. Se aislaron ADNc cortados aleatoriamente de aproximadamente 600 pares de bases y se clonaron en un gen de  $\beta$ gal con modificación del marco. Las células de *E. coli* transformadas que se convirtieron en  $\beta$ gal<sup>+</sup> contenían un marco de lectura abierto. B. Fusión con la  $\beta$ gal<sup>+</sup> con modificación del marco de lectura en un huésped Sup<sup>+</sup>. En un vector con un codón de terminación suprimible ámbar entre el extremo 3' del ADNc y el extremo 5' de la  $\beta$ gal, se controló la fusión del ADNc con la  $\beta$ gal por el fenotipo Sup de la cepa de *E. coli*.  
60 C. Fusión con el M13gpIII con modificación del marco de lectura en un huésped Sup<sup>+</sup>. El mismo tipo de esquema de clonación se puede adaptar para fusionar el extremo 5' de la proteína de exposición del fago M13 con el ADNc, en este caso el fago viable indicará una clonación exitosa del marco de lectura abierto.

Figura 6. Recodificación dinámica del extremo 3' del dominio de activación de la levadura. Traducción de los  
65 elementos que controlan el T7 (T7CE por sus iniciales en inglés: T7 controlling elementos). Se recodificó el extremo 3' del dominio de activación (AD por su iniciales en inglés: activation domain) de la levadura para incorporar los elementos que controlan la expresión del gen de *E. coli*. Se muestra un ejemplo de recodificación de los elementos de control necesarios para la expresión de la proteína del bacteriófago T7. El dominio de activación de la levadura

## ES 2 284 245 T3

recodificado se usó entonces conjuntamente con el sistema de clonación de marco de lectura abierto para fusionar el marco de lectura correcto al dominio de activación y, simultáneamente, a una proteína de fusión 3' separada (por ejemplo, *βgal* o M13gp3). Leyenda: 1 = levadura; 2 = inserto de ADNc; 3 = modificación del marco.

- 5      Figura 7. Selección fenotípica en *E. coli* de los marcos de lectura abiertos de ADNc y fusión simultánea de ellos con el extremo 3' del dominio de activación (AD) de la levadura y con el extremo 5' del M13gpIII. A. Fragmentos de ADNc de 500 bp clonados aleatoriamente; transformación T7RNAP<sup>+</sup>, Sup<sup>+</sup> *E. coli*; placas de muestreo. B. Exposición de fago en *E. coli*. C. Doble híbrido en levadura: transformación T7RNAP<sup>+</sup>, Sup<sup>+</sup> *E. coli*; transformación M13 en levadura Sup<sup>-</sup>. Leyenda: 1 = levadura; 2 = sitio del inserto; 3 = modificación del marco.

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Un método para detectar una interacción entre una primera proteína de ensayo y una segunda proteína de ensayo, que comprende:

- (a) proporcionar una célula huésped de levadura que contiene un gen indicador, en la que el gen indicador expresa una proteína detectable cuando el gen indicador es activado por una secuencia de aminoácidos que incluye un dominio de activación transcripcional cuando el dominio de activación transcripcional está suficientemente próximo al gen indicador;
- (b) proporcionar un primer gen híbrido que se expresa en la célula huésped de levadura, comprendiendo el primer gen híbrido una secuencia de ADN que codifica una primera proteína híbrida, comprendiendo la primera proteína híbrida:
  - (i) un dominio de unión al ADN que reconoce un sitio de unión en el gen indicador en la célula huésped; y
  - (ii) una primera proteína de ensayo o uno de sus fragmentos, en la que se va a ensayar la interacción con al menos una segunda proteína de ensayo o uno de sus fragmentos;
- (c) proporcionar un segundo gen híbrido que se expresa en la célula huésped de levadura, comprendiendo el segundo gen híbrido una secuencia de ADN que codifica una segunda proteína híbrida, comprendiendo la segunda proteína híbrida:
  - (i) el dominio de activación transcripcional; y
  - (ii) una segunda proteína de ensayo o uno de sus fragmentos, en la que se va a ensayar la interacción con la primera proteína de ensayo o uno de sus fragmentos; en la que la interacción entre la primera proteína de ensayo y la segunda proteína de ensayo en la célula huésped de levadura hace que el dominio de activación transcripcional active la transcripción del gen indicador;
- (d) introducir el segundo gen híbrido en una cepa de célula huésped de levadura de un primer tipo de unión y subsiguientemente introducir dichas células en un medio de ordenamiento, creando de este modo una placa de biblioteca maestra;
- (e) introducir las células de la placa de biblioteca maestra en un segundo medio de ordenamiento que es un medio de ordenamiento para medios líquidos, creando de este modo un conjunto de unión;
- (f) introducir el primer gen híbrido en una cepa de célula huésped de levadura del tipo de unión opuesto a (d) y subsiguientemente introducir dicha célula en el conjunto de unión, permitiendo de este modo que se produzca la unión en un medio líquido; y
- (g) elegir el producto de la interacción de los genes primero y segundo.

2. Un método para detectar una interacción entre una primera proteína de ensayo y una segunda proteína de ensayo según la reivindicación 1, que comprende las etapas adicionales:

- (h) retirar una parte del conjunto de unión a un tercer medio de ordenamiento, creando de este modo un conjunto de liberación;
- (i) determinar si el gen indicador ha sido expresado en el conjunto de unión; y
- (j) analizar las células de la placa de liberación.

3. El método según la reivindicación 2, en el que la célula huésped es *Saccharomyces cerevisiae*.

4. El método según la reivindicación 2, en el que el gen indicador se elige entre el grupo que consiste en LEU2, lacZ, HIS3, URA3, LYS2, GAL1, *E. coli* galK, GFP, CUP1, CAT, G418 y GUS.

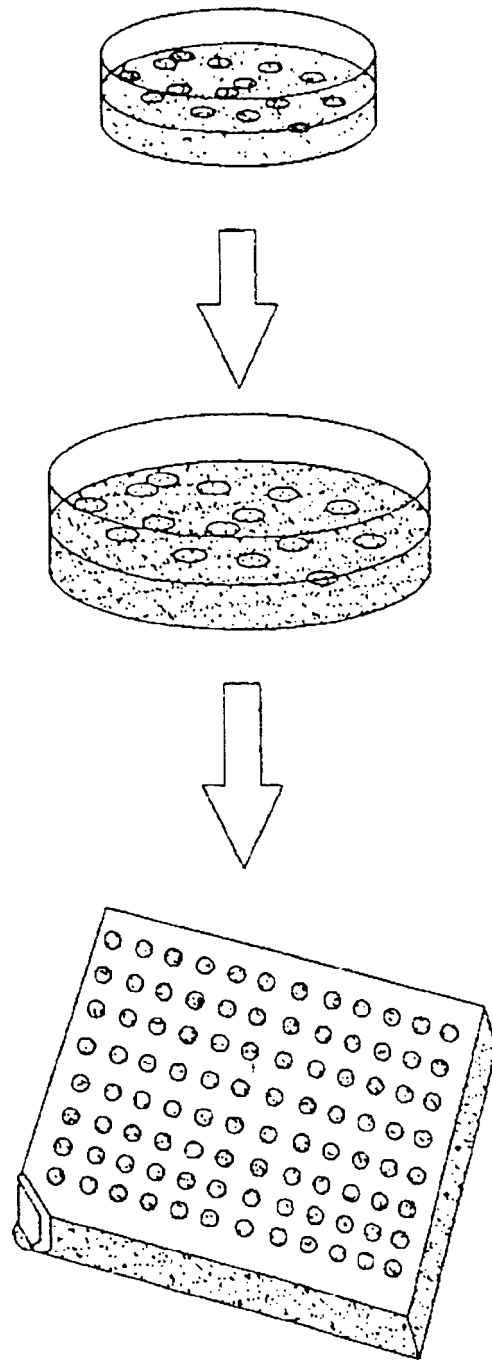


FIG. I

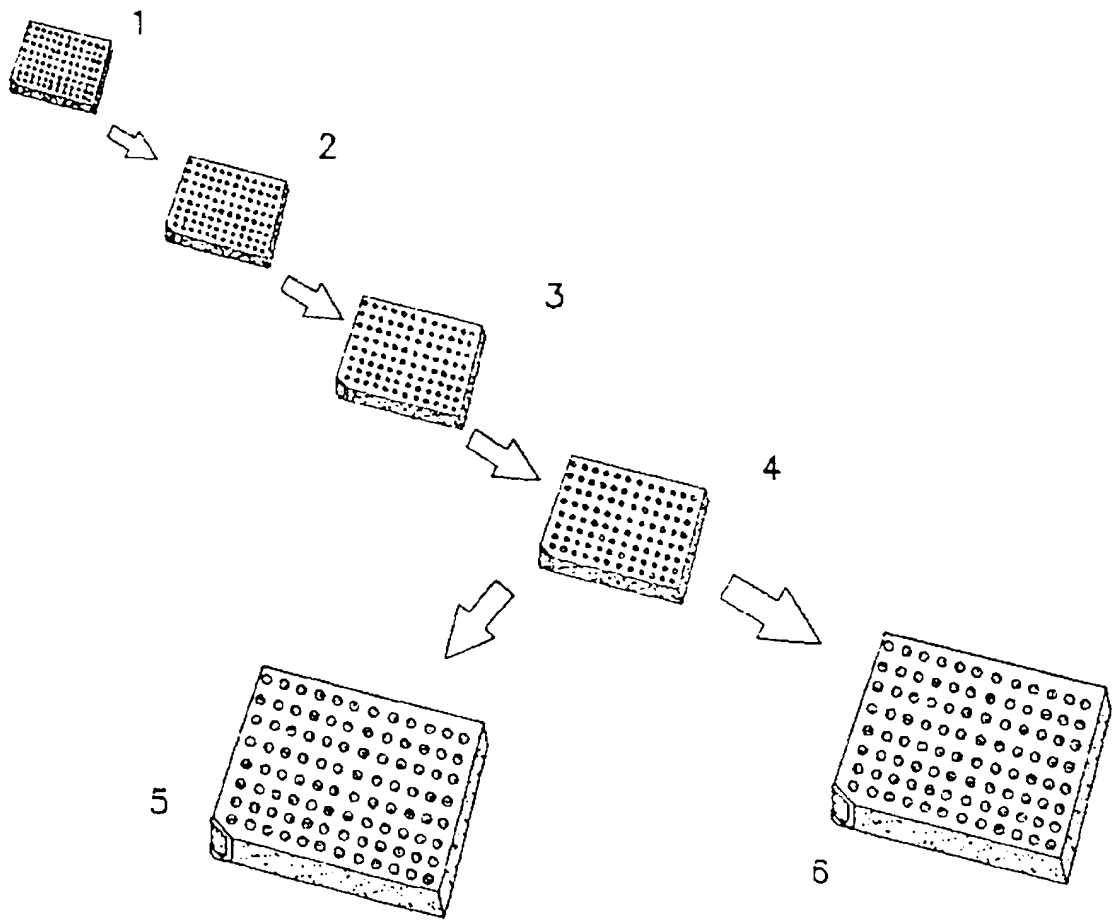


FIG. 2

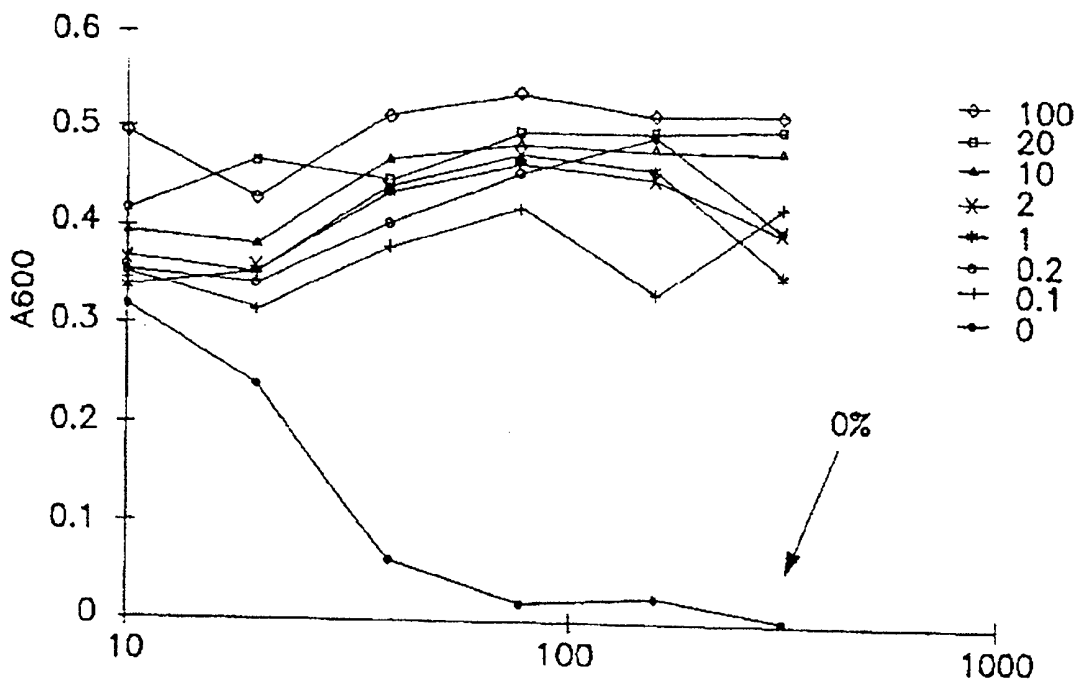


FIG. 3A

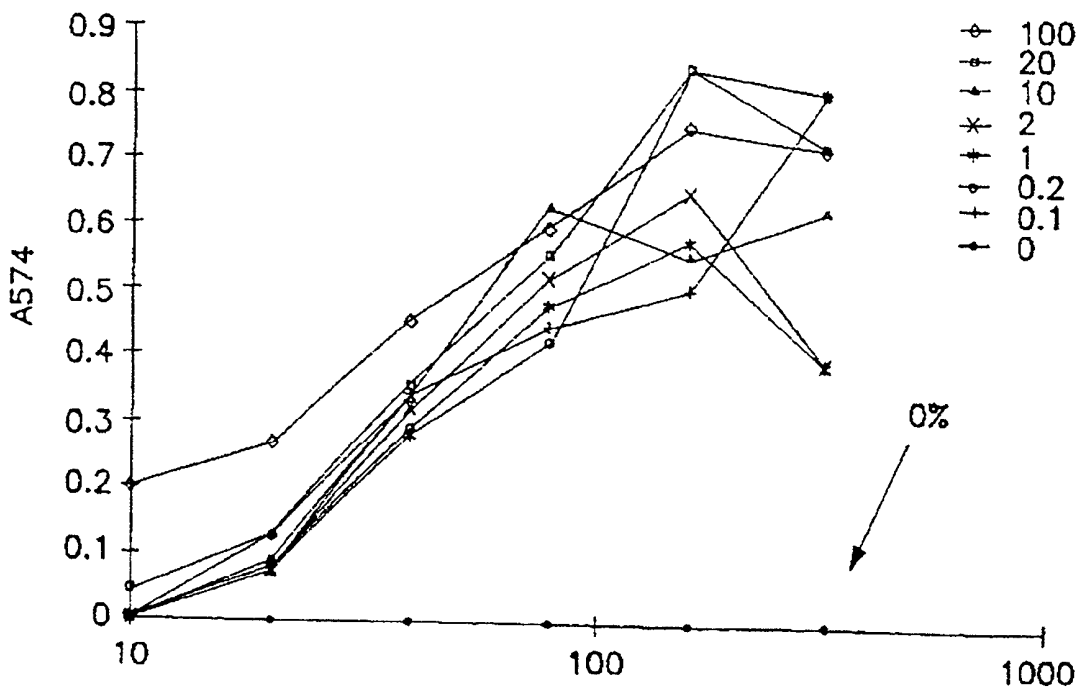


FIG. 3B

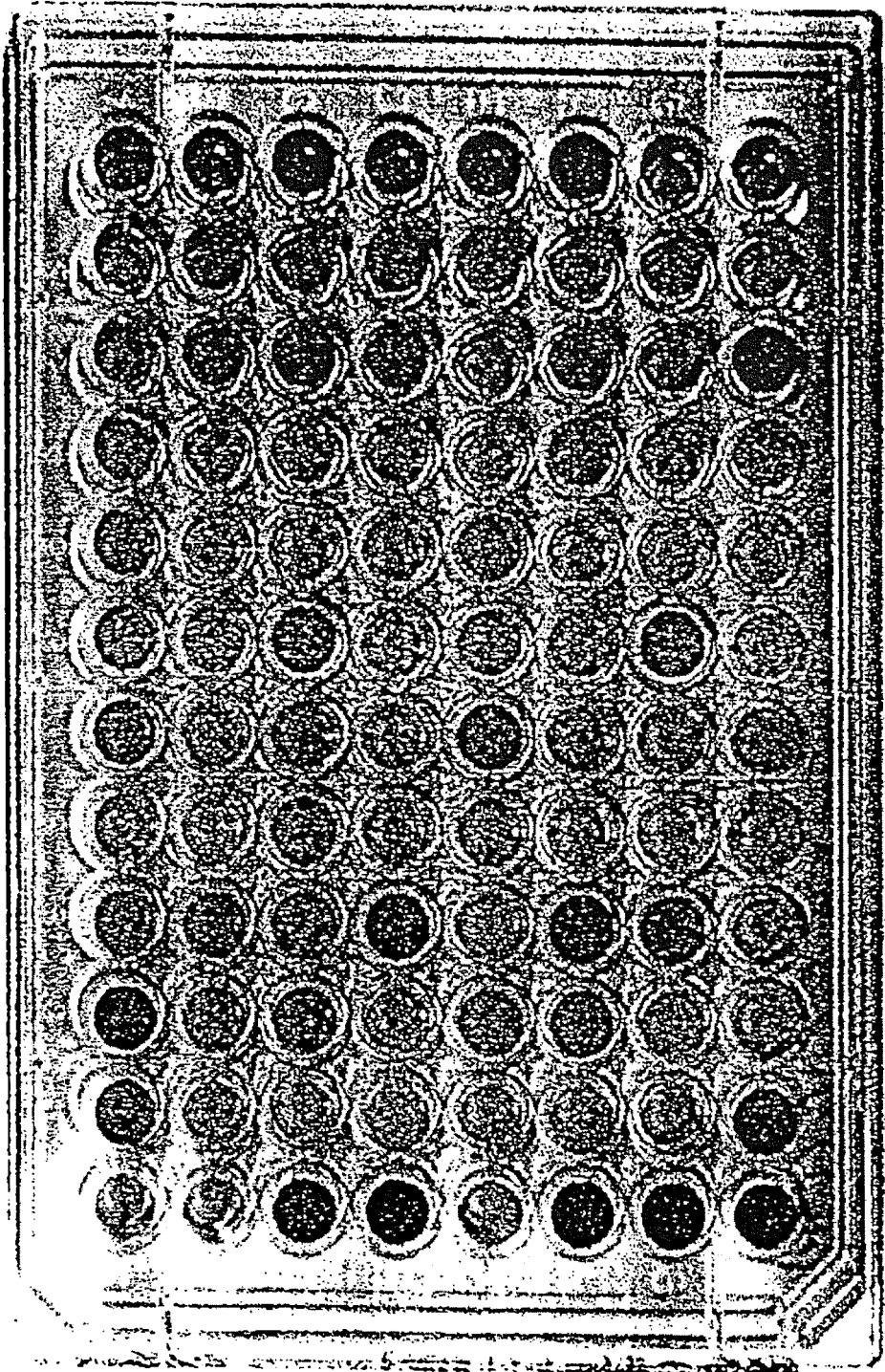


FIG. 4

FIG. 5A

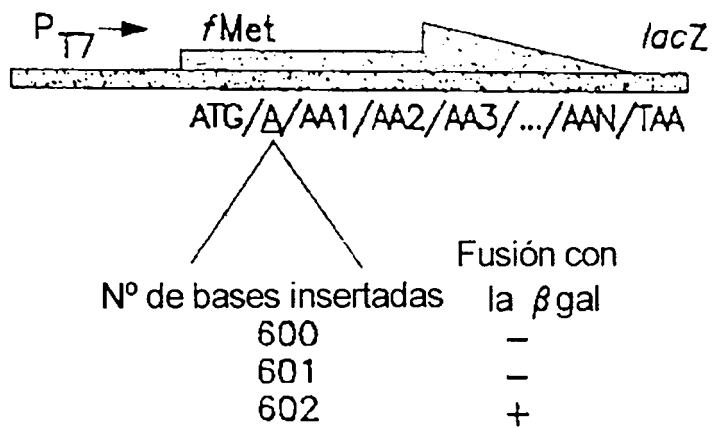


FIG. 5B

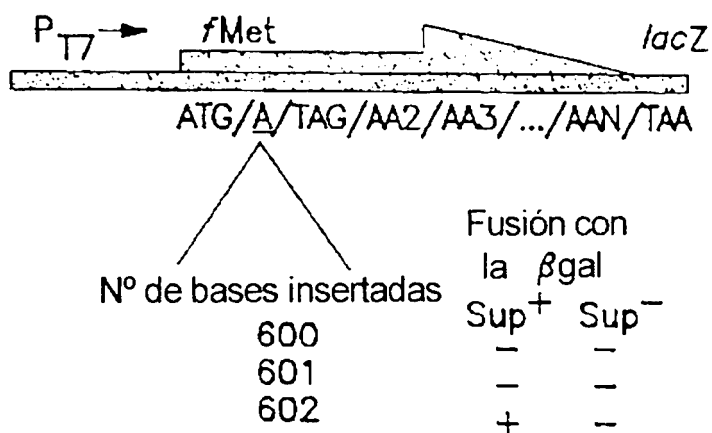
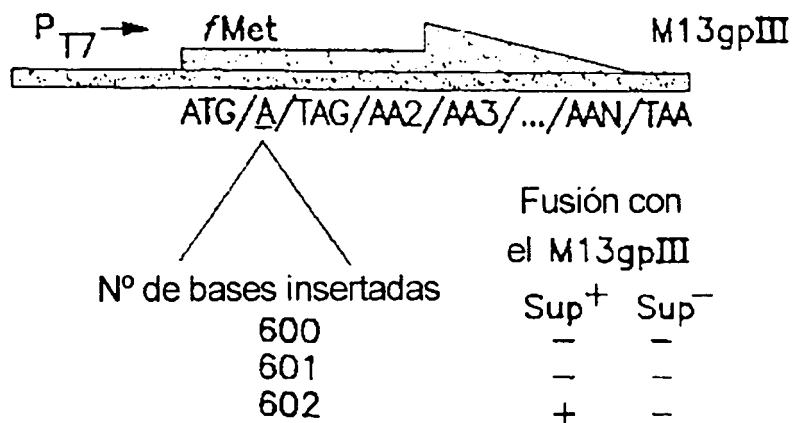


FIG. 5C



PROMOTOR T7-> TTAATACGACTCACTATAGGGAGACCAAAGAAGGAGATATACATCATG... fMet...MCS  
 LeuIleArgLeuThrIleGlyArgProLysLysGluIleTyrIleMet...

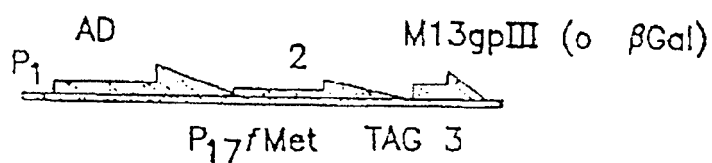
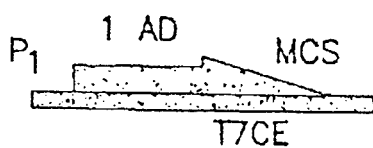


FIG. 6

FIG. 7A

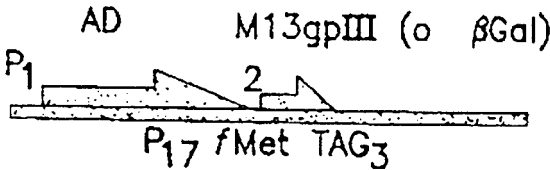


FIG. 7B

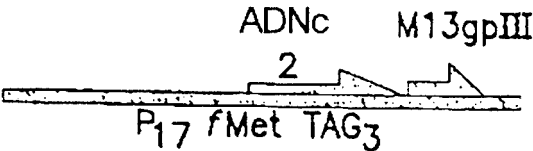


FIG. 7C

