

(12) SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACIÓN EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de la
Propiedad Intelectual
Oficina internacional



(43) Fecha de publicación internacional
16 de noviembre de 2017 (16.11.2017) **WIPO | PCT**

(10) Número de publicación internacional
WO 2017/194810 A1

(51) Clasificación internacional de patentes:

C12Q 1/68 (2006.01) *C12N 15/62* (2006.01)
C12N 15/10 (2006.01)

117, 28006 Madrid (ES). **UNIVERSIDAD AUTÓNOMA DE MADRID** [ES/ES]; Ciudad Universitaria de Cantoblanco, C/ Einstein, 3, 28049 Madrid (ES).

(21) Número de la solicitud internacional:

PCT/ES2017/070298

(72) **Inventor:** VINCENT, Olivier; INSTITUTO DE INVESTIGACIONES BIOMEDICAS ALBERTO SOLS (IIBM), C/ Arturo Duperier, 4, 28029 Madrid (ES).

(22) Fecha de presentación internacional:

11 de mayo de 2017 (11.05.2017)

(74) **Mandatario:** PONS ARIÑO, Ángel; Glorieta de Rubén Dario, 4, 28010 Madrid (ES).

(25) Idioma de presentación:

español

(26) Idioma de publicación:

español

(30) Datos relativos a la prioridad:

P 201630607 11 de mayo de 2016 (11.05.2016) ES

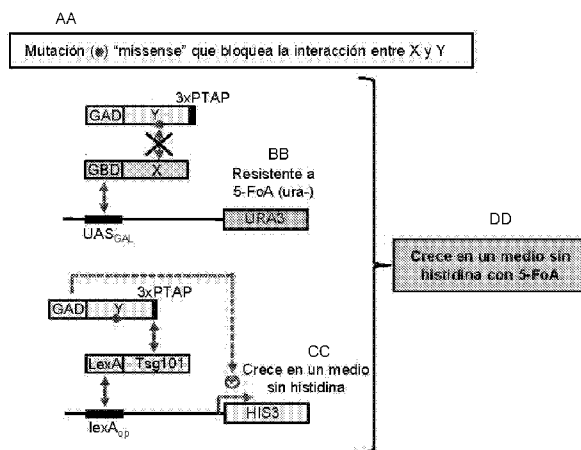
(81) **Estados designados** (a menos que se indique otra cosa, para toda clase de protección nacional admisible): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD,

(71) **Solicitantes:** CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS (CSIC) [ES/ES]; C/ Serrano,

(54) **Title:** REVERSE TWO-HYBRID METHOD FOR THE IDENTIFICATION OF MISSENSE MUTATIONS

(54) **Título:** MÉTODO DE DOBLE HÍBRIDO EN REVERSO PARA LA IDENTIFICACIÓN DE MUTACIONES *MISSENSE*

Fig 1c



AA Missense mutation (-) which blocks the interaction between X and Y
BB Resistant to 5-FoA (ura-)
CC Growth in a histidine-free medium
DD Growth in a medium containing histidine with 5-FoA

(57) **Abstract:** The present invention relates to a reverse two-hybrid method for the identification of missense-type mutations that prevent the binding of two interacting proteins. The invention also describes gene constructs that can be used in the method of the invention, as well as the host cell which comprises said constructs and can be used in the method of the invention.

(57) **Resumen:** La presente invención se refiere a un método de doble híbrido en reverso útil para la identificación de mutaciones de tipo *missense*, es decir, aquéllas mutaciones de cambio de sentido que impiden la unión entre dos proteínas interaccionantes. En la invención se describen además las construcciones génicas útiles en el método de la invención, así como la célula huésped que comprende dichas construcciones y que se utiliza en el método de la invención.

[Continúa en la página siguiente]

ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO,
NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW,
SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Estados designados *(a menos que se indique otra cosa, para toda clase de protección regional admisible)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), euroasiática (AM, AZ, BY, KG, KZ, RU, TJ, TM), europea (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Publicada:

- con informe de búsqueda internacional (Art. 21(3))
- con la parte de lista de secuencias de la descripción (Regla 5.2(a))

MÉTODO DE DOBLE HÍBRIDO EN REVERSO PARA LA IDENTIFICACIÓN DE MUTACIONES *MISSENSE*

DESCRIPCIÓN

5

La presente invención se refiere a un método de doble híbrido en reverso útil para la identificación de mutaciones de tipo *missense*, es decir, aquéllas mutaciones de cambio de sentido que impiden la unión entre dos proteínas interaccionantes. La presente invención describe además las construcciones génicas útiles en el método de la invención, así como la célula huésped que comprende dichas construcciones y que se utiliza en el método de la invención.

10

ESTADO DE LA TÉCNICA

15 El sistema de doble híbrido (Fields S, Song O. Nature. 1989;340(6230):245-6) es una técnica de referencia para analizar las interacciones proteína-proteína. Esta técnica ha sido adaptada a cribados de alto rendimiento ("*high-throughput*") permitiendo caracterizar el interactoma de numerosos organismos, como por ejemplo el interactoma humano (Rolland T. *et al.*, Cell. 2014;159(5):1212-26). Actualmente, dos
20 empresas, Clontech (USA) e Invitrogen (Life Technologies, Thermofisher, USA), comercializan kits del sistema de doble híbrido clásico.

El sistema de doble híbrido clásico ha derivado en el sistema de doble híbrido en reverso (White MA. Proc Natl Acad Sci U S A. 1996;93(19):10001-3). Esta metodología
25 permite la selección de mutaciones que anulan una interacción definida entre un par de proteínas interaccionantes y, en consecuencia, permite también la identificación de los aminoácidos implicados en dicha interacción. La detección de las regiones, y más específicamente, de los aminoácidos concretos implicados en las interacciones proteína-proteína, es necesaria para definir la relevancia fisiológica y las bases
30 moleculares de dichas interacciones. La técnica de doble híbrido en reverso se utiliza también para identificar moléculas que interfieren con dicha interacción proteína-proteína (Vidal M, Endoh H. Trends Biotechnol. 1999;17(9):374-81). Aunque el potencial del sistema de doble híbrido en reverso es evidente, dicho sistema no ha tenido tanto éxito como su predecesor. La razón principal de esta diferencia es que el
35 sistema de doble híbrido en reverso genera un numero demasiado alto de falsos

positivos que corresponden a mutaciones que truncan la proteína, no a mutaciones de cambio de sentido o de tipo *missense*, que son las que interesan al impedir la unión entre dos proteínas interaccionantes, de hecho, >97% de las mutaciones al azar generadas por PCR producen un truncamiento en la proteína.

5

Teniendo en cuenta el alto número de falsos positivos del sistema de doble híbrido en reverso se han propuesto diferentes mejoras para identificar únicamente las mutaciones *missense*, aunque ninguna de dichas mejoras ha conseguido resolver de forma definitiva el problema del alto número de falsos positivos detectados. La primera de dichas mejoras fue dirigida al uso de fusiones C-terminal a la proteína verde fluorescente (GFP-“*Green Fluorescent Protein*”), para seleccionar los mutantes no truncados mediante análisis de la fluorescencia (Endoh H, *et al.*, Methods Enzymol. 2000;328:74-88). Sin embargo, la proteína GFP puede interferir en la interacción entre las proteínas a estudiar y además, el hecho de que la selección mediante fluorescencia se haga a posteriori, incrementa mucho la complejidad de la técnica.

Para evitar el uso de proteínas de tipo GFP y evitar la posible interferencia con la interacción analizada, se propuso utilizar fusiones a epítomos cortos reconocidos por anticuerpos (revisado en Bennett MA. *et al.* Methods Mol Biol. 2015;1278:433-46). De nuevo, este sistema es demasiado largo y complejo ya que requiere validar todos los mutantes mediante la técnica de Western blot. Más recientemente, una fusión a GBD (*Gal4 Binding Domain*) ha sido utilizada en un sistema denominado “one-plus two-hybrid system” (Kim JY, *et al.*, Mol Cell Proteomics. 2007;6(10):1727-40; Kim JY *et al.*; Methods Mol Biol. 2012;812:209-23). En este último caso, el sistema utiliza dos marcadores o reporteros que son reconocidos por dos dominios de unión al ADN distintos, al igual que en estudios previos (Jiang R, *et al.* Genes Dev. 1996;10(24):3105-15; Inouye C. *et al.* Genetics. 1997;147(2):479-92) en los cuales el sistema de doble híbrido fue adaptado para identificar mutaciones que anulan la interacción de una proteína con otra pero no con una tercera. En este sistema, la selección se hace en dos pasos diferentes pero con las mismas células transformadas que comprenden y expresan los reporteros y las proteínas a estudiar. Sin embargo, al igual que la proteína GFP, los dominios de unión al ADN como GBD son dominios estructurales que pueden interferir con la interacción que se está analizando. Este mismo problema tampoco ha sido superado por el método descrito por el equipo de Solomon MJ, que utiliza fusiones que comprenden el gen que codifica la proteína a

analizar unida en su extremo C-terminal al gen *URA3* para detectar las mutaciones que truncan la proteína estudiada y en consecuencia no permiten la síntesis y expresión de la proteína de fusión formada por la proteína a analizar y Ura3. Específicamente, se clona el gen *URA3* en fase con el gen que codifica la proteína a estudiar de tal forma que cuando dichos genes se expresan, se produce una proteína de fusión o proteína quimérica (proteína a estudiar-Ura3). La presencia de una mutación de truncamiento en la proteína estudiada impide la expresión y síntesis de la fusión de dicha proteína a Ura3, de tal forma que una levadura que comprenda el gen *URA3* endógeno inactivado no crecerá en un medio de cultivo en ausencia de uracilo (Burton JL. *et al.* EMBO J. 2011 May 4;30(9):1818-29; Lickfeld M. *et al.* Yeast. 2011 Jul;28(7):535-45).

Otro de los métodos alternativos descritos (Gray PN, *et al.* Mol Cell Proteomics. 2007 Mar;6(3):514-26; WO2006060595), y comercializados en forma de kit por Invitrogen, propuso la expresión de una proteína de fusión que comprende el gen mutado de una de las proteínas del par interaccionante a estudiar unido en el extremo C-terminal a un marcador de resistencia a antibióticos, por lo que era necesaria una doble selección en bacterias y levaduras dotando a la técnica de una gran complejidad metodológica. Dicha tecnología no consiguió superar las carencias antes mencionadas, ya que mediante dicha técnica se seguían detectando al menos un 50% de falsos positivos correspondientes a proteínas no mutadas.

En este sentido, sigue existiendo en el estado de la técnica la necesidad de sistemas de doble híbrido en reverso para la detección exclusiva de mutaciones *missense* que anulan la unión entre proteínas interaccionantes, donde dichos sistemas sean de una simplicidad técnica tal que permita su uso rutinario tanto en clínica como en investigación, y además, sin dar lugar a la detección de falsos positivos.

DESCRIPCIÓN DE LA INVENCION

30

La presente invención soluciona el problema de la técnica mencionado previamente mediante un nuevo sistema de doble híbrido en reverso capaz de detectar exclusivamente las mutaciones *missense* que impiden la unión entre un par de proteínas interaccionantes de las que se desee conocer las regiones concretas implicadas en su unión. El sistema de doble híbrido en reverso de la invención utiliza

35

un plásmido que comprende la secuencia nucleotídica que codifica para un péptido capaz de interactuar con una proteína heteróloga expresada desde una construcción génica que se ha integrado en el genoma de la célula utilizada en el sistema de la invención, y donde la secuencia nucleotídica que codifica para dicho péptido está fusionada a la secuencia nucleotídica que codifica para una de las proteínas interaccionantes sometidas a estudio. La transformación de la célula utilizada en el método de doble híbrido en reverso de la invención con el plásmido que codifica para el péptido que se une a dicha proteína heteróloga expresada desde una construcción génica integrada en el genoma de dicha célula, junto con un sistema dual de genes reporteros de selección positiva y contra-selección, permite la identificación positiva e inequívoca de las mutaciones *missense* que impiden la interacción entre el par de proteínas a estudiar.

La utilización del péptido mencionado anteriormente en lugar de proteínas o dominios proteicos, tal y como se utiliza actualmente en el estado de la técnica para la detección de mutaciones *missense*, tiene la ventaja de reducir una posible interferencia de la fusión en los extremos C-terminal de las proteínas del par interaccionante a estudiar.

Adicionalmente, otra de las ventajas del método de doble híbrido en reverso de la invención es la identificación de los mutantes *missense* en una única etapa, mediante el uso de la técnica de PCR mutagénica unida a la técnica de recombinación *in vivo* ("*gap-repair*") y el sistema dual de reporteros que permite la selección simultánea de mutaciones que son *missense* y que provocan la pérdida de interacción entre el par de proteínas a estudiar.

Tal y como se muestran en los ejemplos incluidos en el presente documento, el sistema de doble híbrido en reverso de la invención para la identificación de mutaciones *missense* entre un par de proteínas interaccionantes sometidas a estudio, además de ser un sistema rápido y sencillo, es un sistema muy eficiente, llegando a identificar correctamente el 100% de las mutaciones *missense* obtenidas, sin dar lugar a falsos positivos.

Por tanto, la presente invención proporciona un método de doble híbrido en reverso capaz de identificar mutaciones *missense* que impiden la unión entre dos proteínas interaccionantes en una única etapa. La invención también presenta células,

preferentemente cepas de levadura y varias construcciones génicas que son útiles para la identificación de las mutaciones que inhiben las interacciones moleculares entre las proteínas a estudiar mediante el método descrito aquí.

5 Método de la invención

En un primer aspecto, la presente invención se refiere a un método *in vitro* para identificar al menos una mutación en una proteína de referencia donde dicha mutación afecta a la capacidad de dicha proteína de referencia de unirse a otra proteína diana, donde dicho método comprende:

a) al menos una célula huésped que comprende:

- i) Una primera secuencia nucleotídica que codifica para un gen reportero, donde dicha secuencia nucleotídica está operativamente unida a un promotor que comprende una secuencia de reconocimiento para una proteína que se une al ADN,
- ii) Una segunda secuencia nucleotídica que codifica para un segundo gen reportero, donde dicha segunda secuencia nucleotídica está operativamente unida a un promotor que comprende una secuencia de reconocimiento para una proteína que se une al ADN, con la condición de que dicha secuencia de reconocimiento es diferente de la secuencia de reconocimiento de i), y
- iii) Una tercera secuencia nucleotídica que codifica para una primera proteína de fusión que comprende el dominio de unión al ADN de ii) y una proteína heteróloga capaz de unirse a un péptido funcional localizado en el extremo C-terminal de la proteína de referencia de estudio, donde la secuencia nucleotídica que codifica para dicha primera proteína de fusión está operativamente unida a un promotor,

b) Pre-transformar la célula de la etapa a) con un plásmido que comprende la secuencia nucleotídica que codifica para una segunda proteína de fusión que comprende el dominio de unión al ADN de i) y la proteína diana, donde la secuencia nucleotídica que codifica para dicha segunda proteína de fusión está operativamente unida a un promotor,

- c) Cultivar la célula pre-transformada de la etapa b) bajo condiciones que permitan exclusivamente el crecimiento de las células que hayan incorporado el plásmido de la etapa b),
- 5 d) Transformar la célula de la etapa c) con un vector linearizado y al menos un fragmento de ADN que previamente ha sido sometido a mutagénesis y que comprende la secuencia nucleotídica que codifica para la proteína de referencia con al menos una mutación, donde entre el vector linearizado y el fragmento de ADN sometido a mutagénesis se produce recombinación
- 10 homóloga obteniéndose un plásmido que comprende la secuencia nucleotídica que codifica para una tercera proteína de fusión que comprende el dominio de transactivación de Gal4 y la proteína de referencia con al menos una mutación, donde la secuencia nucleotídica que codifica para dicha tercera proteína de fusión está operativamente unida a un promotor,
- 15 e) Cultivar la célula de la etapa d) bajo condiciones tales que permitan exclusivamente el crecimiento de las células que presentan mutaciones *missense* que impiden la unión entre la proteína de referencia y la proteína diana,
- 20 f) Comparar la secuencia de la proteína de referencia mutada con la secuencia de la proteína de referencia sin mutar (nativa o *wild-type*) e identificar la mutación que impide la unión de la proteína de referencia con la proteína diana.

25

A efectos de la presente invención, los términos proteína, o molécula de ADN “de referencia”, se refiere a aquella molécula capaz de unirse o asociarse de forma transitoria o constitutiva con otra proteína, o molécula de ADN a la que a efectos de la presente invención se le denomina “diana”. En los ejemplos que se muestran en el

30 presente documento para poner de manifiesto la validez del método de la invención se ha utilizado como proteína de referencia la proteína humana glucoquinasa (GK) de SEQ ID NO: 10, codificada por la secuencia nucleotídica SEQ ID NO: 9, y como proteína diana a la que se une la proteína de referencia, se ha utilizado la proteína reguladora de la glucoquinasa humana (GKRP) de SEQ ID NO: 2, codificada por la

35 secuencia nucleotídica SEQ ID NO: 1. Mediante el método de la invención se puede

utilizar cualquier par de proteínas interaccionantes de las que se quiera analizar la presencia de mutaciones *missense* que impidan su interacción.

5 A efectos de la presente invención, el término “dominio de unión al ADN” se entiende como un conjunto de amino ácidos que es capaz de dirigir la unión específica de un polipéptido a una secuencia particular de ADN, por ejemplo a una secuencia de reconocimiento para un factor de transcripción que interactúa específicamente con una secuencia de ácido nucleico en el promotor de un gen. Alternativamente, el dominio de unión al ADN puede ser cualquier dominio proteico que interactúa específicamente con
10 una secuencia que se puede producir de forma natural o se inserta artificialmente en el promotor de un gen reportero. El dominio de unión al ADN puede estar unido covalentemente a un dominio de transactivación, de manera que la unión de la proteína al ADN en un sitio localizado dentro del promotor de un gen reportero elegido, activa su transcripción.

15 A efectos de la presente invención, existe una gran variedad de dominios de unión al ADN y de dominios de transactivación que son adecuados para su uso en los diversos aspectos de la invención. En general, cualquier dominio de unión al ADN o cualquier dominio de transactivación de cualquier factor de transcripción puede ser utilizado en
20 la presente invención. El dominio de unión al ADN y el dominio de transactivación pueden pertenecer o no a distintos factores de transcripción. Secuencias de reconocimiento útiles incluyen, sin ser limitativos, los sitios de unión de los factores de transcripción Gal4 y Ace1 de levadura y LexA de bacteria. Estos sitios de unión pueden utilizarse fácilmente con un promotor reprimido (por ejemplo, los promotores
25 SPAL, SPEX y ESPACE combinan un promotor *SP013* con las secuencias de reconocimiento para Gal4, Ace1 y LexA, respectivamente). Otros factores de transcripción útiles incluyen la proteína Gcn4 de *S. cerevisiae* (véase, por ejemplo, Hope and Struhl, 1986, Cell 46: 885-894) y la proteína Adr1 de *S. cerevisiae* (véase, por ejemplo, Kumar *et al*, 1987, Cell 51: 941-951). La secuencia de reconocimiento
30 debe incluir al menos un sitio de unión para el dominio de unión al ADN del factor de transcripción que se utiliza. El número de sitios de unión se puede ajustar para proporcionar mayor o menor sensibilidad al ensayo.

Por "secuencia de reconocimiento" se entiende un segmento de ADN que es necesario y suficiente para interaccionar específicamente con un polipéptido dado, tal como por ejemplo, el dominio de unión al ADN de un factor de transcripción.

- 5 Por "unido operativamente" se entiende que un gen y una secuencia reguladora (s) (por ejemplo, un promotor) están conectados de tal manera que permiten la expresión génica cuando las moléculas apropiadas (por ejemplo, proteínas que incluyen dominios que activan la transcripción) se unen a las secuencias reguladoras.
- 10 Por "unido covalentemente" se entiende cuando dos moléculas, por ejemplo proteínas, están unidas directamente por uniones covalentes. Por ejemplo, proteínas o dominios proteicos unidos covalentemente se pueden localizar inmediatamente contiguos, o pueden localizarse separados por residuos de uno o más aminoácidos dentro de la misma proteína híbrida.

15

- Tal y como se utiliza en la presente invención, el término "gen reportero" utilizado a lo largo del documento, se refiere a un gen cuya expresión puede analizarse como una medida de la capacidad de dos moléculas de ensayo para interactuar (es decir, como una medida de las interacciones proteína/proteína, proteína/ARN, ARN/ARN o
- 20 proteína/ADN). Los genes reporteros descritos en este documento pueden localizarse en un plásmido o pueden integrarse en el genoma de una célula haploide o diploide. El gen reportero se une preferentemente de manera operativa a un promotor que tiene secuencias que dirigen la transcripción del gen reportero. El gen reportero está posicionado de tal manera que se expresa cuando un dominio de transactivación de
- 25 un factor de transcripción se pone en estrecha proximidad con el gen (por ejemplo, mediante el uso de proteínas híbridas para reconstituir un factor de transcripción, o por unión covalente del dominio de transactivación a un dominio de unión al ADN). El gen reportero también puede estar unido operativamente a secuencias reguladoras que le hacen muy sensible a la presencia o ausencia de un factor de transcripción. Por
- 30 ejemplo, en ausencia de un factor de transcripción específico, el gen reportero *URA3* unido operativamente a un promotor altamente sensible confiere un fenotipo *Ura⁻ Foa^r* en la célula. En presencia de un factor de transcripción específico, esta construcción génica confiere un fenotipo *Ura⁺ Foa^s* en la célula. Métodos conocidos por el experto medio en el presente campo técnico pueden ser utilizados para conectar un gen
- 35 reportero a un promotor y para introducir el gen reportero en una célula.

A efectos de la presente invención, un gen reportero útil tiene en su promotor una secuencia de reconocimiento para el dominio de unión al ADN de un factor de transcripción que puede ser reconstituido mediante la interacción de una proteína con un dominio de unión al ADN y otra proteína con un dominio de transactivación. Tales genes incluyen, sin limitación, *lacZ*, genes biosintéticos aminoacídicos tales como por ejemplo, los genes de levadura *LEU2*, *HIS3*, *LYS2*, *LYS5*, o *TRP1*, el gen *URA3*, genes biosintéticos de ácidos nucleicos, el gen que codifica para la cloranfenicol-transacetilasa bacteriana (*cat*), y el gen bacteriano *gus*. También se incluyen los genes que codifican marcadores fluorescentes, tales como el gen de la proteína fluorescente verde (GFP). Ciertos genes reporteros son considerados como genes reporteros "seleccionables" o "contra-seleccionables" tal y como se describe a continuación.

En una realización más preferida, el método de doble híbrido en reverso según se describe en la presente invención se caracteriza por que la célula utilizada en dicho método comprende al menos dos genes reporteros, donde el primero de dichos genes es un gen de selección positiva y el segundo de dichos genes es un gen de contra-selección.

Por marcador "seleccionable" se entiende un gen que, cuando se expresa, confiere una ventaja de crecimiento sobre una célula que lo contiene. Ejemplos de marcadores seleccionables incluyen, sin limitación, *LEU2*, *TRP1*, e *HIS3*. Ciertos marcadores seleccionables descritos en este documento pueden usarse para promover el crecimiento de células que comprenden un plásmido que expresa este marcador seleccionable. En este caso, el promotor unido operativamente a este marcador seleccionable es el promotor natural para el marcador. Por otra parte, el marcador puede ser diseñado para ser unido operativamente a un promotor distinto de aquel al que está naturalmente ligado. Así, en el caso de los genes reporteros utilizados en esta invención, el marcador seleccionable está unido operativamente a un promotor con una secuencia de reconocimiento para un factor de transcripción que puede ser reconstituido mediante la interacción de una proteína con dominio de unión al ADN y otra proteína con dominio de transactivación.

Por marcador "contra-seleccionable" se entiende un gen que, cuando se expresa, impide el crecimiento de la célula que lo contiene. Ejemplos de marcadores contra-seleccionable incluyen *URA3*, *LYS2*, *GAL1*, *CYH2* y *CAN1*.

Por gen reportero "seleccionable" se entiende un gen reportero que, cuando se expresa bajo un determinado conjunto de condiciones, confiere una ventaja de crecimiento sobre las células que lo contienen. Ejemplos de genes reporteros seleccionables incluyen los marcadores *LEU2*, *TRP1*, e *HIS3*. En la presente invención, el gen reportero seleccionable preferido es el gen *HIS3*.

Por gen reportero "contra-seleccionable" se entiende un gen reportero que, cuando se expresa bajo un determinado conjunto de condiciones, impide el crecimiento de una célula que lo contiene. Ejemplos de genes reporteros contra-seleccionables incluyen los marcadores *URA3*, *LYS2*, *GAL1*, *CYH2* y *CAN1*. En la presente invención, el gen reportero contra-seleccionable preferido es el gen *URA3*.

Por gen reportero "detectable" se entiende un gen cuya expresión puede ser detectada en una célula por un medio distinto al de conferir una ventaja de crecimiento selectivo en una célula. Un ejemplo de un gen reportero detectable es el gen *lacZ*. Si se desea, un gen reportero detectable puede ser integrado en el genoma de una célula, preferentemente de una célula de levadura. Un gen reportero detectable se puede utilizar en la invención para medir la capacidad de dos moléculas de interactuar. Así, el promotor que está unido operativamente a un gen reportero detectable debe contener una secuencia de reconocimiento para un factor de transcripción que puede ser reconstituido mediante la interacción de una proteína con dominio de unión al ADN y otra proteína con dominio de transactivación.

Preferiblemente, cada uno de los genes reportero esta operativamente unido a un promotor que porta una secuencia represora que impide la transcripción en ausencia de un motivo de activación génica. Por lo tanto, el gen reportero debe colocarse de tal manera que su expresión sea altamente sensible a la presencia o ausencia de un factor de transcripción. Por ejemplo, se prefiere que cuando se utiliza el gen reportero que codifica para el alelo *URA3*, dicho alelo confiere un fenotipo $\text{Ura}^- \text{Foa}^r$ en ausencia del factor de transcripción, y un fenotipo $\text{Ura}^+ \text{Foa}^s$ en su presencia. Ciertos promotores, tales como el promotor *SP013*, contienen de forma natural una secuencia represora aguas arriba. Otros promotores pueden ser diseñados o modificados utilizando métodos convencionales de clonación, para que comprendan dichas secuencias. Cuando se utiliza un gen reportero de contra-selección, la expresión del

gen puede ser detectada mediante la detección de la inhibición del crecimiento celular. Cuando se emplea más de un gen reportero, los genes reporteros pueden estar unidos operativamente a promotores que son idénticos entre sí sólo en sus secuencias de reconocimiento. Preferiblemente, el gen reportero es uno que permite la selección titulable; por lo tanto, el crecimiento celular se puede medir en un rango de condiciones (por ejemplo, concentraciones de 5-FoA).

Marcadores "*contra-seleccionables*": Mientras que los marcadores seleccionables se han utilizado para, bajo ciertas condiciones, promover el crecimiento de sólo aquellas células que expresan los marcadores seleccionables, el marcador *contra-seleccionable* se han utilizado, bajo ciertas condiciones, para promover el crecimiento de sólo aquellas células que no expresan el marcador *contra-seleccionable*. Los marcadores *contra-seleccionables* cuando están presentes en los plásmidos se pueden utilizar para seleccionar las células que han perdido el plásmido. Por ejemplo, la expresión del gen *URA3*, que codifica para la orotidina-5'-fosfato, es letal en presencia de un medio que contiene ácido 5-fluoro-orótico (5-FOA). Las células que expresan *URA3* también se pueden seleccionar positivamente haciéndolas crecer en un medio sin uracilo. Por lo tanto, dependiendo de las condiciones de crecimiento, el marcador *URA3* se puede utilizar ya sea para seleccionar de manera positiva o negativa. El gen *LYS2*, que codifica la α -aminoadipato reductasa, también se puede utilizar para *contra-selección*. Las células de levadura que expresan *LYS2* no crecen en un medio que contiene α -aminoadipato como fuente de nitrógeno primario. Del mismo modo, la expresión de *LYS5* en un medio que contiene α -aminoadipato es letal. Estos genes, que están implicados en la biosíntesis de lisina, se pueden seleccionar de una manera positiva en un medio libre de lisina. Otro marcador *contra-seleccionable* es el gen *CAN1* que codifica una permeasa de arginina. La expresión de este gen en ausencia de arginina y en presencia de canavanina es letal. Del mismo modo, la expresión del marcador de *contra-selección* *CYH2* es letal en presencia de cicloheximida. La expresión de un gen marcador *contra-seleccionable* se ha utilizado para identificar las mutaciones en el dominio de activación de receptor de estrógeno que inhiben su capacidad para activar la transcripción (Pierrat *et al*, 1992, Gene 119: 237-245).

En otra realización preferida del método de la invención, este se caracteriza por que los genes reporteros de selección positiva se seleccionan del grupo que consiste en

cualquiera de los siguientes marcadores seleccionables *HIS3*, *LEU2*, *URA3*, *ADE2*, *TRP1*, *LYS2* y *LYS5*. En otra realización más preferida aún, el gen reportero de selección positiva es preferentemente el gen *HIS3*.

- 5 En otra realización preferida, los genes reporteros de contra-selección se seleccionan del grupo que consiste en cualquiera de los marcadores contra-seleccionables *URA3*, *TRP1*, *LYS2*, *LYS5*, *CYH2*, *CAN1*, *GAL1* y *mazF*. En otra realización más preferida aún, el gen reportero de contra-selección es preferentemente el gen *URA3*.
- 10 En otra realización preferida, si se desea, la construcción génica con el gen reportero (por ejemplo, *SPAL10:URA3* o *(lexAop)4:HIS3*) se puede integrar en el genoma de una célula haploide o diploide. En una realización más preferida aún, se pueden integrar en el genoma de una célula haploide o diploide, más de un gen reportero, preferiblemente al menos cuatro genes reporteros, al menos tres genes reporteros y
- 15 más preferiblemente, al menos dos genes reporteros. Si se desea, una combinación de genes reporteros se puede integrar cromosómicamente en el genoma de la célula y otros genes reporteros pueden localizarse en un plásmido episomal y expresarse a partir del mismo.
- 20 En una realización preferida del método de doble híbrido en reverso descrito en la presente invención, este se caracteriza por que la célula utilizada en dicho método comprende las secuencias de reconocimiento para proteínas que se unen al ADN seleccionadas de entre cualquiera de los siguientes factores de transcripción: Gal4, LexA y Ace1.
- 25 A efectos de la presente invención, el método de doble híbrido en reverso se caracteriza por que la célula utilizada en dicho método comprende integrado en su genoma al menos dos genes reporteros, preferentemente uno de selección positiva y otro de contra-selección. En otra realización más preferida aún, el gen reportero de
- 30 selección positiva es *HIS3* y el gen reportero de contra-selección es *URA3*.

En otra realización preferida, la célula utilizada en el método de la invención también se caracteriza por que comprende, preferentemente integrado en su genoma vía transformación con un plásmido integrativo, la secuencia nucleotídica que codifica

35 para una primera proteína de fusión, donde dicha primera proteína de fusión

comprende el mismo dominio de unión al ADN que el comprendido en el apartado ii), y una proteína heteróloga que se une al péptido utilizado en el método de la invención y que está comprendido en un plásmido que comprende a su vez la secuencia nucleotídica que codifica para la proteína de referencia sometida a estudio. En una
5 realización más preferida aún, el plásmido integrativo comprende las secuencias nucleotídicas de la construcción génica donde el promotor de *ADH1* esta operativamente unido a la secuencia nucleotídica que codifica para la fusión del dominio de unión al ADN de LexA con la proteína heteróloga que se une al péptido de la invención. A efectos de la presente invención y a modo de ejemplo, sin querer ser
10 limitativo se ha utilizado como proteína heteróloga expresada desde una construcción génica integrada en el genoma de la célula de la invención, la proteína humana TSG101 (SEQ ID NO: 32, codificada por la secuencia nucleotídica SEQ ID NO: 31). Así, la célula de la invención expresa una primera proteína de fusión que comprende el dominio de unión al ADN de LexA y TSG101 (LexA-TSG101), siendo su secuencia
15 nucleotídica la SEQ ID NO: 48 que codifica para la proteína de fusión de SEQ ID NO: 49.

La proteína TSG101 (del inglés Tumor Susceptibility Gene 101 pertenece a un grupo de enzimas aparentemente inactivas para conjugación con ubiquitina (ubiquitin-
20 conjugating enzymes); en la presente invención TSG101 preferiblemente se refiere al Gene ID: 7251 de *Homo sapiens* aunque cualquier experto en el estado de la técnica entenderá que es posible el uso de otras secuencias procedentes de otras especies, preferiblemente mamíferos, más preferiblemente mamíferos primates, que son homólogas a la secuencia humana. Dicha proteína forma parte del complejo ESCRT-I
25 que reconoce las proteínas ubiquitinadas en la ruta endocítica y está implicada en la formación del cuerpo multivesicular (MVB). TSG101 es a su vez capaz de interaccionar con una pequeña secuencia de un péptido denominado P(S/T)TAP que está presente en la proteína Gag p6 del virus VIH (PNAS (2001) 98, 7724-9; Cell (2001) 107, 55-65). Se ha observado que una secuencia que comprende una triple
30 repetición del péptido PTAP, que denominamos 3xPTAP, a lo largo del presente documento, presente de forma natural en una cepa de VIH aislada (GenBank: ACS76886.1) interacciona con TSG101 más fuertemente que una única repetición del péptido PTAP aislado. A efectos de la presente invención, para aumentar la fuerza de interacción de la proteína de fusión LexA-TSG101 (SEQ ID NO: 49) se ha utilizado
35 esta triple repetición del péptido PTAP (3xPTAP) de SEQ ID NO: 12. Se podrían

utilizar otras secuencias que comprendan el péptido P(S/T)AP y que interaccionan con TSG101, u otras combinaciones de proteína y péptido que interaccionan entre si. A modo de ejemplo, se podría utilizar la combinación del péptido YPX(L/I) y la proteína humana ALIX o cualquier homólogo de esta proteína en otros organismos que une este péptido (por ejemplo la proteína PalA del hongo filamentoso *Aspergillus nidulans*) (Mol Cell Biol (2003) 23, 1647-55).

El péptido 3xPTAP (SEQ ID NO: 12) une con gran afinidad la proteína humana TSG101 (SEQ ID NO: 32) en el sistema de doble híbrido en reverso utilizado en la presente invención. De esta manera, se ha llevado a cabo la fusión del péptido 3xPTAP en el extremo C-terminal de la proteína de referencia (que es sometida a mutagénesis), para identificar mutaciones que anulan la interacción con una proteína diana, sin afectar a la interacción mediada por el péptido 3xPTAP con la proteína TSG101, lo que permite descartar todos los falsos positivos que corresponden a mutaciones que truncan la proteína. La proteína de fusión con el péptido 3xPTAP en el extremo C-terminal de la proteína diana mutada, se encuentra fusionada al dominio de transactivación de Gal4 (GAD) en su extremo N-terminal, permitiendo la interacción en el sistema de doble-híbrido de la invención con la proteína de fusión LexA-TSG101 integrada en el cromosoma de la célula, y en consecuencia la activación del reportero (*lexAop*)4-*HIS3*. Este sistema permite una selección positiva e inequívoca de las mutaciones *missense* ya que todas la mutaciones que truncan, inestabilizan o impiden la entrada de la proteína en el núcleo bloquean su interacción con la proteína heteróloga TSG101 (SEQ ID NO: 32) y en consecuencia la activación del gen reportero *HIS3*, lo que se traduce en una falta de crecimiento en un medio sin histidina (ver **Figura 1**). La selección de los mutantes que han incorporado las mutaciones *missense*, se realiza en un solo paso, con un sistema doble de reporteros, tal y como hemos mencionado anteriormente, el primer reportero permite la selección de los mutantes que pierden la interacción, mientras que el segundo selecciona de forma simultánea los mutantes que no truncan la proteína.

En una realización más preferida aún, la célula utilizada en el método de la invención comprende, preferentemente integrado en su genoma los reporteros LexAop-HIS3 ((*lexAop*)4-*HIS3*) y *UASGal-URA3* (*SPAL10::URA3*), que codifican para los marcadores de selección positiva y contra-selección *HIS3* y *URA3*, respectivamente.

Adicionalmente, tal y como se ha mencionado anteriormente, la célula utilizada en el

método descrito en la presente invención comprende la construcción génica donde el promotor de *ADH1* esta operativamente unido a la secuencia nucleotídica que codifica para la fusión del dominio de unión al ADN de LexA con la proteína heteróloga, preferentemente dicha proteína heteróloga es la proteína humana TSG101, que se une al péptido de la invención. Así, la célula de la invención expresa la proteína de fusión LexA-TSG101 de SEQ ID NO: 49.

Los términos “hibrida” o “de fusión”, en relación a proteínas o ADN, hacen referencia a una quimera de al menos dos polipéptidos, o dos moléculas de ADN, unidas covalentemente.

Por “dominio de transactivación”, se entiende al conjunto de aminoácidos capaces de inducir la expresión de un gen de la región a cuyo promotor está unido.

A efectos de la presente invención se entiende por “péptido”, “etiqueta”, “epítipo”, “cola”, o “tallo”, funcional en el extremo C-terminal de la proteína de referencia, a un conjunto de amino ácidos ubicados en el extremo C-terminal de la proteína de referencia. A efectos de la presente invención se utiliza como péptido funcional en el extremo C-terminal de la proteína de referencia analizada a una secuencia peptídica de pequeño tamaño que interactúa con gran afinidad con la proteína heteróloga que expresa la célula de la invención de la etapa a).

En una realización preferida, el péptido funcional en el extremo C-terminal de la proteína de referencia se refiere a la secuencia SEQ ID NO: 12 que se corresponde con tres repeticiones del péptido PTAP (3xPTAP). El péptido PTAP utilizado en la presente invención se encuentra en la proteína Gag p6 del virus VIH, según se ha indicado anteriormente.

En una realización más preferida, la célula utilizada en el método de doble híbrido en reverso de la invención, comprende integrado en su genoma, las construcciones génicas *SPAL10-URA3*, *(lexAop)4-HIS3* y *ADH1::LexA-TSG101*. Adicionalmente, también puede comprender la construcción génica que comprende el gen reportero detectable *GAL1-LacZ*, que, al igual que el gen reportero seleccionable *SPAL10-URA3*, está bajo el control de un promotor con la secuencia de reconocimiento *UASGal*. La presencia de dicho reportero detectable permite, si fuera necesario,

confirmar o validar los resultados obtenidos con el reportero *SPAL10-URA3* en el método de la invención, con ensayos de la actividad enzimática β -galactosidasa.

En otra realización preferida del método de la invención, el plásmido utilizado en la etapa b) comprende la secuencia nucleotídica que codifica para el dominio de unión al ADN de Gal4 fusionada a la secuencia nucleotídica que codifica para la proteína diana a ensayar. A efectos de la presente invención y, a modo de ejemplo, se ha utilizado como proteína diana a ensayar la proteína humana GKRP de SEQ ID NO: 2, todo ello bajo el control de un promotor constitutivo, preferentemente y a modo de ejemplo, se ha utilizado el promotor de *ADH1*.

Según se describe en los ejemplos incluidos en el presente documento, la célula utilizada en el método de la invención se pre-transforma con dicho plásmido para asegurar la expresión de la proteína diana GKRP, de esta manera, se asegura que la célula expresa la proteína diana, y posteriormente tal y como se indica en la etapa c), se cultiva en condiciones que permiten exclusivamente el crecimiento de las células que han incorporado dicho plásmido.

Los medios de cultivo utilizados para el crecimiento de las células descritas en la presente invención y utilizados en el método aquí descrito, son medios de cultivo conocidos para el crecimiento preferentemente de células de levadura. Preferentemente, los medios de cultivo utilizados en la presente invención son el medio de cultivo mínimo (SD) y el medio de cultivo completo (YPAD). La preparación de dichos medios de cultivo se describe en *Methods in yeast genetics — A laboratory course manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 1990; pp 198. El medio de cultivo completo (YPAD) contiene extracto de levadura al 1% (p/v), bactopectona al 2% (p/v), 0.02% (p/v) de adenil(hemi)sulfato y como fuente de carbono, glucosa al 2%. El medio de cultivo mínimo (SD) se compone de base nitrogenada para levadura (YNB) sin aminoácidos al 0,17% (p/v), y sulfato amónico, $(\text{NH}_4)_2\text{SO}_4$, al 0,5% (p/v) suplementado con glucosa al 2% (p/v). Cuando no se especifica lo contrario, se añaden los requerimientos siguientes a una concentración final de 20 mg/l (histidina, triptófano, uracilo, arginina, metionina), 30 mg/l (isoleucina, lisina, tirosina), 50 mg/l (adenina, fenilalanina), 100 mg/l (leucina), 150 mg/l (valina) o 200 mg/l (treonina).

En otra realización preferida del método de la invención, la célula pre-transformada según se ha mencionado anteriormente, se transforma con un vector linealizado y un fragmento de ADN que comprende la secuencia nucleotídica que codifica para la proteína de referencia que previamente ha sido sometida a un procedimiento de mutagénesis, *in vivo* o *in vitro*, y por lo tanto comprende al menos una mutación. Así, el cultivo de la célula pre-transformada con el vector linealizado y el fragmento de ADN previamente mencionado hace que entre ambas moléculas se produzca recombinación homóloga ("*gap-repair*") y se obtenga un plásmido que comprende el dominio de transactivación de Gal4 fusionado a la secuencia nucleotídica que codifica para la proteína de referencia, y estando la secuencia nucleotídica comprendida en dicho plásmido bajo el control de un promotor. A efectos de la presente invención y a modo de ejemplo, sin ser limitativo, la proteína de referencia es la proteína GK de SEQ ID NO: 10, todo ello bajo el control del promotor constitutivo de ADH1.

En una realización preferida del método de la invención, las mutaciones en la secuencia nucleotídica que codifica para la proteína de referencia se obtienen utilizando preferentemente la técnica de PCR mutagénica y recombinación *in vivo* ("*gap-repair*"). A efectos de la presente invención, el método de mutagénesis mediante reacción en cadena de la polimerasa (PCR) proporciona un método conveniente para inducir mutagénesis de una secuencia elegida (Muhlrاد *et al*, 1992, Yeast 8: 79-82), a efectos de la presente invención en la secuencia de la proteína de referencia. En el método de PCR mutagénica y recombinación *in vivo* "*gap-repair*", el ADN que codifica la secuencia de la proteína que se muta se amplifica en una reacción de PCR en condiciones que favorecen la incorporación de nucleótidos incorrectos en la molécula de ADN. Tales condiciones incluyen niveles relativamente altos de manganeso y/o una mezcla desigual de los diferentes nucleótidos. Los cebadores de PCR que se utilizan en el método de PCR mutagénica y recombinación *in vivo* "*gap-repair*" generan productos de PCR lineales que tienen en sus extremos secuencias homólogas a los extremos del plásmido linealizado, lo que favorece la recombinación entre ambos después de su co-transformación en levadura.

En una realización preferida del método de doble híbrido en reverso descrito en la presente invención se caracteriza por que las etapas b) y d) se pueden llevar a cabo simultáneamente. Así, las células de la invención se co-transformaron con el plásmido

que codifica para la proteína diana y simultáneamente también con el plásmido linealizado y los productos obtenidos de la PCR mutagénica.

5 A efectos de la presente invención, el término "mutado" o "mutación" se refiere a la secuencia modificada, bien mediante mutagénesis dirigida o mutagénesis al azar, respecto a la misma secuencia encontrada de manera natural en la naturaleza (secuencia nativa o secuencia *wild-type*) y que no presenta mutaciones. Las secuencias mutadas incluyen aquellas secuencias que tienen mutaciones puntuales, inserciones, deleciones o reordenamientos.

10

A efectos de la presente invención, el término "promotor" se refiere a una mínima secuencia suficiente para dirigir la transcripción; tales elementos pueden estar situados en los extremos 5' o 3' de la secuencia del gen nativo. Promotores adecuados para la expresión de un gen reportero son aquellos que, cuando se enlaza con el gen
15 reportero, puede dirigir la transcripción del mismo en presencia de moléculas apropiadas (es decir, proteínas que tienen dominios de transactivación). Un ejemplo de un promotor útil es el promotor de *ADH1* o el promotor de *SP013*. Otros promotores útiles incluyen aquellos promotores que contienen secuencias represoras aguas arriba (véase, por ejemplo, Vidal *et al*, 1995, Proc Natl Acad Sci EE.UU. 92:2370-2374), y
20 que inhiben la expresión del gen reportero en ausencia de un activador transcripcional. La capacidad de un promotor para la transcripción de un gen reportero se puede medir con métodos convencionales de ensayos de expresión de genes (por ejemplo, la detección del producto del gen o su ARNm, o la detección del crecimiento celular bajo condiciones donde se requiere la expresión del gen reportero para el crecimiento de
25 una célula).

Técnicas convencionales de biología molecular pueden ser usadas para construir derivados de promotores que incluyen una o más secuencias de reconocimiento para proteínas que se unen el ADN. Por ejemplo, el promotor de *SP013* puede diseñarse
30 para incluir una o más copias del sitio de unión de Gal4. Los sitios de unión naturales de Gal4 han sido ampliamente caracterizados, permitiendo la creación de una secuencia sintética a la que se une Gal4 con una afinidad relativamente alta. Otras secuencias de reconocimiento útiles para proteínas que se unen al ADN incluyen los sitios de unión de LexA y Ace1. Además, donde se mide la capacidad de una proteína
35 para unirse a una secuencia de ADN, la secuencia de reconocimiento de una proteína

que se une al ADN puede ser de tipo nativo o *wild-type*, o puede tener cualquier diseño, tanto intencionalmente como al azar, para probar la capacidad de la secuencia de ADN para interactuar con esta proteína.

- 5 En una realización preferida del método de doble híbrido en reverso descrito en la presente invención, los promotores utilizados son preferentemente promotores constitutivos. En una realización más preferida aún, dichos promotores constitutivos se seleccionan de entre cualquiera de la lista que consiste en: *ADH1*, *PGK1*, *TEF1*, *TPL1*, *HXT7*, *TDH3* y *PYK1*.

10

- En una realización preferida el método de doble híbrido en reverso de la invención se caracteriza por que el plásmido utilizado en la etapa d), antes de transformar la célula con él, es sometido a un procedimiento de mutagénesis, *in vivo* o *in vitro*, para inducir mutaciones al azar en la secuencia nucleotídica que comprende dicho plásmido y que
- 15 codifica para la proteína de referencia. En una realización más preferida aún, el procedimiento de mutagénesis se lleva a cabo mediante PCR mutagénica de la secuencia nucleotídica que codifica para la proteína de referencia y recombinación *in vivo* ("*gap-repair*") del producto obtenido en un plásmido. Preferiblemente, la etapa b) y d) del método de la invención, pueden realizarse por separado o simultáneamente,
- 20 siendo preferido llevarlas a cabo por separado.

- A efectos de la presente invención, en la PCR mutagénica utilizada para introducir mutaciones al azar en la secuencia nucleotídica que codifica para la proteína GK se pueden utilizar diferentes ADN polimerasas con diferentes tasas de mutación cada una
- 25 de ellas, con la finalidad de comparar la idoneidad de cada polimerasa en el método de la invención. En los ejemplos mostrados en el presente documento, se han mostrado los resultados utilizando la ADN polimerasa Taq procedente de Takara o Roche, y la ADN polimerasa Mutazyme II procedente del kit de mutagénesis al azar Genemorph II de Agilent. A efectos de la presente invención es preferible usar una
- 30 ADN polimerasa que incluya una única mutación en cada plásmido.

Otras técnicas conocidas para generar mutaciones al azar, y que pueden ser utilizadas en el método de la invención, pueden ser el uso de mutágenos físicos y/o químicos, cepas celulares mutadoras, preferentemente cepas bacterianas, tal como por ejemplo,

la cepa de *Escherichia coli* mutadora descrita en Rasila TS. *et al.* Anal Biochem. 2009;388(1):71-80.

Técnicas convencionales de transformación celular, preferentemente de
5 transformación en células de levadura pueden ser utilizadas a efectos de la presente invención. A modo de ejemplo y sin pretender limitar, electroporación, método de transformación en presencia de litio, método de transformación con esferoplastos, método de transformación con bolas de vidrio, etc.

10 En una realización preferida del método de la invención, a continuación en la etapa e) se cultiva la célula transformada de la etapa d) bajo condiciones que permiten el crecimiento de las células que presentan mutaciones *missense* que provocan la pérdida de interacción entre el par de proteínas a estudiar. El medio mínimo selectivo utilizado es un medio sin triptófano, leucina, adenina e histidina y con 0.1 % 5-FoA y
15 de 1 a 5 mM 3-AT (SD-AHTL+5-FoA+3-AT). Este medio carece de adenina ya que la cepa utilizada es protótrofa para este requerimiento. La ausencia de triptófano y leucina permite seleccionar los transformantes que han incorporado los dos plásmidos que expresan la proteína diana (marcador *TRP1*) y la proteína de referencia (marcador *LEU2*). La ausencia de histidina permite seleccionar los transformantes que activan el
20 reportero (*(lexAop)4::HIS3*), y la presencia de 3-AT (3-aminotriazol), un inhibidor de *HIS3*, impide el crecimiento de los transformantes con un nivel basal de activación de este reportero. Finalmente, la presencia de 5-FoA impide el crecimiento de los transformantes que activan el reportero *SPAL10::URA3*. En consecuencia, el medio SD-AHTL+5-FoA+3-AT permite el crecimiento únicamente de los transformantes que
25 han incorporado los dos plásmidos, el que expresan la proteína diana y el que expresa la proteína de referencia, y que por otra parte activan el reportero (*(lexAop)4::HIS3*), pero no activan el reportero *SPAL10::URA3*. Dado que la interacción entre la proteína diana y la proteína de referencia activa el reportero *SPAL10::URA3*, los transformantes que no tienen mutaciones, o los que si las tienen pero que no bloquean la interacción entre
30 las dos proteínas, activarán el reportero *SPAL10::URA3* y no podrán crecer en presencia de 5-FoA (**Figura 1a**). Por otra parte, la activación del reportero (*(lexAop)4::HIS3*) depende de la interacción entre TSG101 y el péptido 3xPTAP situado en el extremo C-terminal de la proteína de referencia. En consecuencia, los plásmidos con mutaciones en la secuencia codificante de la proteína de referencia que provocan
35 su truncamiento, y en consecuencia eliminan el péptido 3xPTAP, así como los

plásmidos recircularizados sin incorporar el producto de PCR mutagénico en el proceso de recombinación *in vivo* (o los que se habían quedado sin digerir) no activarán el reportero (*lexAop*)4:*HIS3* y no podrán crecer en ausencia de histidina (**Figura 1b**). En consecuencia, los únicos transformantes que podrán crecer en este medio son los que portan una mutación que bloquea la interacción entre la proteína diana y la proteína de referencia sin que esta mutación produzca un truncamiento en la proteína, es decir que sea una mutación *missense* (**Figura 1c**).

En una realización preferida, el método de doble híbrido en reverso se caracteriza por que la célula se selecciona de la lista que consiste en: célula de levadura, célula bacteriana y célula de mamífero. En una realización más preferida, la célula es preferentemente una célula de levadura. En una realización más preferida aún, la célula de levadura se selecciona del grupo que consiste en *Yarrowia lipolytica*, *Pichia pastoris* y *Saccharomyces cerevisiae*. Las células de la especie *S. cerevisiae* son particularmente útiles en el método de la invención. Las cepas utilizadas en la presente invención se mantienen en cultivos estándar mediante métodos estándar.

En otra realización preferida del método de la invención, en la etapa f) se compararán e identificarán las mutaciones *missense* presentes en los plásmidos que portan las células que han crecido en la etapa e) del método de la invención, respecto a la secuencia nativa o *wild-type* de la proteína de referencia.

En una realización preferida, la identificación de las mutaciones *missense* obtenidas mediante el método de la invención se llevan a cabo mediante la extracción del plásmido que comprende la secuencia con la mutación *missense*, mediante técnicas comúnmente conocidas en el presente campo técnico, y posteriormente se procede a la secuenciación de dicha secuencia mutada, directamente desde el ADN plasmídico, para determinar la mutación concreta comparándola con la secuencia nativa o *wild-type* de la proteína analizada.

Dada la eficiencia del sistema de doble híbrido en reverso descrito en la presente invención, para detectar el 100% de las mutaciones *missense* entre un par de proteínas interaccionantes de estudio, sin falsos positivos, sería posible aislar en masa, y de una sola vez el ADN de todas los clones positivos para luego amplificar la secuencia codificante de la proteína de referencia estudiada (GK en los ejemplos

mostrados) mediante PCR y secuenciar de una sola vez todos los mutantes obtenidos, con un sistema de secuenciación masiva (*Next Generation Sequencing*), de tipo Illumina. El análisis bioinformático del resultado permitiría generar una huella genética del sitio de interacción en la proteína del par interaccionante analizado y que ha sido
5 sometida a mutagénesis.

Tal y como hemos mencionado previamente, la célula de levadura utilizada en la presente invención, ha integrado en su genoma un gen reportero contra-seleccionable, siendo preferido el gen reportero *URA3*, que está unido operativamente a un promotor
10 que comprende una secuencia de reconocimiento para una proteína que se une al ADN, siendo preferida la secuencia UASGal reconocida por el dominio de unión al ADN de Gal4. Además, la célula utilizada comprende también integrado en su genoma otro gen reportero seleccionable, siendo preferido el gen reportero *HIS3*, que está unido operativamente a un promotor que comprende una secuencia de reconocimiento
15 para una proteína que se une al ADN, siendo preferida la secuencia *lexAop* reconocida por el dominio de unión al ADN de LexA. Adicionalmente, la célula de la invención comprende integrado en su genoma la construcción génica *ADH1::LexA-TSG101* (SEQ ID NO: 48), que codifica para la proteína de fusión de SEQ ID NO: 49.

20 Cuando los tres genes mencionados anteriormente (es decir, un gen reportero contra-seleccionable, un gen reportero seleccionable, preferiblemente de selección positiva, y la proteína heteróloga a la que se une el péptido funcional del extremo C-terminal de la proteína de referencia sometida a mutagénesis y recombinación *in vivo*, que forma la proteína de fusión de SEQ ID NO: 49), se integran en el genoma de una célula, los
25 promotores de los dos genes reporteros son distintos, específicamente en la región de los sitios de reconocimiento de la unión ADN-proteína (**Figura 1**), mientras que el resto del promotor puede ser similar.

Célula huésped de la invención.

30

A efectos de la presente invención se utilizan los términos “célula”, “célula huésped”, “cepa”, indistintamente a lo largo del presente documento.

Tal y como se demuestra en el presente documento, los inventores han construido un
35 conjunto de células o cepas de levadura que tienen las siguientes características:

- i) Una primera secuencia nucleotídica que codifica para un gen reportero, donde dicha secuencia nucleotídica está operativamente unida un promotor que comprende una secuencia nucleotídica reconocida por una proteína que se une al ADN,
- 5 ii) Una segunda secuencia nucleotídica que codifica para un segundo gen reportero, donde dicha segunda secuencia nucleotídica está operativamente unida a un promotor que comprende una secuencia de reconocimiento para una proteína que se une al ADN, con la condición de que dicha secuencia sea reconocida por un dominio de unión al ADN distinto del de i), y
- 10 iii) Una tercera secuencia nucleotídica que codifica para una primera proteína de fusión que comprende el dominio de unión al ADN de ii) y una proteína heteróloga capaz de unirse a un péptido funcional localizado en el extremo C-terminal de la proteína de referencia de estudio, donde la secuencia nucleotídica que codifica para dicha primera proteína de fusión está
- 15 operativamente unida a un promotor.

En una realización preferida, la célula de la invención comprende dichas construcciones integradas en su genoma.

- 20 En otra realización más preferida, la célula de la invención se caracteriza por que la primera secuencia nucleotídica que comprende integrada en su genoma comprende un gen reportero que se selecciona de entre genes reporteros de selección positiva y/o, contra-selección, preferentemente un gen reportero de contra-selección, y más preferentemente seleccionado de entre cualquiera de la siguiente lista: *URA3*, *TRP1*,
- 25 *LYS2*, *LYS5*, *CYH2*, *CAN1*, *GAL1*, *mazF*. En otra realización preferida, dicho gen reportero está operativamente unido a una secuencia de reconocimiento para una proteína que se une al ADN, según se ha definido previamente de igual manera para el método de la invención. En una realización más preferida aún, la primera secuencia nucleotídica que comprende la célula de la invención se caracteriza por que
- 30 comprende la construcción génica *UASGal-URA3* (= *SPAL10::URA3*), que expresa el marcador de contra-selección *URA3*.

- En otra realización preferida de la célula de la invención, ésta se caracteriza por que la
- 35 un gen reportero que se selecciona de entre genes reporteros de selección positiva y/o

contra-selección, preferentemente un gen reportero de selección positiva, y más preferentemente seleccionado de entre cualquiera de la siguiente lista: *HIS3*, *LEU2*, *URA3*, *ADE2*, *TRP1*, *LYS2*, *LYS5*. En otra realización preferida, dicho gen reportero está operativamente unido a una secuencia de reconocimiento para una proteína que se une al ADN, con la condición de que dicha secuencia de reconocimiento es diferente de la secuencia de reconocimiento comprendida en la primera secuencia nucleotídica que comprende la célula, es decir, diferente a la secuencia de reconocimiento de i). En una realización más preferida aún, la segunda secuencia nucleotídica que comprende la célula de la invención se caracteriza por que comprende la construcción génica (*lexAop*)4-*HIS3*, que expresa el reportero de selección positiva *HIS3*.

Tal y como se ha descrito a lo largo del presente documento, los dominios de unión al ADN que pueden comprender las cepas de la invención se seleccionan de entre cualquiera de los siguientes factores de transcripción: Gal4, LexA y Ace1.

En otra realización preferida de la célula de la invención, ésta se caracteriza por que comprende además integrado en su genoma, una tercera secuencia nucleotídica que codifica para una proteína de fusión, denominada aquí primera proteína de fusión, que comprende el dominio de unión al ADN de LexA y la proteína heteróloga que se une al péptido funcional descrito y utilizado en la presente invención, todo ello bajo el control del promotor de *ADH1*. A efectos de la presente invención, y a modo de ejemplo, sin querer ser limitativo se ha utilizado como proteína heteróloga con la que se transforma la célula de la invención y que se integra en su genoma la proteína humana TSG101 (SEQ ID NO: 32, codificada por la secuencia nucleotídica SEQ ID NO: 31). Así, la célula de la invención expresa la proteína de fusión LexA-TSG101 de SEQ ID NO: 49, a partir de la secuencia nucleotídica SEQ ID NO: 48.

En una realización preferida, la construcción génica que codifica para la proteína de fusión de la célula de la invención aquí descrita, está unida operativamente a un promotor, preferentemente constitutivo y que se selecciona de entre cualquiera de los descritos anteriormente, siendo preferido el promotor de *ADH1*.

En otra realización preferida de la célula de la invención, ésta se caracteriza por que comprende integrado en su genoma, al menos dos genes reporteros, preferentemente

uno de selección positiva y otro de contra-selección. En otra realización más preferida aún, el gen reportero de selección positiva es *HIS3* y el gen reportero de contra-selección es *URA3*.

- 5 En una realización más preferida, la célula de la invención, comprende integrado en su genoma, las construcciones génicas *SPAL10-URA3*, *(lexAop)4-HIS3* y *ADH1::LexA-TSG101*. Adicionalmente, también puede comprender la construcción génica que expresa para el gen reportero detectable *GAL1-LacZ*, que está bajo el control de un promotor con la secuencia de reconocimiento *UASGal*. La presencia de dicho
- 10 reportero detectable permite, si fuera necesario, confirmar o validar los resultados obtenidos con el reportero *URA3* en el método de la invención, con ensayos de la actividad β -galactosidasa.

- En una realización preferida, el método de doble híbrido en reverso se caracteriza por
- 15 que la célula se selecciona de la lista que consiste en: célula de levadura, célula bacteriana y célula de mamífero. En una realización más preferida, la célula es preferentemente una célula de levadura. En una realización más preferida aún, la célula de levadura se selecciona del grupo que consiste en *Yarrowia lipolytica*, *Pichia pastoris* y *Saccharomyces cerevisiae*. Las células de la especie *S. cerevisiae* son
- 20 particularmente útiles en el método de la invención. Las cepas utilizadas en la presente invención se mantienen en cultivos estándar mediante métodos estándar.

- En una realización más preferida aún, la célula de la invención es la célula OVY216. Dicha célula OVY216 se caracteriza por el genoma: *MATa ade2-101*, *his3- Δ 200*, *leu2-3,112*, *trp1-901*, *gal4 Δ* , *gal80 Δ* , *LYS2:(lexAop)4-HIS3*, *SPAL10::URA3*, *GAL1-lacZ*
- 25 *ADE2::LexA-TSG101*, que comprende integrado cromosómicamente las construcciones génicas *(lexAop)4-HIS3* y *UASGal-URA3 (=SPAL10::URA3)*, que expresan los reporteros *HIS3* y *URA3*, y la construcción génica *ADH1::LexA-TSG101*, que expresa la proteína de fusión LexA-TSG101 (SEQ ID NO: 49) que es la proteína
- 30 que se une al péptido funcional 3xPTAP (SEQ ID NO: 12). Adicionalmente, tal y como se han indicado anteriormente, la cepa OVY216 de la presente invención también comprende la construcción que expresa el gen reportero detectable *GAL1-lacZ*. Este gen reportero, al igual que la construcción *SPAL10::URA3*, está bajo el control de *UASGal*. La presencia del reportero detectable *GAL1-lacZ* permite, si fuese necesario,

confirmar o validar los resultados obtenidos con el reportero *URA3*, con ensayos de la actividad β -galactosidasa.

Construcciones génicas

5

Otro aspecto descrito en la presente invención se refiere a las construcciones génicas utilizadas para obtener la célula huésped descrita en la presente invención, así como para llevar a cabo el método de doble híbrido en reverso aquí descrito.

10 Una de las construcciones génicas aquí descritas, a partir de ahora la denominaremos primera construcción génica de la invención, comprende las secuencias nucleotídicas que codifican para:

i) un promotor

ii) un dominio de unión al ADN, y

15 iii) una proteína heteróloga capaz de unirse a un péptido funcional localizado en el extremo C-terminal de la proteína de referencia.

En una realización preferida, la primera construcción génica de la invención se caracteriza por que comprende un promotor, preferentemente promotor constitutivo,
20 que puede seleccionarse de entre cualquiera de los descritos previamente en la presente invención. Adicionalmente, dicha primera construcción génica comprende también un dominio de unión al ADN, que se selecciona de entre cualquiera de los descritos a lo largo del presente documento. De la misma manera, la primera construcción génica de la invención, comprende también la secuencia nucleotídica que
25 codifica para la proteína heteróloga que se une al péptido funcional descrito en la presente invención. En una realización preferida, dicha proteína heteróloga es la proteína humana TSG101 (SEQ ID NO: 32, codificada por la secuencia nucleotídica SEQ ID NO: 31). La proteína TSG101 es capaz de interaccionar con un pequeño péptido denominado PTAP que está presente en la proteína Gag p6 del virus VIH
30 (PNAS (2001) 98, 7724-9; Cell (2001) 107, 55-65).

En una realización más preferida aún, la primera construcción génica de la invención es preferentemente el plásmido pRS402-LexA-TSG101, más preferentemente el plásmido pRS402-LexA-TSG101 que comprende la secuencia SEQ ID NO: 30.

35

En otra realización preferida, la primera construcción génica descrita en la presente invención se encuentra integrada en el cromosoma de la célula huésped de la invención y codifica para la primera proteína de fusión descrita en la presente invención, la proteína de fusión LexA-TSG101 de SEQ ID NO: 49.

5

Otro aspecto descrito en la presente invención se refiere a una segunda construcción génica, a partir de aquí la denominaremos segunda construcción génica de la invención, que comprende las secuencias nucleotídicas que codifican para:

- i) un promotor,
- 10 ii) un dominio de unión al ADN, y
- iii) una proteína diana a la que se une la proteína de referencia.

En una realización preferida, la segunda construcción génica de la invención, comprende por tanto, un promotor, preferentemente constitutivo, pudiendo utilizarse
15 cualquiera de los descritos en la presente invención, un dominio de unión al ADN, que se selecciona de entre cualquiera de los descritos a lo largo del presente documento, y la secuencia de la proteína diana que se une a la proteína de referencia. A efectos de la presente invención, la proteína diana ejemplificada es la proteína reguladora de la glucoquinasa humana de SEQ ID NO: 2 codificada por la SEQ ID NO: 1. En una
20 realización más preferida aún, la tercera construcción génica de la invención se refiere al plásmido pGBKT7-GKRP, más preferentemente al plásmido pGBKT7-GKRP que comprende la SEQ ID NO: 6.

En otra realización preferida, la segunda construcción génica descrita en la presente
25 invención codifica para la segunda proteína de fusión descrita en la presente invención, la proteína de fusión GBD-GKRP de SEQ ID NO: 51.

Otro aspecto descrito en la presente invención se refiere a una tercera construcción génica, a partir de aquí la denominaremos tercera construcción génica de la invención,
30 que comprende las secuencias nucleotídicas que codifican para:

- i) un promotor
- ii) un dominio de transactivación, y
- iii) una proteína de referencia que además en su extremo carboxilo
terminal comprende la secuencia que codifica para el péptido funcional

3xPTAP que se une específicamente a la proteína heteróloga codificada por la primera construcción génica de la invención.

En una realización preferida, la tercera construcción génica de la invención,
5 comprende por tanto, un promotor, preferentemente constitutivo, pudiendo utilizarse cualquiera de los descritos en la presente invención, un dominio de transactivación, pudiendo utilizarse cualquiera de los descritos en la presente invención, y la secuencia de la proteína de referencia que se quiera estudiar, fusionada al péptido funcional 3xPTAP de SEQ ID NO: 12 en su extremo C-terminal. A efectos de la presente
10 invención, la proteína de referencia ejemplificada es la proteína glucoquinasa humana de SEQ ID NO: 10 codificada por la SEQ ID NO: 9. En una realización más preferida aún, la segunda construcción génica de la invención se refiere al plásmido pACT2-GK-3xPTAP, más preferentemente al plásmido pACT2-GK-3xPTAP que comprende la SEQ ID NO: 14.

15

En otra realización preferida, la tercera construcción génica descrita en la presente invención codifica para la tercera proteína de fusión descrita en la presente invención, la proteína de fusión GAD-GK-3xPTAP de SEQ ID NO: 53.

20 A efectos de la presente invención se pueden utilizar otros plásmidos a los descritos aquí, con los mismos dominios de unión al ADN, tal como por ejemplo el plásmido pDESTTM32 (Invitrogen), o con otros (en sustitución de GBD y LexA) en combinación con una cepa con reporteros bajo el control de promotores con las secuencias de reconocimiento correspondientes, en sustitución de UASGal y lexAop, así como otros
25 plásmidos con el mismo dominio de transactivación (GAD), tal como por ejemplo el plásmido pEXPTM-AD502 (Invitrogen), o con otro dominio de transactivación, tal como por ejemplo, VP16.

A efectos de la presente invención, también se podría utilizar como péptido funcional,
30 en sustitución del péptido 3xPTAP utilizado en la presente invención, junto con su proteína interactora, TSG101, cualquier otra pareja, siempre que se cumpla la condición de que el péptido funcional debe ser un péptido pequeño, preferentemente menos de 30 aminoácidos y más preferentemente menos de 25 aminoácidos, que no interfiera con la interacción de las proteínas estudiadas, o que para ser detectado
35 requiera de más etapas incrementando la complejidad de la técnica, así como otro

promotor diferente a *ADH1* para expresar la fusión a TSG101 y otro marcador que *ADE2* para su integración genómica.

A lo largo de la descripción y las reivindicaciones la palabra "comprende" y sus
5 variantes no pretenden excluir otras características técnicas, aditivos, componentes o
pasos. Para los expertos en la materia, otros objetos, ventajas y características de la
invención se desprenderán en parte de la descripción y en parte de la práctica de la
invención. Los siguientes ejemplos y figuras se proporcionan a modo de ilustración, y
no se pretende que sean limitativos de la presente invención.

10

BREVE DESCRIPCIÓN DE LAS FIGURAS

Figura 1. Representación esquemática de los genes reporteros de contra-selección
URA3 y selección positiva *HIS3*, bajo el control de promotores con secuencias de
reconocimiento para distintos dominios de unión al ADN. (a) La ausencia de mutación
15 en la proteína Y (ejemplificada en la proteína humana glucoquinasa) produce la
activación de *URA3* y la falta de crecimiento en presencia de 5-FoA. (b) Una mutación
de truncamiento en Y no permite la activación de *HIS3* y el crecimiento en ausencia de
histidina. (c) Una mutación *missense* en Y produce la activación de *HIS3* pero no la de
URA3, lo que permite el crecimiento en un medio sin histidina y con 5-FoA. X: se
20 refiere a la proteína reguladora de la glucoquinasa (GKRP).

Figura 2. Construcciones génicas utilizadas en el método de la invención. (a)
Construcciones integradas en el genoma de la cepa de levadura OVY216. (b)
Construcciones génicas utilizadas para la transformación de la cepa de la invención
OVY216. X: se refiere a la proteína reguladora de la glucoquinasa humana (GKRP). Y:
25 se refiere a la proteína humana glucoquinasa (GK).

Figura 3. Representación de la técnica de PCR mutagénica y recombinación *in vivo*
("gap repair") para la introducción de mutaciones al azar en el gen Y (ejemplificada en
el gen que codifica para la proteína humana glucoquinasa) comprendido en el
30 plásmido pACT2-Y-3xPTAP de SEQ ID NO: 14.

Figura 4. Ensayo de la actividad β -galactosidasa en filtro para detectar la interacción
en el sistema de doble-híbrido clásico de células OVY216 transformadas con el
plásmido pACT2-GK-3xPTAP sin mutar o sometido a PCR mutagénica y con el

plásmido pGBKT7-GKRP (panel izquierda) o con el plásmido pLexA(1-202)PL-TSG101 (panel derecha). La interacción entre las proteínas GK y GKRP (izquierda) o entre el péptido 3xPTAP y TSG101 (derecha) produce la activación del gen reportero *lacZ*, y en consecuencia la expresión de la β -galactosidasa y la hidrólisis del X-Gal en un compuesto de color azul. Un total de 7-8 células independientes transformadas han sido analizadas para cada interacción.

EJEMPLOS

A continuación se ilustrará la invención mediante unos ensayos realizados por los inventores, que pone de manifiesto la efectividad del producto de la invención.

En los ejemplos que se detallan a continuación se muestra como mediante el sistema de doble híbrido en reverso, utilizando la célula huésped descrita en la presente invención, así como las construcciones génicas específicas, diseñadas de acuerdo a los ejemplos mostrados, se han localizado mutaciones en la proteína glucoquinasa humana que bloquean su interacción con la proteína humana GKRP. Además, al ser conocida la estructura del complejo GK-GKRP (Choi JM. *et al.* Proc Natl Acad Sci U S A. 2013 Jun 18;110(25):10171-6), ha permitido validar los resultados obtenidos mediante el método descrito en la presente invención, poniendo de manifiesto la consistencia entre los resultados aquí obtenidos con la estructura conocida del complejo.

Ejemplo 1. Análisis de la validez del método de doble híbrido en reverso de la invención mediante el análisis de la interacción entre el par de proteínas GK y GKRP.

1.1. Diseño y obtención de los plásmidos utilizados en el método de la invención.

Las construcción de los plásmidos utilizados en el método de invención ha sido realizada mediante técnicas estándar de ADN recombinante (Sambrook *et al.* Molecular Cloning: A Laboratory Manual, 2d ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.).

Para expresar una de las proteína de fusión descritas en el presente documento y que comprende una de las proteínas del par interaccionante a estudiar, específicamente

para el presente ejemplo se ha seleccionado la proteína reguladora de la glucoquinasa humana (GKRP) de SEQ ID NO: 2, codificada por el gen de la glucoquinasa humana que comprende la SEQ ID NO: 1, se ha utilizado el plásmido comercial pGBKT7 (Clontech) que comprende como marcador de selección *TRP1*. Dicho plásmido
5 comercial pGBKT7 (SEQ ID NO: 3), comprende la secuencia nucleotídica que codifica para el dominio de unión al ADN del activador transcripcional Gal4 (GBD) (SEQ ID NO: 4), bajo el control del promotor de *ADH1* (SEQ ID NO: 5). De esta manera, para obtener el plásmido pGBKT7-GKRP (SEQ ID NO: 6), la secuencia nucleotídica que codifica para la GKRP (SEQ ID NO: 1) se clona en fase, fusionada en el extremo N-
10 terminal al GBD. Brevemente, mediante la técnica de PCR se clona en el plásmido pGBKT7 (Clontech) la secuencia codificante (SEQ ID NO: 1) de la proteína GKRP humana (SEQ ID NO: 2), incluyendo el codón de iniciación y el codón de parada, entre los sitios EcoR1 y BamH1, de tal forma que las secuencias nucleotídicas que codifican para el dominio de unión al ADN de Gal4 (GBD) y la secuencia codificante (SEQ ID
15 NO: 1) para la proteína GKRP (SEQ ID NO: 2) quedan en fase.

Para la amplificación de la secuencia codificante (SEQ ID NO: 1) de la proteína GKRP (SEQ ID NO: 2) mediante PCR, se utilizó como molde el plásmido pFlag-ctc-hGKRP-FlagC (Brocklehurst K, *et al.* Biochem. J. 2004;378(Pt2):693-7) y se diseñaron
20 cebadores específicos capaces de hibridar con las regiones 5' y 3' del gen que codifica para la proteína GKRP (SEQ ID NO: 1) y con los sitios de las endonucleasas de restricción EcoR1 y BamH1 en los extremos. Los cebadores utilizados para dicho fin son los cebadores de SEQ ID NO: 7 (hGKRP-EcoR1) (5'-GGAATTCATGCCAGGCACAAAACGGTTTC-3') y de SEQ ID NO: 8 (hGKRP-BamH1)
25 (5'-GGGATCCTACTGAACGTCAGGCTCTAGGATTTC-3').

Por otro lado, para expresar la proteína de fusión que comprende la otra proteína del par interaccionante a estudiar, específicamente para el presente ejemplo, se trata de la proteína glucoquinasa humana (GK) de SEQ ID NO: 10, codificada por la secuencia
30 nucleotídica SEQ ID NO: 9, y además las tres repeticiones del péptido PTAP (3xPTAP) (SEQ ID NO: 12: EPEPTAPPEPTAPPEPTAPPAE), codificado por la secuencia nucleotídica SEQ ID NO: 11, se ha utilizado el plásmido comercial pACT2 (Clontech) de SEQ ID NO: 13, que comprende como marcador de selección *LEU2*. Así, para la obtención del plásmido final pACT2-GK-3xPTAP (SEQ ID NO: 14), la secuencia
35 codificante SEQ ID NO: 9 de la proteína GK (SEQ ID NO: 10) se fusiona en el extremo

N-terminal al dominio de transactivación de Gal4 (GAD) (SEQ ID NO: 15), y en el extremo C-terminal a las tres repeticiones del péptido PTAP (SEQ ID NO: 11), todo ello bajo el control del promotor de *ADH1* (SEQ ID NO: 16). Brevemente, el plásmido comercial pACT2 (Clontech) de SEQ ID NO: 13 ha sido modificado mediante la
5 inserción en el sitio Xho1 del *polylinker*, de la secuencia nucleotídica (SEQ ID NO: 11) que codifica para el péptido 3xPTAP de SEQ ID NO: 12 (EPEPTAPPEPTAPPEPTAPPAE). Esta secuencia nucleotídica (SEQ ID NO: 11) ha sido obtenida mediante el anillamiento de dos oligonucleótidos complementarios de
SEQ ID NO: 17 (OV570) (5'-
10 TCGAGAGCCAGAACCCACAGCACCGCCTGAGCCTACCGCCCCACCCGAACCGAC
GGCGCCTCCAGCTGAGTAA-3') y de SEQ ID NO: 18 (OV571) (5'-
TCGATTACTCAGCTGGAGGCGCCGTCGGTTCGGGTGGGGCGGTAGGCTCAGGC
GGTGCTGTGGGTTCTGGCTC-3'). El oligonucleótido de doble cadena resultante presenta extremos protuberantes 5' (TCGA) que permiten su clonaje en el sitio Xho1
15 del plásmido pACT2. Así, el plásmido resultante tras el clonaje de dicha secuencia nucleotídica se denomina pACT2-3xPTAP (SEQ ID NO: 19) y presenta el mismo *polylinker* que el plásmido pACT2 original de SEQ ID NO: 11, pero limitado en cada extremo N-terminal y C-terminal con las secuencias que codifican para el dominio de activación de Gal4 (GAD) y para el péptido 3xPTAP, respectivamente, en fase uno con
20 el otro. A continuación, el plásmido pACT2-GK-3xPTAP de SEQ ID NO: 14 se obtiene mediante clonaje por PCR de la secuencia codificante de la GK humana (SEQ ID NO: 9), incluyendo el codón de iniciación, pero excluyendo el codón de parada, en el sitio BamH1 del *polylinker* del plásmido pACT2-3xPTAP de SEQ ID NO: 19, obtenido previamente, de modo que las secuencias nucleotídicas que codifican para el dominio
25 de activación de Gal4 (GAD), para la proteína GK y para el péptido 3xPTAP, quedan en fase.

Para la amplificación por PCR de la secuencia nucleotídica SEQ ID NO: 9, que codifica para la proteína GK humana (SEQ ID NO: 10), se utilizó como molde el plásmido
30 pGEX-5X-hGKi (Alvarez E, *et al.* J. Neurochem. 2002;80(1):45-53) y se diseñaron cebadores específicos capaces de hibridar con las regiones 5' y 3' del gen que codifica para la proteína GK y con los sitios de la endonucleasa de restricción BamH1 en cada extremo. Los cebadores utilizados para dicho fin son los cebadores de SEQ ID NO: 20 (OV500) (5'-GGGATCCTGATGCTGGACGACAGAGCCAGG-3') y de SEQ ID NO: 21
35 (OV562) (5'-GGGATCCACTGGCCCAGCATACAGGCCTTC-3').

Por otro lado, para la expresión y síntesis de la proteína TSG101 humana (SEQ ID NO: 32) por parte de la célula OVY216 descrita en la presente invención, se sintetizó el plásmido pRS402-LexA-TSG101 (SEQ ID NO: 30) en dos etapas. En una primera etapa, se construyó el plásmido pRS402-LexA de SEQ ID NO: 22. Para ello, se clonó la secuencia nucleotídica SEQ ID NO: 23 que codifica el dominio de unión al ADN del represor transcripcional LexA (aminoácidos 1-202) de SEQ ID NO: 24, junto con el promotor (SEQ ID NO: 25) y terminador (SEQ ID NO: 26) de *ADH1*, entre los sitios Kpn1 y Sac1 del plásmido integrativo pRS402 (Brachmann CB. *et al.* Yeast. 1998;14(2):115-32). Esta secuencia génica se amplificó mediante PCR utilizando como molde el plásmido pLexA(1-202)PL (Ruden DM, *et al.* Nature. 1991;350(6315):250-2) de SEQ ID NO: 27 y cebadores capaces de hibridar al inicio del promotor de *ADH1* (SEQ ID NO: 25) y al final del terminador de *ADH1* (SEQ ID NO: 26), y con los sitios de las endonucleasas de restricción Kpn1 y Sac1 en los extremos. Los cebadores utilizados para dicho fin fueron los cebadores de SEQ ID NO: 28 (OV502) (5'-GGGTACCTTTTGTGTTTCCGGGTGTAC-3') y SEQ ID NO: 29 (OV503) (5'-GGAGCTCGCATGCCGGTAGAGGTGTGGTC-3'). En la segunda etapa, se obtuvo el plásmido definitivo denominado pRS402-LexA-TSG101 de SEQ ID NO: 30 mediante PCR y clonaje de la secuencia nucleotídica SEQ ID NO: 31 que codifica para la proteína TSG101 (SEQ ID NO: 32) (incluyendo el codón de iniciación y el codón de parada), en el sitio BamH1 del plásmido previamente construido pRS402-LexA (SEQ ID NO: 22). Como se mostrará a continuación, el plásmido pRS402-LexA-TSG101 (SEQ ID NO: 30) se integra en el cromosoma de la célula de levadura OVY216 descrita en la presente invención, específicamente en el locus del marcador *ADE2* de dicha célula.

Para la amplificación del gen *TSG101* (SEQ ID NO: 31) por PCR, se utilizó como molde una librería de ADNc humano (Clontech ref 638805) y cebadores capaces de hibridar con las regiones 5' y 3' de gen *TSG101* (SEQ ID NO: 31) y con sitios de la endonucleasa de restricción BamH1 en cada extremo. Los cebadores utilizados para dicho fin fueron los cebadores de secuencias SEQ ID NO: 34 (OV310) (5'-GGGATCCTCATGGCGGTGTCGGAGAGCCAGC-3') y SEQ ID NO: 35 (OV311) (5'-GGGATCCTCAGTAGAGGTCACTGAGACCGGC-3').

Por otra parte, el plásmido pLexA(1-202)PL-TSG101 (SEQ ID NO: 33) utilizado en el sistema de doble híbrido clásico, que se utiliza para validar el método de la invención y que se describe a continuación, se obtuvo de la misma manera que el plásmido pRS402-LexA-TSG101 (SEQ ID NO: 30) mencionado previamente, pero utilizando el
5 plásmido pLexA(1-202)PL (SEQ ID NO: 27), en vez del plásmido pRS402-LexA (SEQ ID NO: 22), como vector receptor.

1.2. Diseño y obtención de la cepa OVY216 de la invención

La cepa de levadura OVY216 de la presente invención comprende integrado
10 cromosómicamente las construcciones génicas *(lexAop)4p-HIS3* y *UASGal-URA3* (=SPAL10::URA3), que expresan los genes reporteros *HIS3* (utilizado en la presente invención como gen reportero de selección positiva) y *URA3* (utilizado en la presente invención como gen reportero de contra-selección), y la construcción génica *ADH1::LexA-TSG101*, que expresa la proteína de fusión LexA-TSG101 (SEQ ID NO:
15 49), que se une al péptido 3xPTAP (SEQ ID NO: 12) expresado por el plásmido pACT2-GK-3xPTAP (SEQ ID NO: 14).

La cepa de levadura OVY216 de la presente invención se ha obtenido mediante el cruce de la cepa comercial procedente de Invitrogen MAV203 (*MAT α* , *leu2-3,112*, *trp1-901*, *his3- Δ 200*, *gal4 Δ* , *gal80 Δ* , *can1R*, *cyh2R*, *LYS2::GAL1-HIS3*, *GAL1-lacZ*,
20 *SPAL10::URA3*) que comprende el gen reportero *URA3* bajo el control de UASGal, y de la cepa comercial procedente de Invitrogen L40-ura (*MAT α* , *leu2-3,112*, *trp1-901*, *his3- Δ 200*, *ade2-101*, *gal80 Δ* , *LYS2:(lexAop)4-HIS3*, *ura3:(lexAop)8-lacZ*), que comprende el gen reportero *HIS3* bajo el control de LexAop.

25

Los medios de cultivo de levadura utilizados en la presente invención para el crecimiento de las cepas son el medio de cultivo mínimo (SD) y el medio de cultivo completo (YPAD). La preparación de dichos medios de cultivo se describe en *Methods in yeast genetics — A laboratory course manual*. Cold Spring Harbor Laboratory Press,
30 Cold Spring Harbor, New York. 1990; pp 198.

Así, para obtener la cepa OVY216 se mezclaron cantidades equivalentes de cada cepa L40-ura y MAV203, en una placa de medio de cultivo completo (YPAD) y se incubaron a 30°C durante 5 horas. La obtención de organismos diploides (*ade2-101/ade2-101*, *his3- Δ 200/his3- Δ 200*, *leu2-3,112/leu2-3,112*, *trp1-901/trp1-901*,
35

gal4Δ/GAL4, *gal80Δ/gal80Δ*, *LYS2::GAL1-HIS3/LYS2:(lexAop)4-HIS3*, *ura3:(lexAop)8-lacZ/SPAL10::URA3*, *GAL1-lacZ/-*) se comprobó por observación al microscopio y varios de ellos fueron recogidos con un micromanipulador (Modelo MSM system SINGER) y depositados en otra placa de cultivo en presencia de medio de cultivo completo YPAD. Trascurridos dos días de cultivo a 30°C, los diploides fueron transferidos a una placa de medio de pre-esporulación (0.8% extracto de levadura, 0.3% bactopectona, 10% glucosa, 2% agar) e incubados a 30°C durante 24 horas, para luego ser transferidos a una placa de medio de esporulación (1% acetato potásico, 0.1% extracto de levadura, 0.05% glucosa, 2% agar) e incubados a 30°C durante 3 días. Posteriormente, se preparó una suspensión de cada uno de los diploides y se comprobó la obtención de tétradas por observación al microscopio. Para la disección de tétradas, se incubó una de las suspensiones de los individuos diploides en presencia de beta-glucuronidasa durante 15 minutos a 25°C y las esporas fueron separadas con el micromanipulador MSM system SINGER y cultivadas en una placa de medio completo YPAD.

De dicho cultivo se obtuvieron 70 segregantes haploides viables que se caracterizaron mediante análisis genotípico/fenotípico. Para la caracterización fenotípica se seleccionaron aquellos segregantes que no pueden utilizar la galactosa como fuente de carbono (*gal4Δ*) y que requieren adenina en el medio de cultivo (*ade2-101*). Estos cultivos se llevaron a cabo a 30°C en placas que comprendían medio de cultivo completo con un 2% de galactosa como fuente de carbono o medio mínimo SD sin adenina. Por otra parte, el análisis genotípico se realizó mediante técnicas de PCR, seleccionando aquellos segregantes que comprenden las construcciones génicas *SPAL10-URA3*, *(lexAop)4-HIS3* y *GAL1-lacZ*. Para la detección de las construcciones génicas mencionadas se amplificó vía PCR un fragmento de 300-400 pb del genoma de los segregantes utilizando como cebadores los descritos a continuación:

Construcción génica	Cebador directo (5'-3')	Cebador inverso (5'-3')
<i>SPAL10-URA3</i>	SEQ ID NO: 36 (OV713) (GCGAGGCATATTTATGG TGAAGG)	SEQ ID NO: 37 (OV714) (CATTTCCGTGCAAAGGTA CTAAC)
<i>(lexAop)4-HIS3</i>	SEQ ID NO: 38 (OV731) (CTGTATATAAAACCACT GGTTATATGTAC)	SEQ ID NO: 39 (OV732) (TCGAGTGCTCTATCGCTA GGG)

Construcción génica	Cebador directo (5'-3')	Cebador inverso (5'-3')
<i>GAL1-lacZ</i>	SEQ ID NO: 40 (OV715) (CCATAGGATGATAATGC GATTAG)	SEQ ID NO: 41 (OV741) (CGCTTCTGGTGCCGAAAA CC)

De los 70 segregantes haploides obtenidos del cruce de las cepas de levadura L40-ura (Invitrogen) y MAV203 (Invitrogen), se ha identificado un único segregante denominado en la presente invención cepa OVY211, que presenta el genotipo: *MATa ade2-101, his3-Δ200, leu2-3,112, trp1-901, gal4Δ, gal80Δ, LYS2:(lexAop)4-HIS3, SPAL10::URA3, GAL1-lacZ*.

En el locus *ADE2* de esta cepa OVY211 se ha integrado el plásmido pRS402-LexA-Tsg101 de SEQ ID NO: 30 (obtenido previamente según se ha descrito) que codifica para la proteína de fusión LexA-Tsg101 (SEQ ID NO: 49), dando lugar a la cepa de levadura de la invención, OVY216 (*MATa ade2-101, his3-Δ200, leu2-3,112, trp1-901, gal4Δ, gal80Δ, LYS2:(lexAop)4-HIS3, SPAL10::URA3, GAL1-lacZ ADE2::LexA-TSG101*) que comprende integrado cromosómicamente las construcciones génicas (*lexAop*)4-HIS3 y *UASGal-URA3* (=SPAL10::URA3), que expresan los genes reporteros *HIS3* (selección positiva) y *URA3* (contra-selección), y la construcción génica *ADH1::LexA-TSG101*, que expresa la proteína de fusión LexA-TSG101 (SEQ ID NO: 49) que es la proteína que se une al péptido funcional 3xPTAP (SEQ ID NO: 12). (Figura 2a). Brevemente, se transformó la cepa de levadura OVY211 con el plásmido pRS402-LexA-TSG101 (SEQ ID NO: 30) linealizado con la endonucleasa de restricción Stu1 que reconoce un único sitio en el marcador *ADE2* de dicho plásmido. La selección de los transformantes en los cuales dicho plásmido se ha integrado mediante recombinación homóloga en el locus del marcador *ADE2* se llevó a cabo en una placa de medio mínimo SD sin adenina. La correcta integración del plásmido se confirmó mediante amplificación por PCR de un fragmento de 2.3Kb con los cebadores de SEQ ID NO: 42 (OV508) (5'-CAGATTGTACTGAGAGTGCACC-3') y SEQ ID NO: 43 (OV747) (5'-ATTCCTTGCTTCTTGTTACTGG-3'). Por lo tanto, la cepa de levadura de la invención, OVY216, tal y como hemos indicado anteriormente comprende integrado cromosómicamente las construcciones génicas *LexAop-HIS3* y *UASGal-URA3* (=SPAL10::URA3), que expresan los reporteros *HIS3* y *URA3*, y la construcción génica *ADH1::LexA-TSG101*, que expresa la proteína de fusión LexA-TSG101 (SEQ

ID NO: 49) que es la proteína que se une al péptido funcional 3xPTAP (SEQ ID NO: 12) incluido en el plásmido pACT2-GK-3xPTAP (SEQ ID NO: 14).

Adicionalmente, la cepa OVY216 de la presente invención comprende la construcción que expresa el gen reportero inducible *GAL1-lacZ*. Este gen reportero, al igual que la construcción *SPAL10::URA3*, está bajo el control de *UASGal*. La presencia del gen reportero inducible *GAL1-lacZ* permite, si fuese necesario, confirmar o validar los resultados obtenidos con el gen reportero *URA3*, mediante ensayos de actividad β -galactosidasa.

1.3. Transformación de la cepa OVY216 con el plásmido pGBKT7-GKRP de SEQ ID NO: 6

Tras la obtención de la cepa de la invención OVY216 según se ha descrito previamente en 1.2, ésta se transforma con el plásmido pGBKT7-GKRP (SEQ ID NO: 6) para que exprese una de las proteínas del par interaccionante que se estudia en el presente ejemplo, la proteína GKRP (SEQ ID NO: 2). Simultáneamente a la transformación de la célula OVY216 con el plásmido pGBKT7-GKRP de SEQ ID NO: 6 se podría llevar a cabo la transformación de dicha célula, para obtener mediante recombinación *in vivo* "gap repair", el segundo plásmido pACT2-GK-3xPTAP, pero para asegurar la expresión de la proteína de fusión GBD-GKRP (SEQ ID NO: 51), por parte de la célula OVY216, se prefiere llevar a cabo en primer lugar la transformación de dicha célula con el plásmido pGBKT7-GKRP (SEQ ID NO: 6), y una vez que la célula comprende dicho plásmido, posteriormente llevar a cabo la segunda transformación y obtener el plásmido pACT2-GK-3xPTAP recombinado. De esta manera, el plásmido pGBKT7-GKRP (SEQ ID NO: 6) expresa la proteína de fusión GBD-GKRP (SEQ ID NO: 51) a un nivel suficiente como para activar la expresión del gen reportero de selección positiva *URA3* (mediante su interacción con la proteína de fusión *GAD-GK-3xPTAP* (SEQ ID NO: 53)) expresada por el plásmido pACT2-GK-3xPTAP de SEQ ID NO: 14, e impedir el crecimiento, en presencia de 5-FoA, de los clones que comprenden la proteína GK que no presentan mutaciones *missense*.

El protocolo de transformación de la cepa de levadura OVY216 con el plásmido pGBKT7-GKRP (SEQ ID NO: 6) se llevó a cabo mediante la metodología estándar con acetato de litio. Brevemente, la célula de levadura OVY216 se mantiene en cultivo en

medio completo YPAD hasta que llega a la fase logarítmica de crecimiento. A continuación, se centrifuga el cultivo durante 5 min a 2000 rpm y se desecha el sobrenadante. Se resuspende el pellet donde se encuentra la célula OY216 de la invención en 1 ml de agua y se transfiere a un tubo eppendorf para volver a centrifugarlo durante 5 min a 2000 rpm y volver a desechar el sobrenadante. Posteriormente, se resuspende el pellet de la célula OY216 en la solución TELiAc (100 µl/5 ml cultivo) que comprende el tampón TE con 0.1 M de LiAc. A continuación, se mezcla en un eppendorf 50 µl de suspensión de la célula OY216 de la invención con 2 µl (20 µg) de ADN *carrier* y 100 ng del plásmido pGBKT7-GKRP (SEQ ID NO: 6). Adicionalmente, se añaden 250µl de la solución TELiPEG que comprende TE con 0.1 M LiAc y 40% PEG 3350 y se vortea durante 15 seg. Posteriormente, se incuba dicha mezcla durante 30 min a 30°C y transcurrido dicho tiempo se vuelve a incubar durante 15 min a 42 °C. A continuación, se centrifuga 1 min a 13000 rpm y se elimina todo el sobrenadante. Finalmente, se resuspende el pellet en 50 µl de agua y se siembra en una placa con medio de cultivo mínimo SD sin triptófano (SD-T) dado que el marcador de selección del plásmido pGBKT7-GKRP es *TRP1*. Las cepas transformadas que han incorporado el plásmido pGBKT7-GKRP (SEQ ID NO: 6) son visibles a cabo de 3-4 días de cultivo a 30°C.

1.4. Generación y selección simultánea de mutaciones missense mediante PCR-mutagénica y recombinación *in vivo* ("gap-repair") en la secuencia nucleotídica que codifica para la proteína GK comprendida en el plásmido pACT2-GK-3xPTAP (SEQ ID NO: 14).

Para generar mutaciones al azar en la secuencia nucleotídica SEQ ID NO: 9 que codifica para la proteína GK (SEQ ID NO: 10) comprendida en el plásmido pACT2-GK-3xPTAP (SEQ ID NO: 14), se ha utilizado la técnica de PCR mutagénica y recombinación *in vivo* "gap-repair". La generación de mutaciones mediante dicha técnica se basa en la utilización de condiciones de PCR que favorecen la introducción al azar de mutaciones en el producto de PCR obtenido, haciendo uso de cebadores que permiten la amplificación de la secuencia nucleotídica que codifica para la GK, junto con las secuencias flanqueantes presentes en el plásmido pACT2-GK-3xPTAP (SEQ ID NO: 14). La técnica de PCR mutagénica y recombinación *in vivo* se basa en la recombinación, después de transformación, entre un plásmido linearizado y un producto de PCR obtenido en condiciones que favorecen la introducción al azar de

mutaciones. La homología entre los extremos del producto de PCR y los extremos del plásmido pACT2 digerido con BamH1, permiten la recombinación homóloga entre ambos (*“gap repair”*) después de co-transformación en la célula de levadura (**Figura 3**). De esta forma, los transformantes obtenidos comprenden un plásmido, a priori, idéntico a pACT2-GK-3xPTAP, pero con mutaciones introducidas al azar en la

5 secuencia nucleotídica que codifica para la proteína GK.

Los cebadores utilizados en la PCR mutagénica han sido los cebadores de SEQ ID NO: 44 (OV621) (5'-CACTGTCTCACCTGGTTGGACGG-3') y de SEQ ID NO: 45 (OV622) (5'-CTATAGATCAGAGGTTACATGGC-3'), que hibridan 213bp corriente arriba y 174bp corriente abajo de la secuencia nucleotídica (SEQ ID NO: 9) que codifica para la proteína GK (SEQ ID NO: 10) en el plásmido pACT2-GK-3xPTAP (SEQ ID NO: 14).

10

En el presente ejemplo se han llevado a cabo dos PCR mutagénicas para introducir mutaciones al azar en la secuencia nucleotídica que codifica para la proteína GK, siendo la única diferencia entre dichas PCR mutagénicas la enzima ADN polimerasa utilizada en cada una de ellas. Las enzimas ADN polimerasas presentan tasas de mutación diferentes entre ellas, de ahí que se ha querido comparar dos de dichas enzimas para conocer cuál es la más idónea para el método de la invención. En el presente ejemplo se han probado la ADN polimerasa Taq procedente de Takara o Roche, y la ADN polimerasa Mutazyme II procedente del kit de mutagénesis al azar Genemorph II de Agilent.

15

20

Las condiciones para las PCR mutagénicas en función de la ADN polimerasa utilizada se describen a continuación:

25

- ADN polimerasa Mutazyme II (kit de mutagénesis al azar Genemorph II de Agilent). Se utilizan 2.5 µg de ADN molde (pACT2-GK-3xPTAP de SEQ ID NO: 14) y se amplifica durante 20 ciclos para minimizar el número de mutaciones por Kb. Se siguieron las recomendaciones indicadas por el fabricante. Brevemente, la mezcla de reacción (50µl total) comprende: 33.1 µl H₂O + 5 µl buffer 10x + 1 µl dNTPs 10 mM + 0.25 µl de los cebadores de SEQ ID NO: 44 y SEQ ID NO: 45 (100 µM) + 9.4 µl pACT2-GK-3xPTAP (2.5 µg) + 1 µl Mutazyme II (ADN pol). Los ciclos de la PCR fueron: (1) 95 °C 2 min; (2) 95 °C 30 sec + 55 °C 30 sec + 72 °C 1 min/Kb (repetir 20 veces el paso (2)) y (3) 72 °C 10 min.
- 30

- ADN polimerasa Taq (Takara Ref. R001 o Roche Ref 11647679001). Se han utilizado condiciones estándar para minimizar el número de mutaciones por Kb. La mezcla de reacción (50 µl total) comprende: 42 µl H₂O + 5 µl buffer + 1 µl dNTPs 10 mM + 0.25 µl de los cebadores de SEQ ID NO: 44 y SEQ ID NO: 45 (100 µM) + 1 µl pACT2-GK-3xPTAP (265 ng) + 0.5 µl Taq (ADN pol). Los ciclos de la PCR fueron (1) 94 °C 2 min; (2) 94 °C 30 sec + 55 °C 30 sec + 72 °C 1 min/Kb (repetir 30 veces el paso (2)) y (3) 72 °C 7 min.

Después de cada reacción de PCR mutagénica, se añadió 1 µl de la enzima *dpn1* (Roche) y se incubó la mezcla durante 1 h a 37°C para eliminar el ADN molde. Posteriormente se precipitó el producto de PCR con acetato de sodio y etanol, para después resuspenderlo en 20 µl H₂O.

Para la técnica de la PCR mutagénica y recombinación *in vivo* se lineariza el plásmido pACT2 (SEQ ID NO: 13) (Clontech) en el *polylinker* con el enzima de restricción BamH1 (**Figura 3**). Para dicha linearización del plásmido pACT2 en el *polylinker* se podrían utilizar otras enzimas de restricción tales como Nco1, Sma1, EcoR1, Sac1 o Xho1. La mezcla de digestión (100 µl total) comprende: 80 µl H₂O + 10 µl buffer 10x + 7.4 µl pACT2 (1.74 µg) + 2.5 µl (25u) BamH1 (Roche). Se incubó la mezcla durante 2h a 37°C. Posteriormente se precipitó el ADN con acetato de sodio y etanol, para después resuspenderlo en 20 µl H₂O.

A continuación, se lleva a cabo la transformación de la célula OVY216, previamente pre-transformada con el plásmido pGBKT7-GKRP de SEQ ID NO: 6, con el plásmido pACT2 linearizado con BamH1 y los diferentes productos obtenidos en las PCR mutagénicas, según se ha descrito anteriormente.

La generación de mutaciones en el plásmido pACT2-GK-3xPTAP (SEQ ID NO: 14) mediante la técnica de PCR mutagénica y recombinación *in vivo*, así como la selección de las mutaciones “*missense*” en la proteína GK que producen una pérdida de interacción con la proteína GKRP, se llevan a cabo en un solo paso mediante la utilización combinada de genes reporteros, específicamente en la presente invención mediante los genes reporteros *SPAL10-URA3* e *(lexAop)4-HIS3*, y el medio selectivo SD-AHTL+5FoA+3-AT, que permiten exclusivamente el crecimiento de los transformantes en los que la mutación en el gen de la GK bloquea su interacción con

GKRP, pero no trunca la proteína (mutación *missense*). El compuesto 3-AT (3-aminotriazol) es un inhibidor competitivo del producto del gen *HIS3* que se utiliza en el presente ejemplo para evitar que la actividad basal del reportero *LexA(op)-HIS3* sea suficiente para permitir el crecimiento de la célula en ausencia de histidina. Se ha
5 utilizado una concentración del compuesto 3-AT de 1 mM pero se ha observado que utilizando dicha concentración existen células OVY216 transformadas que son capaces de crecer en ausencia de histidina debido a la actividad basal del reportero *LexA(op)-HIS3*, aunque las colonias obtenidas son mucho más pequeñas que las obtenidas con los transformantes que comprenden la secuencia que codifica para la
10 proteína GK con mutaciones *missense*. Para evitar este inconveniente se ha probado una concentración de 3-AT de 5mM. Esta mayor concentración es capaz de inhibir completamente el crecimiento de estos transformantes.

Brevemente, la célula de levadura de la invención OVY216, previamente transformada
15 con el plásmido pGBKT7-GKRP (SEQ ID NO: 6), se transforma ahora con el plásmido pACT2 linearizado con BamH1 junto con los productos de la PCR mutagénica. Para ello, dicha célula se inocula en 5ml de medio de cultivo mínimo sin triptófano (SD-T) y se mantiene en cultivo hasta que su crecimiento llegue a la fase logarítmica. Este cultivo asegura el crecimiento de las células que comprenden el plásmido pGBKT7-
20 GKRP y en consecuencia, expresarán la proteína de fusión GBD-GKRP (SEQ ID NO: 51). A continuación, se centrifuga dicho cultivo y se resuspende el pellet de las células en 15 ml de medio completo (YPAD) para mejorar la eficiencia de la transformación. Se mantiene dicho cultivo hasta que llegue de nuevo a fase logarítmica (alrededor de 3-4h). A continuación, se sigue el protocolo de transformación con acetato de litio
25 según se ha explicado previamente, utilizando como mezcla de transformación: 50µl de suspensión de la célula de levadura en TELiAc + 2 µl (20 µg) de ADN carrier + 2 µl (175 ng) de pACT2 linearizado con BamH1 + 2 µl (Taq) o 4 µl (Mutazyme II) de PCR mutagénico. Se utiliza una concentración alta del producto de PCR mutagénico respecto al plásmido linearizado para favorecer su integración mediante
30 recombinación o “*gap repair*” (**Figura 3**). Para incrementar aún más la eficiencia del procedimiento de transformación y asegurar la expresión de los marcadores de selección *URA3* e *HIS3* al final del protocolo de transformación, se resuspende el pellet de las células de levadura transformadas en 5ml de medio completo YPAD y se cultiva a 30°C durante 150 min. Posteriormente, se centrifuga durante 1 min a 13000
35 rpm y se elimina todo el sobrenadante. A continuación, se resuspende el pellet de

levadura en 50µl de agua y se siembra el 1% de la mezcla de transformación en una placa control con medio mínimo no selectivo sin triptófano, leucina, ni adenina (SD-TLA) para calcular el número total de transformantes que contienen los dos plásmidos pGBKT7-GKRP (SEQ ID NO: 6) y pACT2-GK-3xPTAP obtenido mediante
5 recombinación ("*gap repair*"), incluyendo los transformantes que no llevan mutaciones o los que llevan mutaciones sin el efecto deseado. Este medio selectivo carece de adenina ya que la cepa es protótrofa para este requerimiento y carece de triptófano y leucina dado que los marcadores de los dos plásmidos son *TRP1* y *LEU2*. Se ha obtenido una media de 500 transformantes en esta placa control, lo que indica que el
10 número total de transformantes rastreados es de 50000, al haberse sembrado el 1% de la mezcla de transformación.

Para seleccionar las células transformadas que expresan la secuencia mutante de la proteína GK que bloquea su unión a la proteína GKRP (mutaciones *missense*) (**Figura**
15 **1**), se siembra el resto de la mezcla de transformación (99%) en una placa de cultivo con medio mínimo selectivo sin triptófano, leucina, adenina e histidina y con 0.1% 5-FoA y 1mM 3-AT (SD-AHTL+5-FoA+3-AT). En dicho medio selectivo solo crecerán aquellas células OVY216 transformadas según se describe en el presente ejemplo, que presenten mutaciones *missense* en la secuencia que codifica para la proteína GK
20 y que, por tanto, impidan su unión a la proteína GKRP. Por el contrario, las células OVY216 transformadas con el plásmido pACT2 que ha quedado sin digerir o que se ha recircularizado, así como las células transformadas con el plásmido pACT2-GK-3xPTAP no mutado (SEQ ID NO: 14) o con mutaciones que producen un truncamiento en la proteína GK, no son capaces de crecer en dicho medio selectivo SD-AHTL+5-
25 FoA+3-AT. Adicionalmente, también se ha observado que con 1mM 3-AT en el medio de cultivo selectivo, el bloqueo del crecimiento de los transformantes que no activan el reportero *HIS3* no es absoluto, existiendo una pequeña actividad basal del reportero (*lexAop*)4-*HIS3*, aunque las colonias obtenidas son mucho más pequeñas que las obtenidas con los transformantes que comprenden la secuencia que codifica para la
30 proteína GK con mutaciones *missense*. Para evitar este inconveniente se ha probado una concentración de 3-AT de 5mM. Esta mayor concentración es capaz de inhibir completamente el crecimiento de estos transformantes.

Es interesante mencionar que el número de transformantes obtenidos en el medio
35 selectivo SD-AHTL+5-FoA+3-AT varía según la ADN polimerasa utilizada para

introducir mutaciones al azar en la secuencia nucleotídica que codifica para la proteína GK en la PCR mutagénica. Así, cuando se utiliza la ADN polimerasa Mutazyme II, el número de transformantes en el medio selectivo es el 1.25% del total de los transformantes rastreados. Sin embargo, con la ADN polimerasa Taq, se han obtenido 10 veces menos clones positivos (0.125% del total de transformantes) que los obtenidos con Mutazyme II. Esta diferencia se debe probablemente al hecho de que con la ADN polimerasa Taq se han utilizado condiciones estándar que no favorecen la aparición de mutaciones, con el objetivo de limitar el número de mutaciones por plásmido.

1.5. Análisis de las células OVY216 transformadas según se ha descrito previamente que han crecido en el medio selectivo SD-AHTL+5-FoA+3-AT y purificación de los plásmidos que comprenden las mutaciones *missense* en la secuencia que codifica para la proteína GK humana.

Se ha extraído el ADN plasmídico de 19 transformantes obtenidos en el medio selectivo SD-AHTL+5-FoA+3-AT. En estos 19 transformantes (10 obtenidos con la DNA polimerasa mutazyme II y 9 obtenidos con la DNA polimerasa Taq), que presentan las mutaciones en la secuencia nucleotídica que codifica para la proteína GK, obtenidas mediante la técnica de PCR mutagénica y recombinación *in vivo* "gap-repair", y seleccionadas mediante el método de doble híbrido en reverso descrito en la presente invención, se ha aislado el plásmido pACT2-GK-3xPTAP mutado que comprendían dichos transformantes.

Brevemente, dado que *LEU2* es el marcador para seleccionar este plásmido en levaduras, cada transformante se cultiva en 5mL de medio mínimo sin leucina (SD-L) hasta que llega a la fase estacionaria de su crecimiento. A continuación, se aísla el ADN plasmídico mediante el kit comercial "High Pure Plasmid Isolation Kit" (Roche), que se ha adaptado para la extracción de plásmidos en levadura, ya que dicho kit es específico para purificación de plásmidos en bacterias. Se resuspendió el pellet de los transformantes de levaduras seleccionados en el medio selectivo SD-AHTL+5-FoA+3-AT en 250 µl de la solución 1 del kit "High Pure Plasmid Isolation Kit" (Roche) y se añadieron 10µl de zimoliasa 50 mg/ml que se incubó durante 30 min a 37 °C para digerir la pared celular de la levadura. Posteriormente, se añadió la solución 2 de dicho kit y se siguieron las instrucciones del fabricante hasta la obtención del ADN

plasmídico purificado. El ADN plasmídico purificado de cada transformante comprende una mezcla del plásmido pGBKT7-GKRP (que comprende un marcador de resistencia a kanamicina) y del plásmido pACT2-GK-3xPTAP con la secuencia génica GK mutada (que comprende un marcador de resistencia a ampicilina). Posteriormente, de dicha
5 mezcla se aísla el plásmido con la secuencia nucleotídica que codifica para la GK mutada. Para ello, se transforma mediante la metodología estándar del cloruro de rubidio la cepa de *E. coli* DH5 α con 10 μ l de la mezcla del ADN plasmídico extraído tal y como se ha mencionado anteriormente. Después de un choque térmico de 2 min a 37°C, se cultivan dichas células bacterianas transformadas en una placa con medio de
10 cultivo LB en presencia de 50 μ g/ml ampicilina para seleccionar únicamente aquéllos transformantes bacterianos que habían incorporado el plásmido pACT2-GK-3xPTAP que presentaba las mutaciones *missense*.

A continuación, se purificó el plásmido pACT2-GK-3xPTAP que comprende las
15 mutaciones *missense* con el kit “*High Pure Plasmid Isolation Kit*” (Roche). Para ello, se cultivó un transformante bacteriano de cada placa de transformación en 5mL de medio LB con ampicilina y se purificó el plásmido siguiendo las instrucciones del fabricante. Finalmente, se comprobó con el enzima de restricción BamH1 que el plásmido pACT2-GK-3xPTAP mutado presenta el mismo patrón de restricción que el plásmido parental
20 pACT2-GK-3xPTAP (SEQ ID NO: 14) no mutado. En el presente ejemplo, los 19 plásmidos purificados que presentan la secuencia nucleotídica GK mutada mediante la técnica de PCR mutagénica y recombinación *in vivo* “*gap-repair*”, liberaron 2 fragmentos de 8.2 Kb y 1.4 Kb cada uno, al igual que el plásmido pACT2-GK-3xPTAP parental (SEQ ID NO: 14) que no presentaba las mutaciones.

25

1.6. Validación del método de doble híbrido en reverso de la invención con el sistema de doble híbrido clásico.

Mediante el sistema de doble híbrido clásico (ensayo de actividad β -galactosidasa en filtro) se ha confirmado que los 19 plásmidos pACT2-GK-3xPTAP que comprenden la
30 secuencia de la proteína GK mutada, no interaccionan con el plásmido pGBKT7-GKRP, es decir existe una pérdida de la interacción entre las proteínas GK-GKRP, esto indica que la mutación en la secuencia nucleotídica que codifica para la proteína GK es una mutación *missense* que bloquea su interacción con la proteína GKRP. Pero, por otro lado, mediante el sistema de doble híbrido clásico se ha confirmado que
35 los plásmidos pACT2-GK-3xPTAP que comprenden la secuencia de la proteína GK

mutada, si interaccionan con el plásmido pLexA(1-202)PL-TSG101 de SEQ ID NO: 33, a través de la unión a TSG101-3xPTAP, lo que demuestra que estas mutaciones son *missense* y no producen un truncamiento de la proteína.

- 5 Para validar los resultados obtenidos mediante el sistema de doble híbrido en reverso descrito en la presente invención, ejemplificado en la detección de mutaciones *missense* que bloquean la interacción entre las proteína GK codificada por el plásmido pACT2-GK-3xPTAP y la proteína GKPR codificada por el plásmido pGBKT7-GKRP, mediante el sistema de doble híbrido clásico, se ha utilizado la cepa de levadura Y187
10 (Clontech) que comprende el reportero *lacZ* bajo el control de UASGal. Por otro lado, para validar la unión entre la proteína TSG101 codificada por el plásmido pLexA(1-202)PL-TSG101 (SEQ ID NO: 33), y el péptido 3xPTAP codificado por el plásmido pACT2-GK-3xPTAP, que comprende la secuencia GK mutada, se ha utilizado la cepa de levadura CTY10-5d (*Cellular interactions in development: a practical approach* ed.
15 D.A. Hartley Oxford: Oxford University Press 153-17) que comprende el reportero *lacZ* bajo el control de *lexAop*. En ambos casos, se han realizado ensayos de la actividad β -galactosidasa en filtro para cuantificar la activación de los reporteros.

Brevemente, para llevar a cabo dicho análisis de interacción, se transformaron las
20 cepas de levadura mencionadas, Y187 (Clontech) y CTY10-5d, siguiendo el protocolo con acetato de litio descrito previamente, pero utilizando las mezclas siguientes:

- 50 μ l de una suspensión de la cepa de levadura Y187 (Clontech) en TELiAc + 2 μ l (20 μ g) de ADN carrier + 1 μ l (100 ng) del plásmido pGBKT7-GKRP (SEQ ID NO: 6) + 1 μ l (100 ng) del plásmido pACT2-GK-3xPTAP mutante aislado según
25 se ha descrito previamente. Como controles negativos y positivos, hemos co-transformado también pGBKT7-GKRP + pACT2, así como pGBKT7-GKRP + pACT2-GK-3xPTAP (original, no mutado). La mezcla de transformación se sembró en una placa de medio mínimo sin triptófano, leucina y uracilo (SD-TLU) ya que la cepa Y187 es capaz de sintetizar uracilo y los marcadores de
30 los dos plásmidos son *TRP1* y *LEU2*.
- 50 μ l de suspensión de la cepa de levadura CTY10-5d en TELiAc + 2 μ l (20 μ g) de ADN carrier + 1 μ l (100ng) del plásmido pLexA(1-202)PL-TSG101 + 1 μ l (100 ng) del plásmido pACT2-GK-3xPTAP mutante. Como controles negativos y positivos, hemos co-transformado también pLexA(1-202)PL-TSG101 +
35 pACT2, así como pLexA(1-202)PL-TSG101 + pACT2-GK-3xPTAP (original, no

mutado). La mezcla de transformación se sembró en una placa de medio mínimo sin histidina, leucina y uracilo (SD-HLU) ya que la cepa CTY10-5d es capaz de sintetizar uracilo y los marcadores de los dos plásmidos son *HIS3* y *LEU2*.

- 5 Después de 3-4 días de cultivo a una temperatura de 30°C, replicamos de 8-10 transformantes de cada mezcla de transformación a placas con los medios de cultivo SD-TLU o SD-HLU, y se mantuvieron en cultivo durante 2 días a 30 °C. Transcurrido dicho tiempo se llevó a cabo el ensayo de actividad β -galactosidasa en filtro (Breedon L, Nasmyth K. Cold Spring Harb Symp Quant Biol. 1985;50:643-50) para cuantificar la
- 10 activación de los reporteros. Para ello se colocó un filtro de nitrocelulosa encima de la placa de cultivo para transferir, mediante aplicación suave, la levadura al filtro. Se incubó el filtro a -80 °C durante 1 h para permeabilizar las células y posteriormente se colocó el filtro encima de papel Whatman 3MM mojado en tampón Z (Na_2HPO_4 60 mM, NaH_2PO_4 40 mM, KCl 10 mM, MgSO_4 1 mM, β -mercaptoetanol 38 mM) con 0.1% X-
- 15 Gal y se incubó durante 1 h a 30 °C. La interacción entre las proteínas produce la activación del gen reportero *lacZ*, y en consecuencia la expresión de la β -galactosidasa y la hidrólisis del X-Gal en un compuesto de color azul.

- Los resultados obtenidos indican que en todos los casos, es decir con los 19
- 20 plásmidos pACT2-GK-3xPTAP que comprenden la secuencia codificante de la proteína GK mutada, no se detecta actividad β -galactosidasa en la cepa Y187 co-transformada con cada uno de estos plásmidos y con pGBKT7-GKRP, lo que confirma el hecho de que las proteínas GKRP y GK mutadas no interaccionan en el sistema de doble-híbrido clásico y no activan el reportero que comprende *lacZ* bajo el control de
- 25 UASGal. Este resultado indica que la selección de mutaciones en GK que bloquean su interacción con GKRP ha sido eficiente al 100% con este método. Por otra parte, los resultados obtenidos indican que en todos los casos, se detecta actividad β -galactosidasa en la cepa CTY10-5d co-transformada con cada uno de estos plásmidos y pLexA(1-202)PL-TSG101, lo que confirma el hecho de que las proteínas TSG101 y
- 30 GK mutadas siguen interaccionando mediante la unión entre TSG101 y el péptido 3xPTAP fusionado al extremo C-terminal de GK y activan el reportero que comprende *lacZ* bajo el control de *lexAop*. Este resultado demuestra que la selección de mutaciones *missense* en GK que no truncan la proteína y en consecuencia no eliminan el péptido 3xPTAP ha sido eficiente al 100% con este método. En conjunto, estos
- 35 resultados muestran que la selección de mutaciones *missense* en GK que bloquean su

interacción con GGRP ha sido eficiente al 100% con este método. En la **Figura 4**, se presentan los ensayos de la actividad β -galactosidasa en filtro para 10 de los mutantes seleccionados.

5 Ejemplo 2. Identificación de las mutaciones *missense* presentes en la secuencia que codifica para la proteína GK y que impiden su unión a la proteína GGRP.

Las mutaciones *missense* en los 19 plásmidos mutados aislados y validados con el sistema de doble-híbrido clásico según se ha descrito anteriormente, han sido
10 identificadas mediante secuenciación del inserto que comprende el gen de la GK en cada plásmido. Los oligonucleótidos utilizados para la secuenciación y la identificación de las mutaciones *missense* en la secuencia nucleotídica del gen que codifica para la proteína GK fueron los siguientes: SEQ ID NO: 46 (OV284) (5'-CGATGATGAAGATACCCCACC-3'), que hibrida 67bp corriente arriba del *polylinker*
15 donde la secuencia nucleotídica del gen GK está clonada y SEQ ID NO: 47 (OV285) (5'-GAGATGGTGCACGATGCACAG-3'), que hibrida 120bp corriente abajo del *polylinker* donde la secuencia nucleotídica del gen GK está clonada.

Por otro lado, también se han identificado mediante secuenciación las mutaciones
20 presentes en 5 transformantes adicionales, diferentes a los 19 anteriores y procedentes de la transformación con el producto de PCR obtenido con la ADN polimerasa Taq. Para estos 5 transformantes adicionales, el plásmido analizado no ha sido extraído ni validado previamente en el sistema de doble-híbrido clásico. El inserto que comprende el gen de la GK mutado ha sido amplificado directamente desde el
25 ADN plasmídico extraído de los 5 transformantes adicionales y secuenciado. Teniendo en cuenta la alta eficiencia del sistema de doble híbrido en reverso descrito en la presente invención (100% de positivos detectados-100% mutaciones de tipo *missense* con los 19 transformantes), el hecho de poder amplificar y secuenciar directamente el ADN plasmídico, sin tener que aislar el plásmido y validarlo permite ahorrar tiempo. La
30 extracción del plásmido mutado y su validación con el sistema de doble-híbrido clásico puede llevarse a cabo posteriormente, una vez que la mutación haya sido identificada, lo que evita validar mutantes repetidos o ya conocidos.

La identificación de las mutaciones en los 5 transformantes adicionales ha sido
35 realizada de la forma siguiente: Extracción del ADN plasmídico de los transformantes

de levadura seleccionados mediante el sistema de doble híbrido en reverso de la invención (ver apartado 1.5) y utilización de este ADN como molde para amplificar el gen de la GK mediante PCR con los oligonucleótidos utilizados para la secuenciación (SEQ ID NO: 46 (OV284) y SEQ ID NO: 47 (OV285)). Posteriormente se secuenciar el producto de PCR amplificado con estos mismos oligonucleótidos para identificar la ó las mutaciones presentes en el inserto.

La secuenciación del total de los 24 clones analizados (19+5) junto con los ensayos de validación con el sistema de doble híbrido clásico ha demostrado que todos ellos comprenden mutaciones *missense* que provocan una pérdida de interacción entre la GK y la GKRP. Esto demuestra la eficacia del método ya que el 100% de los clones analizados, un total de 24, comprenden mutaciones *missense* en el gen de la GK que bloquean su interacción con la proteína GKRP.

Los resultados obtenidos con la secuenciación de los 24 mutantes analizados ha permitido la identificación de 21 mutaciones *missense* responsables de la pérdida de la interacción de la proteína GK con la proteína GKRP (Tabla 1). Varias de estas mutaciones están repetidas y algunos clones que comprendían varias mutaciones no han sido caracterizados.

Tabla 1. *Mutaciones missense identificadas en la proteína GK que impiden la interacción con la proteína GKRP (substituciones aa).*

Substitución	Repetición
A201V	1
T60A	2
G72W	1
G72E	1
S64P	4
L75P	4
T209P	1
C220R	1
C233Y	1
C252Y	3
G407D	1

Substitución	Repetición
L306R	1

La secuenciación de los 24 mutantes muestra que la frecuencia de mutaciones *missense* en cada clon es distinta en función de la ADN polimerasa utilizada. La situación óptima sería obtener una única mutación *missense* en la secuencia codificante de la GK (1.5 Kb). En el caso de que sean varias, es necesario analizar en el sistema de doble híbrido clásico el efecto de cada mutación de forma aislada para saber cuál es la responsable de la pérdida de interacción. En este caso, los plásmidos que comprendían una sola mutación se han obtenido mediante mutagénesis dirigida.

Utilizando la ADN polimerasa Mutazyme II, se han obtenido cerca de 2 mutaciones *missense* por Kb de media. En el caso de la ADN polimerasa Taq, se han obtenido una media de 1 mutación *missense* por Kb (la mitad que con Mutazyme II). Esta diferencia es consistente con el hecho de que se han obtenido 10 veces menos transformantes en el medio selectivo SD-AHTL+5-FoA+3-AT cuando se ha utilizado la ADN polimerasa Taq que cuando se ha utilizado la ADN polimerasa Mutazyme II. Aunque el resultado con Taq es mejor en cuanto a frecuencia de mutaciones, esta polimerasa presenta un espectro mutacional con una fuerte tendencia para los cambios AT a GC, lo que hace que muchas de las mutaciones obtenidas se repitan. Así, tal y como muestran los resultados obtenidos, la ADN polimerasa Taq da lugar a un 70% de mutaciones *missense* repetidas, mientras que la ADN polimerasa Mutazyme II solo da lugar a un 40% de mutaciones repetidas.

Una vez identificadas las mutaciones *missense* que impiden la unión entre las proteínas GK y GKRP, se procedió a comparar dichas mutaciones con la estructura ya conocida del complejo GK-GKPR (Choi JM *et al.* Proc Natl Acad Sci U S A. 2013 Jun 18;110(25):10171-6), poniéndose de manifiesto la buena correlación existente entre los resultados obtenidos en la identificación de mutaciones *missense* mediante el método de doble híbrido en reverso de la invención y la estructura del complejo, ya que:

- los residuos aminoacídicos Ala201 y Thr60 se localizan junto con el bolsillo hidrofóbico de la GK que media la interacción con GKRP (Proc Natl Acad Sci U S A. 2006 Sep 5;103(36):13368-73),

- el residuo aminoacídico Gly72 está próximo a los residuos aminoacídicos Ala201 y Thr60, y el puente de hidrogeno que forma con el residuo aminoacídico Tyr215 parece estabilizar la conformación súper-abierta de la GK que interacciona con GKRP (J Biol Chem. 2006 Dec 29;281(52):40201-7),
- 5 – las sustituciones a prolina de los residuos S64P, L75P y T209P, pueden tener un efecto indirecto en la rotura de la interacción (vía Thr60, Gly72 y Ala201), ya que el residuo Ser64 se localiza en la lámina beta adyacente a Thr60 mientras que Leu75 y Thr209 se localizan en láminas beta adyacentes a los motivos conteniendo Gly72 y Ala201, y debido a la interrupción de estas láminas beta
- 10 por los residuos de prolina,
- los residuos de cisteína: Cys220, Cys233 y Cys252, forman parte de un anillo de 5 residuos de cisteína, próximo al bolsillo hidrofóbico y probablemente implicados en la formación de puentes disulfuro y en el mantenimiento de la conformación de la GK (Arch Biochem Biophys. 2000 Mar 15;375(2):251-60).
- 15 Además, el residuo aminoacídico Gly407 se localiza junto al anillo de residuos de cisteína.

Ejemplo 3. Comparación entre sistema de doble híbrido en reverso de la invención que comprende el gen reportero de contra-selección *URA3* respecto al mismo sistema de doble híbrido en reverso pero utilizando como sistema de contra-selección el gen represor TetR.

Para poner de manifiesto que los reporteros de selección utilizados en el sistema de doble híbrido de la invención presentan ventajas adicionales sorprendentes respecto de otros reporteros conocidos en el estado de la técnica, se ha probado un sistema alternativo a la contra-selección con el gen reportero *URA3* y 5-FoA, basado en el gen represor TetR (Shih HM *et al.* Proc Natl Acad Sci U S A. 1996;93:13896-901). En este sistema de doble híbrido en reverso alternativo, la unión de las proteínas interaccionantes a estudiar activa la transcripción de TetR que reprime la transcripción del reportero *ADE2*. En consecuencia, siguiendo con el ejemplo mostrado en la presente invención, una mutación *missense* en la secuencia codificante de la proteína GK que bloquea su interacción con GKRP activa el reportero *ADE2*, lo que permite el crecimiento de la cepa en ausencia de adenina. Para probar este sistema, se ha integrado la proteína de fusión *ADH1::LexA-TSG101* en el locus *URA3* de la cepa LY26 (Thomas LR *et al.* J Biol Chem. 2002;277:34343-8) para obtener la cepa

OVY158 (*MATalpha can1 his3 leu2 met15 trp1 ura3 gal4::hisG gal80::hisG LYS2::LexA(op)-HIS3 TetO-ADE2 ho::KanMX::GAL1-TetR URA3::LexA-TSG101*).

En este caso, la pérdida de interacción entre las proteínas GK y GKRP se selecciona
5 mediante crecimiento en un medio selectivo con ausencia de adenina (medio de
cultivo mínimo (SD) con arginina y metionina como únicos requerimientos y 1 mM 3-
AT) en vez de crecimiento en presencia de 5-FoA. Se ha observado que este sistema
de doble híbrido en reverso utilizando como método de contra-selección TetR en lugar
de *URA3* y 5-FoA no es tan eficiente como el sistema de doble híbrido en reverso con
10 5-FoA, ya que el crecimiento de los clones no mutados no está completamente
bloqueado apareciendo en los cultivos colonias más pequeñas. Para intentar
solucionar este problema se ha utilizado sacarosa como fuente de carbono, en lugar
de glucosa, intentando así eliminar el efecto negativo de la glucosa sobre el promotor
GAL1p y, en consecuencia la expresión del reportero TetR, se ha resuelto sólo
15 parcialmente este problema, pero este sistema es menos eficiente que el sistema que
utiliza *URA3* y 5-FoA, ya que mediante el uso de TetR se obtienen un 15% de falsos
positivos, mientras que, como hemos mencionado anteriormente, mediante el sistema
de doble híbrido en reverso utilizando *URA3*, *HIS3* y 5-FoA, descrito en la presente
invención no se obtienen falsos positivos.

REIVINDICACIONES

1. Método *in vitro* para identificar al menos una mutación *missense* en una proteína de referencia donde dicha mutación afecta a la capacidad de unión de
5 dicha proteína de referencia con otra proteína diana, donde dicho método comprende:
- a) al menos una célula huésped que comprende integrado en su genoma:
- i) Una primera secuencia nucleotídica que codifica para un gen
10 reportero, donde dicha secuencia nucleotídica está operativamente unida a un promotor que comprende una secuencia nucleotídica reconocida por una proteína que se une al ADN,
- ii) Una segunda secuencia nucleotídica que codifica para un
15 segundo gen reportero, donde dicha segunda secuencia nucleotídica está operativamente unida a un promotor que comprende una secuencia nucleotídica reconocida por una proteína que se une al ADN, con la condición de que dicha secuencia nucleotídica sea reconocida por un dominio de unión al ADN distinto del de i), y
- 20 iii) Una tercera secuencia nucleotídica que codifica para una primera proteína de fusión que comprende el dominio de unión al ADN de ii) y una proteína heteróloga, capaz de unirse a un péptido funcional localizado en el extremo C-terminal de la proteína de referencia de estudio, donde la secuencia
25 nucleotídica que codifica para dicha primera proteína de fusión está operativamente unida a un promotor,
- b) Pre-transformar la célula de la etapa a) con un plásmido que comprende la secuencia nucleotídica que codifica para una segunda proteína de fusión que comprende el dominio de unión al ADN de i) y la proteína
30 diana, donde la secuencia nucleotídica que codifica para dicha segunda proteína de fusión está operativamente unida a un promotor,
- c) Cultivar la célula de la etapa b) bajo condiciones que permitan

exclusivamente el crecimiento de las células que hayan incorporado el plásmido de la etapa b),

- 5 d) Transformar la célula de la etapa c) con un vector linearizado y al menos un fragmento de ADN que previamente ha sido sometido a mutagénesis y que comprende la secuencia nucleotídica que codifica para la proteína de referencia con al menos una mutación, donde entre el vector linearizado y el fragmento de ADN sometido a mutagénesis se produce recombinación homóloga obteniéndose un plásmido que comprende la secuencia nucleotídica que codifica para una tercera
- 10 proteína de fusión que comprende el dominio de transactivación de Gal4 y la proteína de referencia con al menos una mutación, donde la secuencia nucleotídica que codifica para dicha tercera proteína de fusión está operativamente unida a un promotor,
- 15 e) Cultivar la célula de la etapa d) bajo condiciones tales que permitan exclusivamente el crecimiento de las células que presentan mutaciones *missense* que impiden la unión entre la proteína de referencia y la proteína diana,
- 20 f) Comparar la secuencia de la proteína de referencia mutada con la secuencia de la proteína de referencia *wild-type* e identificar la mutación *missense* que impide la unión de la proteína de referencia con la proteína diana.
2. Método según la reivindicación 1 donde el procedimiento de mutagénesis al que ha sido sometido el fragmento de ADN de la etapa d) es un procedimiento de mutagénesis *in vitro* o *in vivo*.
- 25 3. Método según la reivindicación 1 donde el plásmido de la etapa d) se obtiene mediante recombinación *in vivo* "gap-repair".
4. Método según cualquiera de las reivindicaciones 1 a 3 donde el gen reportero se selecciona de entre genes reporteros de selección positiva y genes reporteros de contra-selección.
- 30 5. Método según la reivindicación 4 donde los genes reporteros de selección positiva se seleccionan del grupo que consiste en: *HIS3*, *LEU2*, *URA3*, *ADE2*, *TRP1*, *LYS2* y *LYS5*; y los genes reporteros de contra-selección se seleccionan

del grupo que consiste en: *URA3*, *TRP1*, *LYS2*, *LYS5*, *CYH2*, *CAN1*, *GAL1* y *mazF*.

6. Método según la reivindicación 5 donde el gen reportero de selección positiva es el gen *HIS3*, y el gen reportero de contra-selección es el gen *URA3*
- 5 7. Método según cualquiera de las reivindicaciones 1 a 6 donde las secuencias de reconocimiento para dominios de unión al ADN comprenden los sitios de unión de una proteína seleccionada de entre cualquiera de las siguientes: Gal4, LexA y Ace1.
- 10 8. Método según cualquiera de las reivindicaciones 1 a 7 donde la secuencia nucleotídica que codifica para un dominio de transactivación comprende los dominios de activación transcripcional de una proteína seleccionada de entre cualquiera de las siguientes: Gal4, VP16 y Ace1.
9. Método según cualquiera de las reivindicaciones 1 a 8 donde los promotores son promotores constitutivos.
- 15 10. Método según la reivindicación 9 donde los promotores constitutivos se seleccionan de entre cualquiera de la lista que consiste en: *ADH1*, *PGK1*, *TEF1*, *TPI1*, *HXT7*, *TDH3* y *PYK1*.
- 20 11. Método según cualquiera de las reivindicaciones 1 a 10 donde la proteína heteróloga es la proteína de SEQ ID NO: 32, y el péptido funcional es el péptido de SEQ ID NO: 12.
12. Método según cualquiera de las reivindicaciones 1 a 11 donde la primera proteína de fusión es la proteína de SEQ ID NO: 49.
13. Método según cualquiera de las reivindicaciones 1 a 12 donde la segunda proteína de fusión es la proteína de SEQ ID NO: 51.
- 25 14. Método según cualquiera de las reivindicaciones 1 a 13 donde la célula se selecciona del grupo que consiste en: célula de levadura, célula bacteriana y célula de mamífero.
- 30 15. Método según la reivindicación 14 donde la célula de levadura se selecciona del grupo que consiste en *Yarrowia lipolytica*, *Pichia pastoris* y *Scacharomyces cerevisiae*, preferentemente la célula de levadura pertenece a la especie *S.*

cerevisiae.

16. Método según cualquiera de las reivindicaciones 1 a 15 donde la célula es OVY216.
- 5 17. Método según cualquiera de las reivindicaciones 1 a 16 caracterizado por que las células que presentan al menos una mutación *missense*, crecen en un medio de cultivo que comprende 5-FoA y 3-AT, y carece de triptófano, leucina, adenina e histidina.
- 10 18. Método según cualquiera de las reivindicaciones 1 a 17 donde la identificación de las mutaciones *missense*, se lleva a cabo mediante técnicas de amplificación y secuenciación.
19. Célula huésped que comprende integrado en su genoma:
- 15 i) Una primera secuencia nucleotídica que codifica para un gen reportero, donde dicha secuencia nucleotídica está operativamente unida a un promotor que comprende una secuencia de reconocimiento para una proteína que se une al ADN,
- 20 ii) Una segunda secuencia nucleotídica que codifica para un segundo gen reportero, donde dicha segunda secuencia nucleotídica está operativamente unida a un promotor que comprende una secuencia de reconocimiento para una proteína que se une al ADN, con la condición de que dicha secuencia de reconocimiento sea diferente de la secuencia de reconocimiento de i), y
- 25 iii) Una tercera secuencia nucleotídica que codifica para una primera proteína de fusión que comprende el dominio de unión al ADN de ii) y una proteína heteróloga capaz de unirse a un péptido funcional localizado en el extremo C-terminal de la proteína de referencia, donde la secuencia nucleotídica que codifica para dicha primera proteína de fusión está operativamente unida a un promotor.
- 30 20. Célula huésped según la reivindicación 17, donde el gen reportero se selecciona de la lista que consiste en genes reporteros de selección positiva y genes reporteros de contra-selección.

21. Célula huésped según la reivindicación 20 donde los genes reporteros de selección positiva se seleccionan del grupo que consiste en: *HIS3*, *LEU2*, *URA3*, *ADE2*, *TRP1*, *LYS2*, *LYS5*; y los genes reporteros de contra-selección se seleccionan del grupo que consiste en: *URA3*, *TRP1*, *LYS2*, *LYS5*, *CYH2*, *CAN1*, *GAL1*, *mazF*.
22. Célula huésped según la reivindicación 21 donde el gen reportero de selección positiva es el gen *HIS3*, y el gen reportero de contra-selección es el gen *URA3*
23. Célula huésped según cualquiera de las reivindicaciones 19 a 22 donde las secuencias de reconocimiento para dominios de unión al ADN comprenden los sitios de unión de una proteína seleccionada de entre cualquiera de las siguientes: Gal4, LexA y Ace1.
24. Célula huésped según cualquiera de las reivindicaciones 19 a 23 donde la secuencia nucleotídica que codifica para un dominio de transactivación comprende los dominios de activación transcripcional de una proteína seleccionada de entre cualquiera de las siguientes: Gal4, VP16 y Ace1.
25. Célula huésped según cualquiera de las reivindicaciones 19 a 24 donde los promotores son promotores constitutivos.
26. Célula huésped según la reivindicación 25 donde los promotores constitutivos se seleccionan de entre cualquiera de la lista que consiste en: *ADH1*, *PGK1*, *TEF1*, *TPI1*, *HXT7*, *TDH3* y *PYK1*.
27. Célula huésped según cualquiera de las reivindicaciones 19 a 26 donde la proteína heteróloga es la proteína de SEQ ID NO: 32, y el péptido funcional es el péptido de SEQ ID NO: 12.
28. Célula huésped según cualquiera de las reivindicaciones 19 a 27 donde la primera proteína de fusión es la proteína de SEQ ID NO: 49.
29. Célula huésped según cualquiera de las reivindicaciones 19 a 28 donde la célula se selecciona del grupo que consiste en: célula de levadura, célula bacteriana y célula de mamífero.
30. Célula huésped según la reivindicación 29 donde la célula de levadura se selecciona del grupo que consiste en: *Yarrowia lipolytica*, *Pichia pastoris* y

Saccharomyces cerevisiae, preferentemente la célula de levadura pertenece a la especie *S. cerevisiae*.

31. Célula huésped según cualquiera de las reivindicaciones 19 a 30 donde la célula es la célula de levadura OY216.

5 32. Construcción génica que comprende las secuencias nucleotídicas que codifican para:

i) un promotor

ii) un dominio de unión al ADN, y

10 iii) la proteína heteróloga capaz de unirse a un péptido funcional localizado en el extremo C-terminal de la proteína de referencia.

33. Construcción génica según la reivindicación 26 donde el promotor es un promotor constitutivo, preferentemente el promotor de *ADH1*, el dominio de unión al ADN es preferentemente el dominio de unión de la proteína LexA, y la proteína heteróloga es preferentemente la proteína humana TSG101 que comprende la SEQ ID NO: 32.

15

34. Construcción génica según cualquiera de las reivindicaciones 32 a 33 caracterizada por que es un plásmido, preferentemente el plásmido pRS402-LexA-Tsg101 que comprende la secuencia SEQ ID NO: 30.

20 35. Construcción génica que comprende las secuencias nucleotídicas que codifican para:

i) un promotor

ii) un dominio de transactivación y

25 iii) una proteína de referencia que en su extremo carboxilo terminal comprende la secuencia que codifica para el péptido funcional 3xPTAP de SEQ ID NO: 12.

36. Construcción génica según la reivindicación 35 donde el promotor es un promotor constitutivo, preferentemente el promotor de *ADH1*, el dominio de transactivación es preferentemente el dominio de transactivación del activador transcripcional Gal4, y la proteína de referencia que comprende en su extremo

carboxilo terminal la secuencia que codifica para el péptido funcional 3xPTAP de SEQ ID NO: 12.

5 37. Construcción génica según las reivindicaciones 35 a 36 caracterizada por que es un plásmido, preferentemente el plásmido pACT2-GK-3xPTAP que comprende la secuencia SEQ ID NO: 14.

38. Construcción génica que comprende las secuencias nucleotídicas que codifican para:

i) un promotor

ii) un dominio de unión al ADN, y

10 iii) una proteína diana a la que se une la proteína de referencia.

39. Construcción génica según la reivindicación 38 donde el promotor es un promotor constitutivo, preferentemente el promotor es el promotor de *ADH1*, el dominio de unión al ADN es preferentemente el del activador transcripcional Gal4, y la proteína diana.

15 40. Construcción génica según cualquiera de las reivindicaciones 38 a 39 caracterizada por que es un plásmido, preferentemente el plásmido PGBKT7-GKRP que comprende la secuencia SEQ ID NO: 6.

Fig 1a

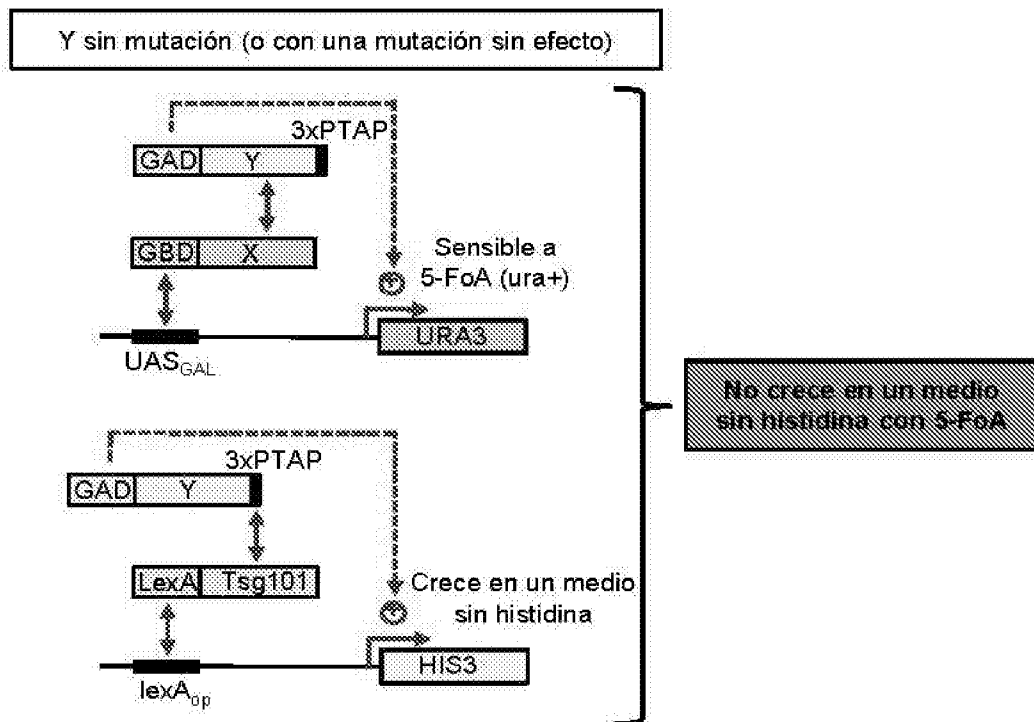


Fig 1b

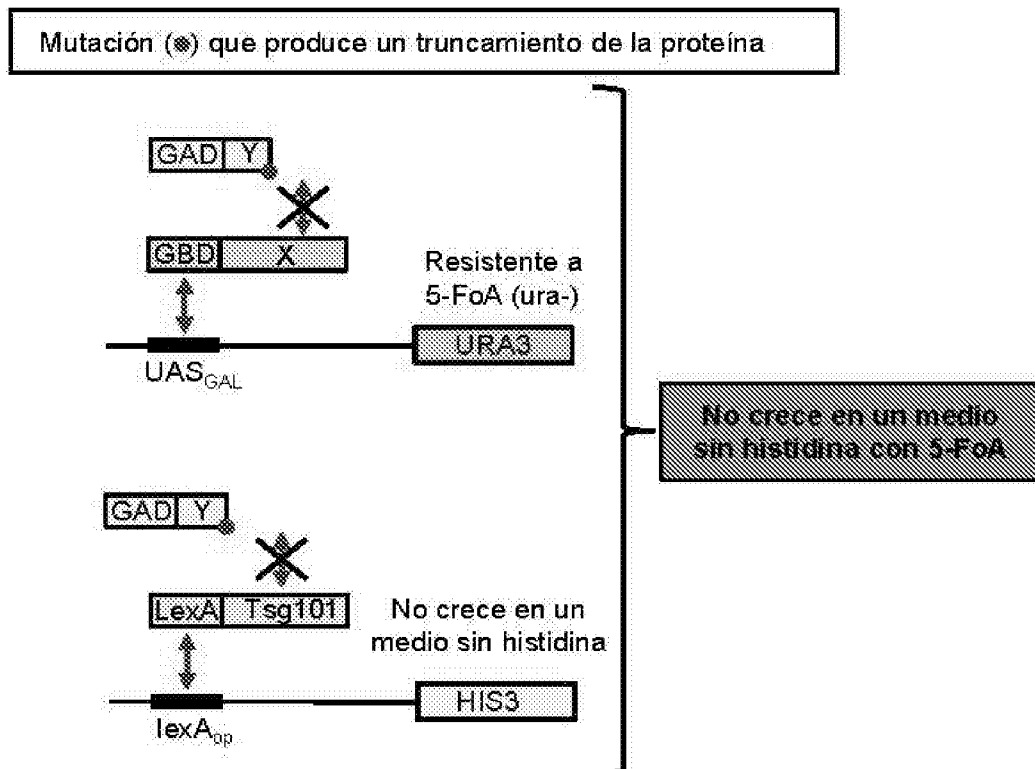


Fig 1c

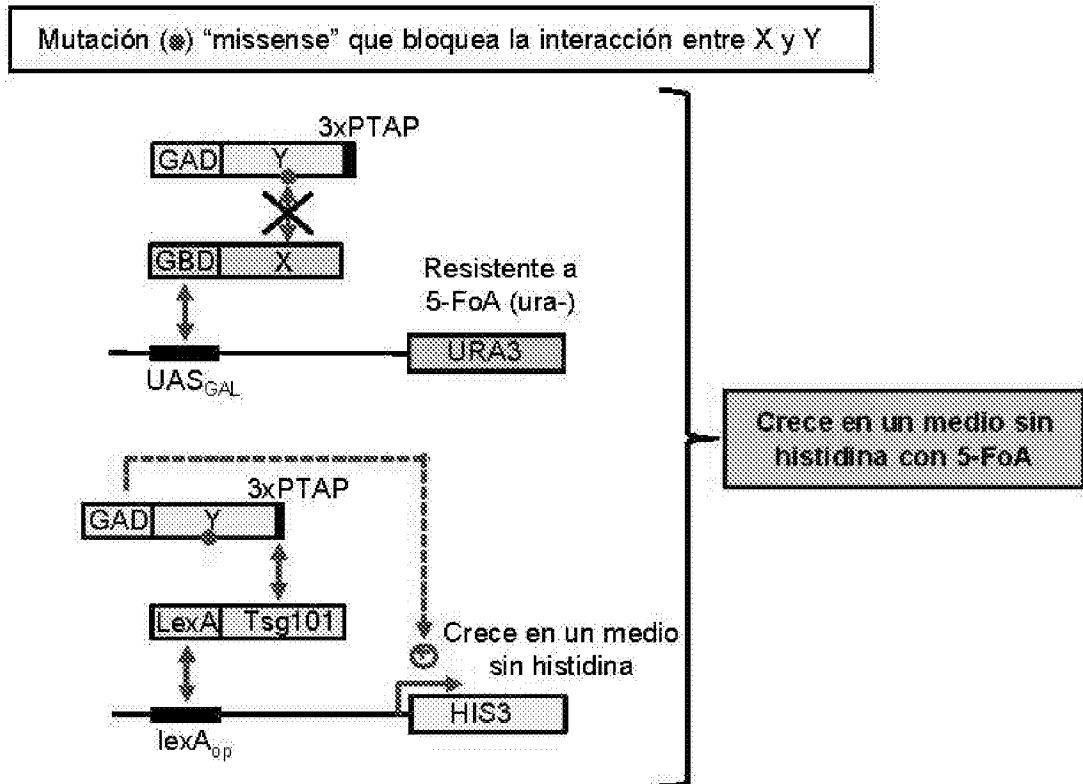


Fig 2a

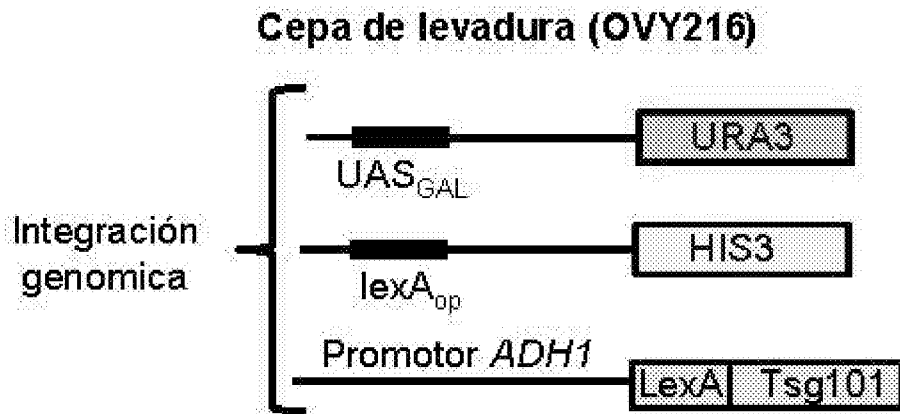


Fig 2b

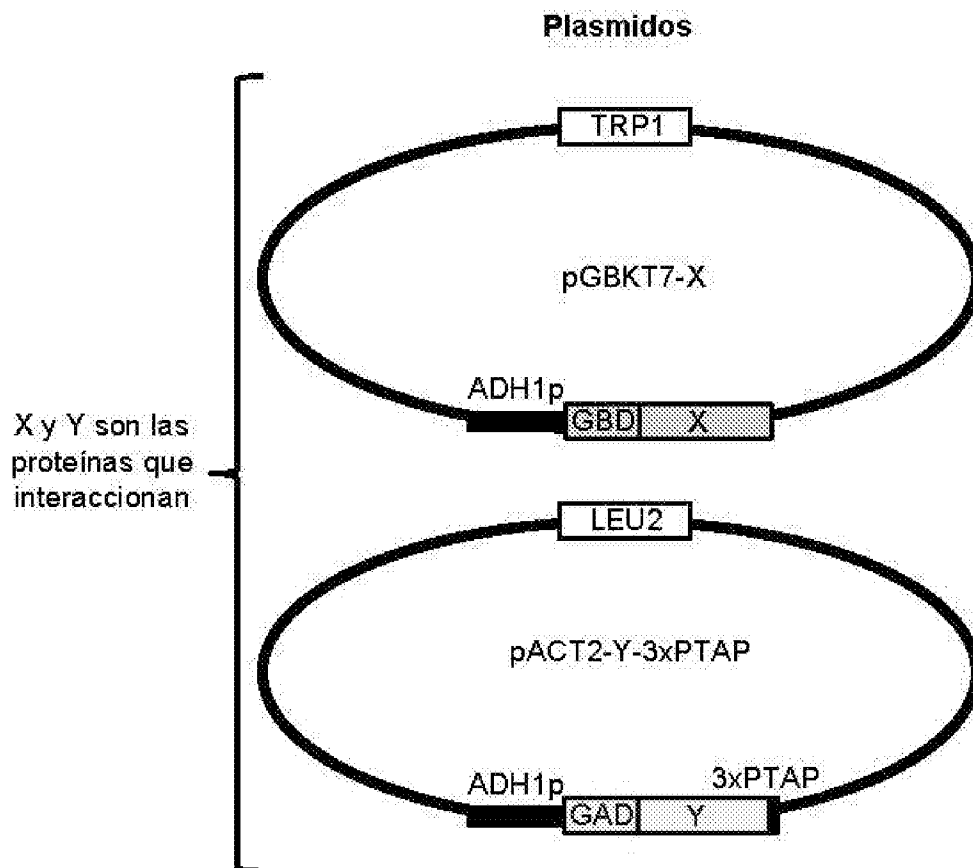


Fig 3

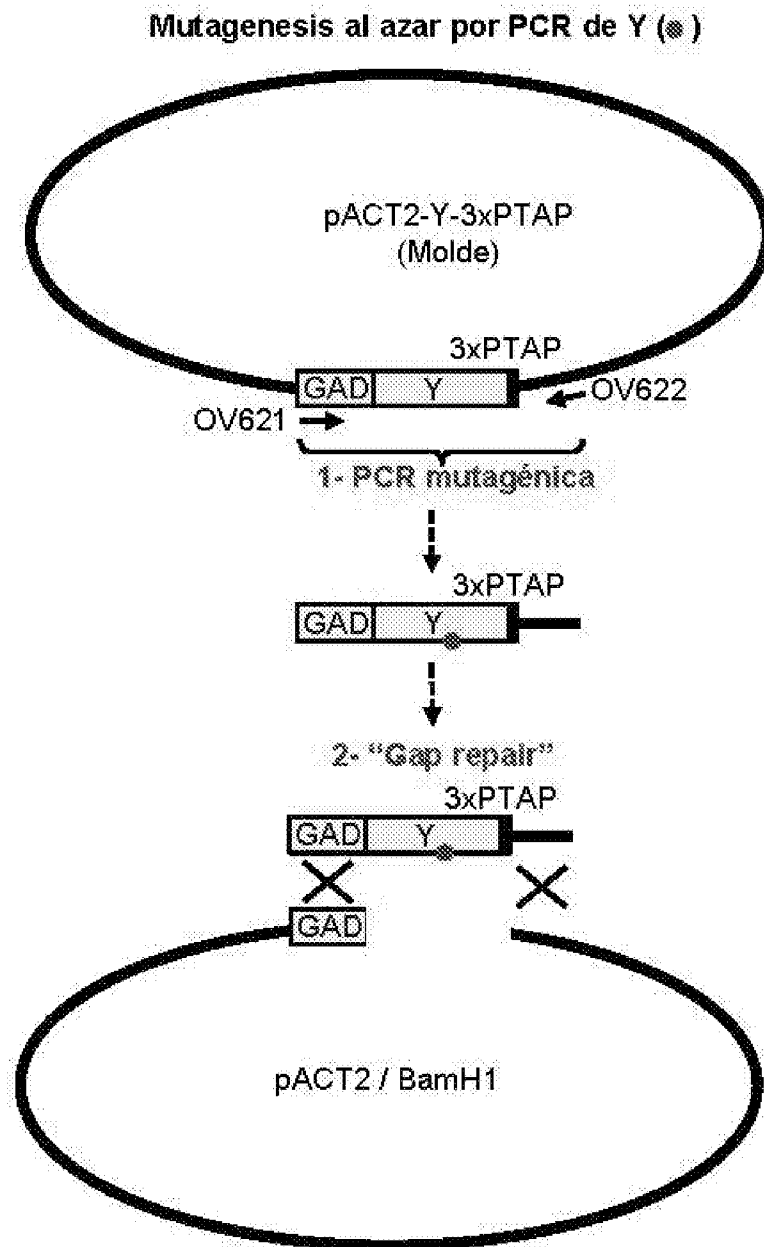
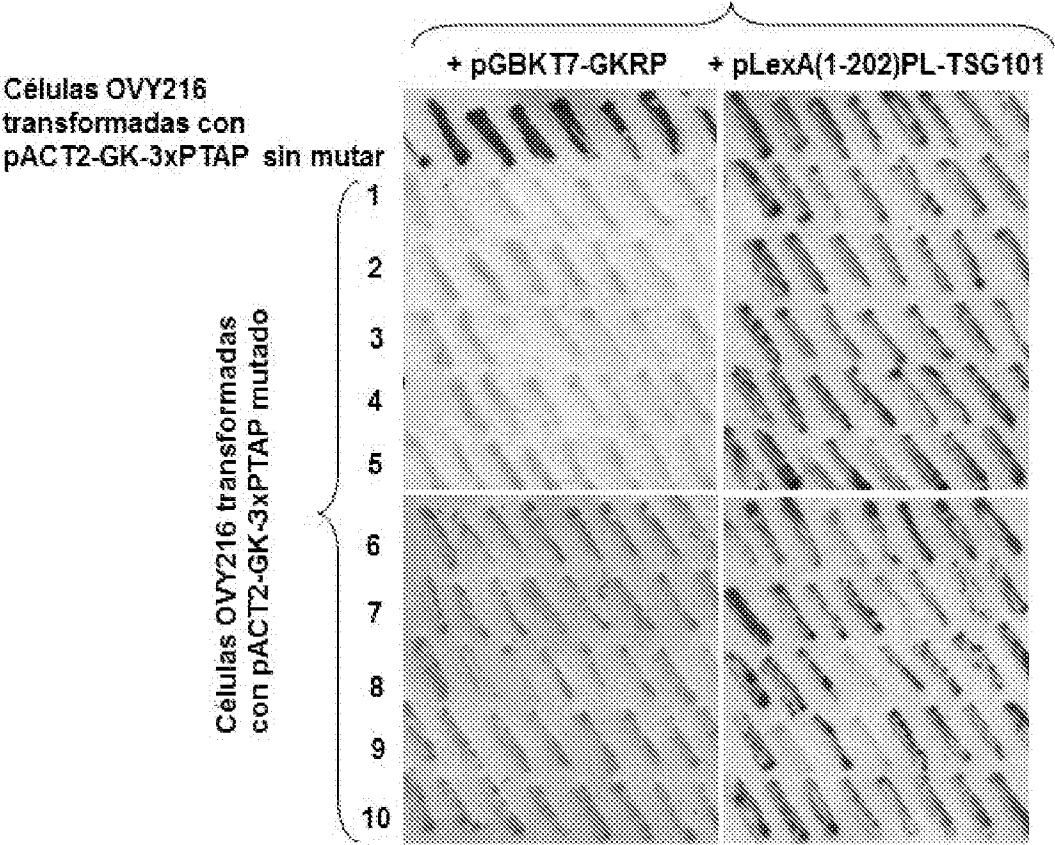


Fig 4



INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2017/070298

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q, C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPODOC, INVENES, WPI, BIOSIS, MEDLINE, EMBASE, EMBL-EBI, INTERNET

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WHITE, M.A. The yeast two-hybrid system: forward and reverse. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 1996, Vol. 93, N° 19, Pages 10001-10003, ISSN: 0027-8424 the whole document, in particular, page 10002, column 1; figure 1.	38-40
X	WO 9723609 A1 (AMERSHAM INT PLC) 03/07/1997, page 8, line 16 - page 11, line 17; page 30, line 28 - page 31, line 13; examples 2 and 3, claims 1-13.	38-40

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance.	
"E" earlier document but published on or after the international filing date	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"O" document referring to an oral disclosure use, exhibition, or other means.	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other documents, such combination being obvious to a person skilled in the art
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search
19/07/2017Date of mailing of the international search report
(28/07/2017)

Name and mailing address of the ISA/

OFICINA ESPAÑOLA DE PATENTES Y MARCAS
Paseo de la Castellana, 75 - 28071 Madrid (España)
Facsimile No.: 91 349 53 04Authorized officer
J. Vizán Arroyo

Telephone No. 91 3498573

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2017/070298

C (continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of documents, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	INOUE, C. et al. Mutational analysis of STE5 in the yeast <i>Saccharomyces cerevisiae</i> : application of a differential interaction trap assay for examining protein-protein interactions. GENETICS, 1997, Vol. 147, N° 2, Pages 479-492, ISSN: 0016-6731 -6731(Print). Results, figure 1; Discussion.	38-40
X	PUTHALAKATH, H. et al. Rapid selection against truncation mutants in yeast reverse two-hybrid screens. BIOTECHNIQUES, 2001, Vol. 30, N° 5, Pages 984-988, ISSN: 0736-6205 (print) Materials and methods; Results.	38-40
X	US 2005053913 A1 (JOHNS HOPKINS UNIVERSITY A MAR) 10/03/2005, figure 4, claims.	38-40
A	KIM, J.Y. et al. One- plus two-hybrid system, a novel yeast genetic selection for specific missense mutations disrupting protein/protein interactions. MOLECULAR & CELLULAR PROTEOMICS, 2007, Vol. 6, N° 10, Pages 1727-1740, ISSN: 1535-9476 the whole document.	1-40
A	SHIH, H.M. et al. A positive genetic selection for disrupting protein-protein interactions: identification of CREB mutations that prevent association with the coactivator CBP. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 1996, Vol. 93, N° 24, Pages 13896-13901, ISSN: 0027-8424 the whole the document.	1-40
A	WO 02072790 A2 (MYRIAD GENETICS INC) 19/09/2002, the whole document.	1-40
A	GALÁN, M. et al. Effects of novel maturity-onset diabetes of the young (MODY)-associated mutations on glucokinase activity and protein stability. BIOCHEMICAL JOURNAL, 2006, Vol. 393, N° Pt 1, Pages 389-396, ISSN: 0264-6021 the whole document.	1-40

INTERNATIONAL SEARCH REPORT

International application No.

Information on patent family members

PCT/ES2017/070298

Patent document cited in the search report	Publication date	Patent family member(s)	Publication date
WO9723609 A1	03.07.1997	US6074829 A EP0868509 A1 EP0790304 A1 AU1165697 A	13.06.2000 07.10.1998 20.08.1997 17.07.1997
-----	-----	-----	-----
US2005053913 A1	10.03.2005	US5965368 A US5955280 A US2006223089 A1 US7601533 B2 US2003068612 A1 US2002197627 A1 WO9632503 A1 US2003124538 A1 US7033768 B2 NZ513455 A NZ502795 A NZ306767 A JP2007075107 A JPH11502717 A EP0830459 A1 EP0830459 A4 CA2217545 A1 AU5541296 A AU716893B B2 AT352638T T	12.10.1999 21.09.1999 05.10.2006 13.10.2009 10.04.2003 26.12.2002 17.10.1996 03.07.2003 25.04.2006 28.03.2003 28.09.2001 27.03.2000 29.03.2007 09.03.1999 25.03.1998 27.02.2002 17.10.1996 30.10.1996 09.03.2000 15.02.2007
-----	-----	-----	-----
WO02072790 A2	19.09.2002	JP2009269918 A ES2328796T T3 AT434936T T US2009035853 A1 US2002173622 A1 US7202329 B2 US2004109861 A1 US7335468 B2 US2002177207 A1 US2005222029 A1 JP2004533815 A CA2441071 A1 CA2441071 C AU2002257067 A1 EP1377165 A2 EP1377165 A4	19.11.2009 18.11.2009 15.07.2009 05.02.2009 21.11.2002 10.04.2007 10.06.2004 26.02.2008 28.11.2002 06.10.2005 11.11.2004 19.09.2002 22.06.2010 24.09.2002 07.01.2004 20.04.2005
-----	-----	-----	-----

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2017/070298

CLASSIFICATION OF SUBJECT MATTER

C12Q1/68 (2006.01)

C12N15/10 (2006.01)

C12N15/62 (2006.01)

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional nº

PCT/ES2017/070298

A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

Ver Hoja Adicional

De acuerdo con la Clasificación Internacional de Patentes (CIP) o según la clasificación nacional y CIP.

B. SECTORES COMPRENDIDOS POR LA BÚSQUEDA

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12Q, C12N

Otra documentación consultada, además de la documentación mínima, en la medida en que tales documentos formen parte de los sectores comprendidos por la búsqueda

Bases de datos electrónicas consultadas durante la búsqueda internacional (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

EPODOC, INVENES, WPI, BIOSIS, MEDLINE, EMBASE, EMBL-EBI, INTERNET

C. DOCUMENTOS CONSIDERADOS RELEVANTES

Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones nº
X	WHITE, M.A. The yeast two-hybrid system: forward and reverse. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 1996, Vol. 93, Nº 19, páginas 10001-10003, ISSN: 0027-8424 todo el documento, en particular, página 10002, columna 1; figura 1.	38-40
X	WO 9723609 A1 (AMERSHAM INT PLC) 03/07/1997, página 8, línea 16 - página 11, línea 17; página 30, línea 28 – página 31, línea 13; ejemplos 2 y 3, reivindicaciones 1-13.	38-40

☒ En la continuación del recuadro C se relacionan otros documentos

☒ Los documentos de familias de patentes se indican en el anexo

* Categorías especiales de documentos citados:	"T" documento ulterior publicado con posterioridad a la fecha de presentación internacional o de prioridad que no pertenece al estado de la técnica pertinente pero que se cita por permitir la comprensión del principio o teoría que constituye la base de la invención.
"A" documento que define el estado general de la técnica no considerado como particularmente relevante.	"X" documento particularmente relevante; la invención reivindicada no puede considerarse nueva o que implique una actividad inventiva por referencia al documento aisladamente considerado.
"E" solicitud de patente o patente anterior pero publicada en la fecha de presentación internacional o en fecha posterior.	"Y" documento particularmente relevante; la invención reivindicada no puede considerarse que implique una actividad inventiva cuando el documento se asocia a otro u otros documentos de la misma naturaleza, cuya combinación resulta evidente para un experto en la materia.
"L" documento que puede plantear dudas sobre una reivindicación de prioridad o que se cita para determinar la fecha de publicación de otra cita o por una razón especial (como la indicada).	"&" documento que forma parte de la misma familia de patentes.
"O" documento que se refiere a una divulgación oral, a una utilización, a una exposición o a cualquier otro medio.	
"P" documento publicado antes de la fecha de presentación internacional pero con posterioridad a la fecha de prioridad reivindicada.	

Fecha en que se ha concluido efectivamente la búsqueda internacional.
19/07/2017

Fecha de expedición del informe de búsqueda internacional.
28 de julio de 2017 (28/07/2017)

Nombre y dirección postal de la Administración encargada de la búsqueda internacional
OFICINA ESPAÑOLA DE PATENTES Y MARCAS
Paseo de la Castellana, 75 - 28071 Madrid (España)
Nº de fax: 91 349 53 04

Funcionario autorizado
J. Vizán Arroyo
Nº de teléfono 91 3498573

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional n°

PCT/ES2017/070298

C (Continuación).		DOCUMENTOS CONSIDERADOS RELEVANTES
Categoría *	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones n°
X	INOUE, C. et al. Mutational analysis of STE5 in the yeast <i>Saccharomyces cerevisiae</i> : application of a differential interaction trap assay for examining protein-protein interactions. GENETICS, 1997, Vol. 147, N° 2, Páginas 479-492, , ISSN: 0016-6731 -6731(Print). Resultados, Figura 1; Discusión.	38-40
X	PUTHALAKATH, H. et al.,. Rapid selection against truncation mutants in yeast reverse two-hybrid screens. BIOTECHNIQUES, 2001, Vol. 30, N° 5, Páginas 984-988, ISSN: 0736-6205 (print) Materiales y métodos; Resultados.	38-40
X	US 2005053913 A1 (JOHNS HOPKINS UNIVERSITY A MAR) 10/03/2005, Figura 4, Reivindicaciones.	38-40
A	KIM, J.Y. et al. One- plus two-hybrid system, a novel yeast genetic selection for specific missense mutations disrupting protein/protein interactions. MOLECULAR & CELLULAR PROTEOMICS, 2007, Vol. 6, N° 10, Páginas 1727-1740, ISSN: 1535-9476 todo el documento.	1-40
A	SHIH, H.M. et al. A positive genetic selection for disrupting protein-protein interactions: identification of CREB mutations that prevent association with the coactivator CBP. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 1996, Vol. 93, N° 24, Páginas 13896-13901, ISSN: 0027-8424 todo el documento.	1-40
A	WO 02072790 A2 (MYRIAD GENETICS INC) 19/09/2002, todo el documento.	1-40
A	GALÁN, M. et al. Effects of novel maturity-onset diabetes of the young (MODY)-associated mutations on glucokinase activity and protein stability. BIOCHEMICAL JOURNAL, 2006, Vol. 393, N° Pt 1, Páginas 389-396, ISSN: 0264-6021 todo el documento.	1-40

INFORME DE BÚSQUEDA INTERNACIONAL

Informaciones relativas a los miembros de familias de patentes

Solicitud internacional nº

PCT/ES2017/070298

Documento de patente citado en el informe de búsqueda	Fecha de Publicación	Miembro(s) de la familia de patentes	Fecha de Publicación
WO9723609 A1	03.07.1997	US6074829 A EP0868509 A1 EP0790304 A1 AU1165697 A	13.06.2000 07.10.1998 20.08.1997 17.07.1997
-----	-----	-----	-----
US2005053913 A1	10.03.2005	US5965368 A US5955280 A US2006223089 A1 US7601533 B2 US2003068612 A1 US2002197627 A1 WO9632503 A1 US2003124538 A1 US7033768 B2 NZ513455 A NZ502795 A NZ306767 A JP2007075107 A JPH11502717 A EP0830459 A1 EP0830459 A4 CA2217545 A1 AU5541296 A AU716893B B2 AT352638T T	12.10.1999 21.09.1999 05.10.2006 13.10.2009 10.04.2003 26.12.2002 17.10.1996 03.07.2003 25.04.2006 28.03.2003 28.09.2001 27.03.2000 29.03.2007 09.03.1999 25.03.1998 27.02.2002 17.10.1996 30.10.1996 09.03.2000 15.02.2007
-----	-----	-----	-----
WO02072790 A2	19.09.2002	JP2009269918 A ES2328796T T3 AT434936T T US2009035853 A1 US2002173622 A1 US7202329 B2 US2004109861 A1 US7335468 B2 US2002177207 A1 US2005222029 A1 JP2004533815 A CA2441071 A1 CA2441071 C AU2002257067 A1 EP1377165 A2 EP1377165 A4	19.11.2009 18.11.2009 15.07.2009 05.02.2009 21.11.2002 10.04.2007 10.06.2004 26.02.2008 28.11.2002 06.10.2005 11.11.2004 19.09.2002 22.06.2010 24.09.2002 07.01.2004 20.04.2005
-----	-----	-----	-----

CLASIFICACIONES DE INVENCION

C12Q1/68 (2006.01)

C12N15/10 (2006.01)

C12N15/62 (2006.01)