

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 374 418**

51 Int. Cl.:  
**C12N 15/90** (2006.01)  
**C12N 15/10** (2006.01)  
**C12N 9/22** (2006.01)  
**C12N 15/82** (2006.01)  
**A61K 48/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07731461 .5**  
96 Fecha de presentación: **15.05.2007**  
97 Número de publicación de la solicitud: **2021485**  
97 Fecha de publicación de la solicitud: **11.02.2009**

54 Título: **PSEUDOTRANSPONES Y SISTEMAS DE TRANSPOSICIÓN RECOMBINANTES HIPERACTIVOS, DERIVADOS DEL TRANSPÓSÓN MOS-1.**

30 Prioridad:  
**15.05.2006 FR 0604285**

45 Fecha de publicación de la mención BOPI:  
**16.02.2012**

45 Fecha de la publicación del folleto de la patente:  
**16.02.2012**

73 Titular/es:  
**CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)  
3, RUE MICHEL ANGE  
75016 PARIS, FR y  
UNIVERSITE FRANCOIS RABELAIS DE TOURS**

72 Inventor/es:  
**BIGOT, Sylvie, Marie-Louise;  
BIGOT, Yves, Philippe, Marcel;  
BONNIN-ROULEUX, Florence;  
GERMON, Stéphanie, Odile, Rébecca y  
JEGOT, Gwenhael**

74 Agente: **No consta**

ES 2 374 418 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

**DESCRIPCIÓN**

La presente invención se refiere al campo de la biología molecular relativa a los elementos transposables. La invención se refiere más particularmente a la mejora de las propiedades de los sistemas de transposición MOS-1 para su utilización en biotecnología.

- 5 La presente invención tiene por objeto un pseudo-transposón *Mos-1* hiperactivo, en el que:
- a) por lo menos una de las dos repeticiones terminales no traducidas (UTR) y/o por lo menos una de las dos repeticiones terminales inversas (ITR) está(n) genéticamente modificada(s); y
  - b) una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa MOS-1 de origen,

10 caracterizado porque se selecciona de entre los pseudo-transposones siguientes:

- α) ITR3'-UTR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3'- (pseudo-transposón 33sec33),
- β) ITR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3' (pseudo-transposón 3sec33),
- γ) ITR3'-UTR5'-secuencia nucleotídica exógena de interés-UTR3'-ITR5' (pseudo-transposón 35sec35),
- δ) los pseudo-transposones que comprenden por lo menos una ITR40 de secuencia SEC ID nº 39, y
- 15 ε) los pseudo-transposones que comprenden por lo menos una ITR46 de secuencia SEC ID nº 38, y en los que ITR3', UTR3', UTR5' e ITR5' designan respectivamente la secuencia de la repetición terminal inversa 3' del transposón *Mos-1*, la secuencia de la repetición terminal no traducida 3' del transposón *Mos-1*, la secuencia de la repetición terminal no traducida 5' del transposón *Mos-1* y la secuencia de la repetición terminal inversa 5' del transposón *Mos-1*.

20 La presente invención tiene además por objeto un sistema de transposición recombinante hiperactivo del transposón *Mos-1*, que comprende por lo menos las dos parejas siguientes:

- a) un pseudo-transposón *Mos-1* hiperactivo según la reivindicación 1; y
- b) una transposasa *Mos-1* suministrada en *trans* de dicho pseudo-transposón, en el que la frecuencia de transposición de dicha secuencia nucleotídica exógena de interés está mejorada en un factor por lo menos igual a 5, preferentemente por lo menos igual a 10, con respecto a la frecuencia de transposición de dicha secuencia nucleotídica exógena de interés contenida por un pseudo-transposón *Mos-1* rodeado por dos ITR3' (3sec3) en presencia de la transposasa *Mos-1* suministrada en *trans*.

Además de dichos sistemas, la invención proporciona unas células hospedantes, unos kits y unas utilizaciones tales como se especifican en las reivindicaciones.

30 La presente invención se refiere asimismo a la utilización de uno o varios de los medios anteriores para realizar unas transposiciones de secuencias, y más particularmente unas transferencias de genes, eficaces en una célula eucariota, con la excepción de una célula en un organismo humano o animal.

35 Los elementos transposables (TE) o elementos genéticos móviles (EGM) son unos fragmentos de ADN de pequeño tamaño, capaces de desplazarse de un sitio cromosómico a otro (Renault *et al.*, 1997). Estos fragmentos de ADN se caracterizan por una secuencia repetida inversa (ITR) situadas en las posiciones 5' y 3' terminales. Una enzima codificada por los propios TE, la transposasa, cataliza el proceso de transposición de estos últimos.

Se han identificado unos TE tanto en los procariotas como en los eucariotas (véase una obra de referencia en este campo: Craig *et al.*, 2002).

40 Los TE se reparten en dos clases según sus mecanismos de transposición. Por un lado, los elementos de clase I, o retrotransposones, transponen a través de la transcripción inversa de un intermedio ARN. Por otro lado, los elementos de clase II transponen directamente de un sitio cromosómico a otro, a través de un intermedio ADN, según un mecanismo de tipo "copiar-pegar".

En los procariotas, un gran número de TE han sido clasificados en la actualidad. Se pueden citar, por ejemplo, unas secuencias de inserción tales como IS1, y unos transposones, tal como Tn5.

45 En los eucariotas, los elementos de clase II comprenden diez familias, P, PiggyBac, hAT, hélitron, Harbinger, En/Spm, Mutator, Transib, Pogo y Tc1-mariner.

Los elementos genéticos móviles *mariner* (o MLE por "*mariner-like elements*") constituyen un gran grupo de TE de clase II, que pertenecen a la superfamilia Tc1-mariner (Plasterk *et al.*, 1999).

La capacidad de las transposasas de TE para movilizar unos fragmentos de ADN más o menos largos, homólogos o heterólogos, que comprenden unas secuencias de interés, para insertarlos en unos ácidos nucleicos diana, en particular en el cromosoma de un hospedante, ha sido y sigue siendo ampliamente explotada en el campo de las biotecnologías, en particular en el campo de la ingeniería genética.

5 Entre los TE, los MLE presentan unas propiedades particularmente ventajosas para una utilización en biotecnología, en particular en ingeniería genética y genómica funcional. Se pueden citar, por ejemplo, de manera no limitativa las propiedades siguientes:

- (i) los MLE son unos transposones de pequeño tamaño, fáciles de manipular.
- 10 (ii) El mecanismo de transposición de los MLE es sencillo. En efecto, la transposasa es capaz de catalizar, por sí misma, todas las etapas de la transposición de los MLE. De hecho, es necesaria y suficiente para asegurar la movilidad de los MLE en ausencia de factores del hospedante (Lampe *et al.*, 1996).
- 15 (iii) Los MLE se caracterizan por una ubiquidad muy amplia en los procariontes y en los eucariotes. El primer MLE, *Dmmar1*, denominado asimismo *Mos-1*, ha sido descubierto en *Drosophila mauritiana* por Jacobson y Harti (1985). A continuación, se han identificado numerosos elementos emparentados en unos genomas que pertenecen en particular a bacterias, protozoarios, hongos, plantas, invertebrados, vertebrados de sangre fría y mamíferos.
- 20 (iv) La actividad transcripcional de los MLE es altamente específica, y no induce ningún mecanismo de "resistencia" del genoma del hospedante, tales como unos fenómenos de interferencias por metilación [MIP; Jeong *et al.* (2002); Martienssen y Colot (2001)] o a través de unos ARN [RNAi; Ketting *et al.* (1999); Tabara *et al.* (1999)]. Los eventos de transposición pueden ser controlados mediante diversos factores, tales como la temperatura, la presencia de ciertos cationes divalentes y el pH.

En términos de estructura, el elemento *Mos-1 de mariner* es un elemento compacto de 1286 pb que contiene un solo marco abierto de lectura (ORF) que codifica para una transposasa de 354 aminoácidos. La transposasa está constituida por un dominio N-terminal implicado en la unión al ADN, la dimerización y la tetramerización y por un dominio C-terminal que contiene el sitio activo en el que un motivo DDE(D) coordina el ión metálico catalítico. El marco de lectura está flanqueado por dos secuencias terminales inversas (ITR) de 28±2 pb. Las dos regiones localizadas entre las ITR y el ORF no están traducidas y se denominan UTR (por "Untranslated Terminal Region"). Las dos ITR de *Mos-1* difieren en secuencia por cuatro nucleótidos, lo cual deja pensar que la configuración natural de este elemento no es óptima para la transposición. Esto ha sido verificado de hecho por unos experimentos que muestran que un pseudo-transposón rodeado por dos ITR 3' de *Mos-1* transpone 10.000 veces mejor *in vivo* en bacteria que el rodeado por las ITR 5' y 3' en configuración natural (Augé-Gouillou *et al.*, 2001b).

Las aplicaciones potenciales de los MLE en biotecnología, en particular como herramientas no víricas de recombinación genética, son considerables.

35 Típicamente, para unas aplicaciones de mutagénesis de inserción *in vitro*, el gen del transposón que codifica para la transposasa se sustituye por un ADN "etiqueta". La transposasa se suministra en *trans* en forma proteica. Para unas aplicaciones de transferencia de genes *in vivo* o *in vitro*, el gen que codifica la transposasa se sustituye por el ADN exógeno a transferir (se obtiene así un "pseudo-transposón"). La transposasa se suministra entonces en *trans* a través de un plásmido de expresión, un ARN mensajero o la propia proteína. Sin embargo, la transferencia de un ADN exógeno con la ayuda de los medios actuales, es decir utilizando un pseudo-transposón *mariner*, no deja de plantear dificultades, 40 en particular debido a una eficacia y a una especificidad de integración del transgén muy limitadas.

El documento de Pledger D y Coates C (2005) en *Insect Biochemistry and Molecular Biology* vol. 35 p. 1199-1207 describe el aumento de la actividad de transposición de un pseudo-transposón *Mos-1* cuando se utiliza una transposasa *Mos-1* mutada hiperactiva en combinación con un pseudo-transposón hiperactivo 3T3 (según la terminología de la presente solicitud).

45 El documento de Augé-Gouillou C *et al* (2001b) en *Molecular Genetics and Genomics* vol. 265 p.51-57 describe que la afinidad de la transposasa *Mos-1* para la ITR3' es superior a aquella para ITR5', y que la actividad de transposición es más elevada en un pseudo-transposón que comprende dos ITR3' (3T3) que en un pseudo-transposón que comprende una ITR5' y una ITR3'.

50 El documento de Pledger D. *et al* en *Molecular Genetics and Genomics* vol. 272 p. 67-75 describe el aumento de actividad de transposición en *E. coli* cuando las dos ITR son de tipo 3', y describe un aumento suplementario cuando se elimina la secuencia interna de *Mos-1*, salvo 10 pares de bases.

Sin embargo, es esencial disponer, en cada una de estas aplicaciones, de sistemas de transposición eficaces.

55 Ahora bien, la eficacia de transposición de los MLE naturales, comparada con la de los demás TE, sigue siendo baja. En particular, los MLE parecen menos activos en los eucariotes que los demás EGM de clase II. En definitiva, el interés práctico de los MLE naturales ha permanecido hasta ahora limitado, puesto que es importante, para el industrial

o el investigador, disponer de herramientas de transposición eficaces, de manera que se reduzca el número de manipulaciones, el coste y el tiempo necesarios para realizar las transposiciones buscadas. A falta de eficacia suficiente de los sistemas de transposición disponibles, el industrial o el investigador es propenso en la actualidad a privilegiar, en la medida de lo posible, la utilización de sistemas víricos de transferencia de genes, a pesar de los inconvenientes que presentan, en detrimento de los sistemas de transposición recombinantes construidos a partir de los transposones MLE.

Por lo tanto, existe en la actualidad una necesidad para un sistema que (i) permita transferir eficazmente unos genes; (ii) presente una buena inocuidad en términos de inmunogenicidad para el hospedante; (iii) garantice la seguridad del hospedante y del medioambiente (ausencia de contaminaciones, en particular ausencia de emergencia de virus recombinantes); (iv) sea fácil de producir.

La presente invención permite precisamente responder a esta necesidad proporcionando un sistema de transposición recombinante que aprovecha las ventajas del elemento *Mos-1* (ubiquidad, actividad transposicional, simplicidad de producción y de utilización, etc.), evitando al mismo tiempo los inconvenientes relacionados con la baja eficacia de transposición de este elemento en el estado salvaje.

Así, con el fin de responder de manera satisfactoria a la necesidad existente, los inventores se han interesado en el elemento MLE *Mos-1* de *Drosophila mauritiana* que es el elemento más caracterizado de esta familia y el único que es naturalmente activo. Además, el elemento *Mos-1* presenta la ventaja considerable de ser activo al mismo tiempo en célula eucariota y en bacteria, lo cual evidencia su interés en el marco de la puesta a punto de un sistema de transposición ubicuo eficaz.

Tal como se desprende de los diferentes aspectos de la presente invención descritos con mayor detalle a continuación, los inventores proponen varias herramientas que mejoran el sistema de transposición *Mos-1* natural. Estas herramientas, que pueden ser utilizadas solas o en combinación en función de las aplicaciones previstas, comprenden:

unos pseudo-transposones *Mos-1* recombinantes hiperactivos.

Se define en este caso un "pseudo-transposón" como un transposón en el que el gen que codifica para la transposasa ha sido sustituido con una secuencia nucleotídica exógena. Un pseudo-transposón posee por lo tanto unos extremos ITR y UTR pero está desprovisto de la actividad transposasa. Por consiguiente, ha perdido la capacidad para transponer, salvo que sea asociado a una transposasa externa.

Así, los inventores se han interesado en unos sistemas de transposición recombinantes hiperactivos derivados del transposón *Mos-1*, que comprende por lo menos las dos parejas siguientes:

- a) un pseudo-transposón *Mos-1* en el que una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa *Mos-1* de origen; y
- b) una transposasa *Mos-1* suministrada en *trans* de dicho pseudo-transposón, estando una por lo menos de dichas parejas genéticamente modificada de manera apropiada para mejorar la frecuencia de transposición de dicha secuencia nucleotídica exógena de interés en un factor por lo menos igual a 5, preferentemente por lo menos igual a 10.

En los sistemas de transposición recombinantes propuestos, las dos parejas, a saber el pseudo-transposón y la transposasa, pueden estar genéticamente modificadas. Así, una u otra de las parejas o las dos puede(n) estar genéticamente modificada(s) y ser hiperactiva(s).

Una "secuencia nucleotídica" o un "ácido nucleico" según la invención está de acuerdo con la acepción habitual en el campo de la biología. Estas dos expresiones abarcan indiferentemente los ADN y los ARN, pudiendo ser los primeros por ejemplo genómicos, plasmídicos, recombinantes, complementarios (ADNc), y los segundos, mensajeros (ARNm), ribosómicos (ARNr), de transferencia (ARNt). De manera preferida, las secuencias nucleotídicas y los ácidos nucleicos de la invención son unos ADN.

Los términos y las expresiones "actividad", "función", "actividad biológica" y "función biológica" son equivalentes y responden a la acepción habitual en el campo técnico de la invención. En el marco de la invención, la actividad en cuestión es la actividad de transposición.

La "hiperactividad" corresponde en este caso a una actividad de transposición superior a la observada utilizando un sistema de transposición *Mos-1* natural que comprende:

- a) un pseudo-transposón *Mos-1* en el que una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa *Mos-1* de origen; y
- b) la transposasa *Mos-1* salvaje, suministrada en *trans* de dicho pseudo-transposón.

Además, la "hiperactividad" en el sentido de la presente invención designa una actividad de transposición superior a la observada utilizando un sistema de transposición *Mos-1* recombinante que comprende:

- a) un pseudo-transposón *Mos-1* en el que
- (i) una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa *Mos-1* de origen, y
  - (ii) la repetición terminal inversa salvaje situada en 5' (ITR5') ha sido mutada de manera que es una copia perfecta de la repetición terminal inversa salvaje situada en 3' (ITR3'); y
- b) la transposasa *Mos-1* salvaje, suministrada en *trans* de dicho pseudo-transposón.

Este pseudo-transposón está por lo tanto rodeado por 2 ITR3' ("pseudo-transposón 2 ITR3'" o "pseudo-transposón 3T3"). La presencia de 2 ITR3' ya ha sido descrita como que permite mejorar la eficacia de transposición con respecto al transposón *Mos-1* natural (Augé-Gouillou *et al.*, 2001b), y el sistema recombinante anterior se utiliza por lo tanto en este caso como referencia para determinar si un sistema de transposición recombinante es "hiperactivo" en el sentido de la presente invención. El sistema recombinante de referencia descrito anteriormente se designa a continuación como el "sistema (de transposición) de referencia" o "sistema (de transposición) 3T3.

La "eficacia de transposición" se determina según la frecuencia de transposición. La eficacia de transposición está por lo tanto mejorada si la frecuencia de la transposición ha aumentado. Se hablará en la continuación de la descripción indiferentemente de "mejora de la actividad (de transposición)", de "mejora de la transposición" o también de "mejora de la eficacia (de transposición)", refiriéndose todas estas expresiones a la "mejora de la frecuencia de transposición", es decir a su aumento.

Con el fin de cuantificar la "hiperactividad", se utiliza en este caso un "factor" o un "factor de hiperactividad" que es igual a la relación de las frecuencias de transposición según la fórmula siguiente:

Factor de hiperactividad (F)= (frecuencia de transposición observada con un sistema de transposición recombinante dado)/(frecuencia de transposición observada con el sistema de transposición de referencia 3T3 definido anteriormente).

En otras palabras, se compara la frecuencia de transposición de una secuencia nucleotídica exógena contenida por un pseudo-transposón *Mos-1* en presencia de la transposasa *Mos-1* suministrada en *trans* con la frecuencia de transposición de esta misma secuencia nucleotídica exógena cuando está contenida por el pseudo-transposón *Mos-1* de referencia en presencia de la transposasa *Mos-1* suministrada en *trans* (sistema 3T3). Este método de evaluación de la actividad de transposición es una práctica de las más corrientes en el campo.

Los sistemas de transposición recombinantes interesantes en el marco de la presente invención permiten por lo tanto mejorar la transposición en un factor por lo menos igual a 5. Preferentemente, el factor de hiperactividad es por lo menos igual a 10, y mejor por lo menos igual a 15. De manera aún más preferida, es por lo menos igual a 20, 25, 30, 35, 40, 45, 50, 55, 60 y más.

Tal como se desprende de los ejemplos siguientes, la o las modificaciones genéticas del pseudo-transposón y/o de la transposasa es (son) precisamente seleccionada(s) por su(s) capacidad(es) para provocar una hiperactividad de transposición. Las mutaciones y las combinaciones de mutaciones que hacen la transposasa y/o el pseudo-transposón y/o el sistema de transposición hiperactivos no son ni aleatorias, ni previsibles. Con el fin de seleccionar las mutaciones apropiadas, los inventores han utilizado un ensayo de transposición bien conocido en el campo. Este ensayo ya ha sido descrito en la bibliografía (en particular en Augé-Gouillou *et al.*, 2001b) y forma parte de los conocimientos generales del experto en la materia que podrá, sin embargo, si lo desea, utilizar cualquier otro ensayo adaptado a su disposición.

Un primer aspecto de la presente invención se refiere a un pseudo-transposón *Mos-1* hiperactivo, en el que:

- a) por lo menos una de las dos repeticiones terminales no traducidas (UTR) y/o por lo menos una de las dos repeticiones terminales inversas (ITR) está(n) genéticamente modificada(s); y
- b) una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa *Mos-1* de origen,

caracterizado porque se selecciona de entre los pseudo-transposones siguientes:

- α) ITR3'-UTR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3' (pseudo-transposón 33sec33),
- β) ITR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3' (pseudo-transposón 3sec33),
- γ) ITR3'-UTR5'-secuencia nucleotídica exógena de interés-UTR3'-ITR5' (pseudo-transposón 35sec35),
- δ) los pseudo-transposones que comprenden por lo menos una ITR40 de secuencia SEC ID n° 39, y
- ε) los pseudo-transposones que comprenden por lo menos una ITR46 de secuencia SEC ID n° 38, y en el que ITR3', UTR3', UTR5' e ITR5' designan respectivamente la secuencia de la repetición terminal inversa 3' del transposón *Mos-1*, la secuencia de la repetición terminal no traducida 3' del transposón *Mos-1*, la secuencia de

la repetición terminal no traducida 5' del transposón *Mos-1* y la secuencia de la repetición terminal inversa 5' del transposón *Mos-1*.

Un segundo aspecto de la presente invención se refiere a un sistema de transposición recombinante hiperactivo derivado del transposón *Mos-1*, que comprende por lo menos las dos parejas siguientes:

- 5 a) un pseudo-transposón *Mos-1* hiperactivo según la reivindicación 1; y
- b) una transposasa *Mos-1* suministrada en *trans* de dicho pseudo-transposón, en el que la frecuencia de transposición de dicha secuencia nucleotídica exógena de interés está mejorada en un factor por lo menos igual a 5, preferentemente por lo menos igual a 10, con respecto a la frecuencia de transposición de dicha secuencia nucleotídica exógena de interés contenida por un pseudo-transposón *Mos-1* rodeado por dos ITR3' (3sec3) en presencia de la transposasa *Mos-1* suministrada en *trans*.
- 10

Otros aspectos de la invención se detallan en las reivindicaciones.

En el marco de la presente invención, los pseudo-transposones se designan de la manera siguiente: ITR - UTR - secuencia nucleotídica exógena de interés - UTR - ITR. En el caso del ejemplo particular, ITR 3' - UTR 3' - secuencia nucleotídica exógena de interés - UTR 3' - ITR 3', abreviado 33sec33 o 33T33. La designación 33T33 podrá eventualmente hacer referencia más particularmente al pseudo-transposón en el que la secuencia nucleotídica exógena de interés es el gen de resistencia a la tetraciclina (véanse en particular los ejemplos siguientes). Evidentemente, esto se desprenderá claramente del contexto.

15

Tal como se ha indicado anteriormente, la transposasa *Mos-1* suministrada en *trans* en este sistema podrá ser una transposasa *Mos-1* mutante y, en particular, una transposasa *Mos-1* mutante hiperactiva. Por ejemplo, una transposasa *Mos-1* mutante hiperactiva apropiada comprenderá por lo menos una mutación a nivel de por lo menos un residuo seleccionado de entre los residuos siguientes de la secuencia SEC ID nº 2: F53, Q91, E137, T216 e Y237.

20

De esta forma, unos sistemas de transposición preferidos en el sentido de la presente invención comprenden por lo menos las dos parejas siguientes:

- 25 a) un pseudo-transposón *Mos-1* hiperactivo que comprende por lo menos una ITR40 de secuencia SEC ID nº 39 y una transposasa *Mos-1* mutante hiperactiva que comprende las mutaciones T216A e Y237C;
- b) un pseudo-transposón *Mos-1* hiperactivo que comprende por lo menos una ITR46 de secuencia SEC ID nº 38 y una transposasa *Mos-1* mutante hiperactiva que comprende las mutaciones T216A e Y237C, o E137K y T216A, o F53Y y T216A e Y237C;
- 30 c) un pseudo-transposón *Mos-1* hiperactivo 3sec33 y una transposasa *Mos-1* mutante hiperactiva que comprende las mutaciones T216A e Y237C, o E137K y T216A, o también F53Y y Q91R, o bien también F53Y y Q91R y E137K y T216A.

Se describe asimismo un sistema de transposición recombinante hiperactivo que comprende por lo menos las dos parejas siguientes:

- 35 a) un pseudo-transposón *Mos-1* en el que una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa *Mos-1* de origen; y
- b) una transposasa *Mos-1* hiperactiva suministrada en *trans* de dicho pseudo-transposón y que comprende por lo menos:
- una mutación a nivel de por lo menos un residuo seleccionado de entre los residuos siguientes de la secuencia SEC ID nº 12: F53, Q91 e Y237, y/o
  - 40 - la mutación T216A,

siendo la frecuencia de transposición de la secuencia nucleotídica exógena de interés de este sistema mejorada en un factor por lo menos igual a 5, preferentemente por lo menos igual a 10.

Preferentemente, la transposasa *Mos-1* hiperactiva comprende por lo menos una mutación seleccionada de entre las mutaciones F53Y, Q91R, T216A, Y237C y sus combinaciones. Puede comprender además una mutación a nivel del residuo E137, en particular la mutación E137K, pero con la exclusión sin embargo de la combinación de mutaciones Q91R+E137K+T216A o F53Y+E137K+T216A.

45

Ventajosamente, en dicho sistema, por lo menos una de las dos repeticiones terminales no traducidas (UTR) y/o por lo menos una de las dos repeticiones terminales inversas (ITR) del pseudo-transposón *Mos-1* podrá(n) estar genéticamente modificada(s).

50 En los sistemas de transposición de acuerdo con la presente invención, la transposasa *Mos-1*, salvaje o

genéticamente modificada, se suministra en *trans* del pseudo-transposón. Así, ésta puede ser codificada por una secuencia nucleotídica dispuesta en un vector, bajo el control de elementos de regulación de la expresión. Ventajosamente, la expresión de la transposasa será entonces inducible. Para ello, el experto en la materia podrá utilizar unos promotores convencionales. Por ejemplo, podrá utilizar el promotor bien conocido *pLac* inducible con el IPTG.

5 Alternativamente, la transposasa podrá ser añadida en el sistema de transposición en forma de una proteína o de una fracción proteica funcional purificada. O bien, puede ser añadida en forma de un ARN mensajero. En este caso, la expresión de la transposasa estará limitada en el tiempo (hasta algunas horas) debido a la labilidad de los ARN mensajeros.

De manera particularmente interesante, la secuencia nucleotídica exógena de interés contenida por el pseudo-transposón *Mos-1* es un gen funcional. En el sentido de la invención, un gen se denomina "funcional" si la secuencia nucleotídica correspondiente comprende por lo menos el marco abierto de lectura (ORF), es decir la secuencia codificante, apta para dar lugar a una secuencia en aminoácidos que presenta la actividad del producto del gen salvaje. Preferentemente, el gen funcional es el gen salvaje. Sin embargo, se puede tratar de un gen que comprende una o varias mutaciones a partir del momento en el que el producto del gen mutado sigue siendo activo (incluso si la actividad del producto del gen mutado es más baja que la del producto del gen nativo). Así, esta definición de "gen funcional" abarca asimismo las secuencias codificantes desprovistas de su promotor. A título de ejemplo, la secuencia nucleotídica exógena de interés puede ser un gen de resistencia a un antibiótico, desprovisto o no de su promotor (por ejemplo el gen de resistencia a la tetraciclina), o cualquier otro marcador de selección apropiado.

En el sentido de la invención, una "mutación" está de acuerdo con la acepción habitual en biotecnología. Así, una mutación puede ser una sustitución, una adición o una deleción de una o varias bases en una secuencia nucleotídica, o de uno o varios aminoácidos en una secuencia proteica. Una "mutación" puede designar en particular una sustitución de por lo menos una base de un codón de una secuencia nucleotídica, provocando dicha sustitución por ejemplo, durante la traducción de la secuencia nucleotídica en cuestión, la incorporación de un aminoácido diferente en el lugar del aminoácido nativo, en la secuencia proteica resultante. Como regla general, se preferirá que la o las mutaciones no provoque(n) ninguna pérdida de la función biológica del producto mutado. Por el contrario, una disminución de actividad podrá eventualmente ser tolerada, salvo si el elemento mutado (por ejemplo el pseudo-transposón y/o la transposasa) es "hiperactivo", en cuyo caso su actividad deberá ser superior a la del elemento salvaje correspondiente.

Se considera así que una transposasa es "hiperactiva" o que un pseudo-transposón es "hiperactivo" si la eficacia de transposición observada cuando se utiliza ha aumentado. El aumento de actividad del elemento en cuestión (pseudo-transposón o transposasa) se determina durante su utilización en un sistema de transposición recombinante de acuerdo con la invención. El factor de hiperactividad puede así ser determinado y, si el nivel de actividad transposicional obtenido alcanza el umbral fijado en el marco de la invención, el elemento se denomina "hiperactivo".

De manera general, una "modificación genética" debe ser entendida en este caso como equivalente a una o varias mutaciones. Si una secuencia codificante está genéticamente modificada, entonces, típicamente, contiene una o varias mutaciones. Estas mutaciones no deben sin embargo provocar una pérdida de la función del producto codificado. Al contrario, en el caso, por ejemplo, de una transposasa *Mos-1* genéticamente modificada, se buscará preferentemente un aumento de su actividad (es decir una transposasa recombinante hiperactiva). Si una secuencia no codificante está genéticamente modificada (por ejemplo una ITR o una UTR) entonces o bien contiene una o varias mutaciones, o bien no es la secuencia individual como tal la que se modifica, sino el número de repeticiones de esta secuencia y/o el orden en la cadena de las secuencias y/o la orientación de esta secuencia con respecto a las demás, que está(n) modificada(s) con respecto a la configuración normal (por ejemplo, con respecto al transposón *Mos-1* salvaje). A título de ejemplos de modificaciones genéticas de una secuencia no codificante tal como una ITR o una UTR, se citará en particular la deleción de la totalidad o parte de la secuencia, la permutación de la secuencia con otras secuencias presentes en el entorno, etc. En resumen, una disposición diferente de secuencias, ya sean codificantes o no, está comprendida en la definición de las "modificaciones genéticas" según la invención.

De acuerdo con la descripción anterior, se puede utilizar, en los sistemas de transposición recombinantes previstos por la presente invención, un pseudo-transposón *Mos-1* hiperactivo, en el que:

- a) por lo menos una de las dos repeticiones terminales no traducidas (UTR) y/o por lo menos una de las dos repeticiones terminales inversas (ITR) está(n) genéticamente modificada(s); y
- b) una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa *Mos-1* de origen, caracterizado porque se selecciona de entre los pseudo-transposones siguientes:
  - α) ITR3'-UTR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3' (pseudo-transposón 33sec33),
  - β) ITR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3' (pseudo-transposón 3sec33),
  - γ) ITR3'-UTR5'-secuencia nucleotídica exógena de interés-UTR3'-ITR5' (pseudo-transposón 35sec35),

- δ) los pseudo-transposones que comprenden por lo menos una ITR40 de secuencia SEC ID nº 39, y
- ε) los pseudo-transposones que comprenden por lo menos una ITR46 de secuencia SEC ID nº 38, y en el que ITR3', UTR3', UTR5' e ITR5' designan respectivamente la secuencia de la repetición terminal inversa 3' del transposón Mos-1, la secuencia de la repetición terminal no traducida 3' del transposón Mos-1, la secuencia de la repetición terminal no traducida 5' del transposón Mos-1 y la secuencia de la repetición terminal inversa 5' del transposón Mos-1.

Se excluye el pseudo-transposón 2 ITR 3' de los pseudo-transposones *Mos-1* susceptibles de ser utilizados en el marco de la invención. Tal como se ha indicado anteriormente, éste se utiliza como referencia para determinar la hiperactividad de los sistemas de transposición según la invención.

En algunos casos, por lo menos una de las dos repeticiones terminales no traducidas (UTR) del pseudo-transposón *Mos-1* hiperactivo está genéticamente modificada. Alternativa o adicionalmente, por lo menos una de las dos repeticiones terminales inversas (ITR) del pseudo-transposón *Mos-1* hiperactivo está genéticamente modificada.

De acuerdo con lo expuesto anteriormente y tal como se ha ilustrado en los ejemplos siguientes, la o las ITR genéticamente modificada(s) del pseudo-transposón *Mos-1* recombinante hiperactivo puede(n) ser ventajosamente seleccionada(s) de entre las secuencias de ITRSelex denominadas ITR40 (SEC ID nº 39) e ITR46 (SEC ID nº 38).

Alternativamente, el pseudo-transposón *Mos-1* recombinante hiperactivo puede comprender una combinación de ITR y UTR seleccionada en particular de entre las combinaciones siguientes: ITR 3' + UTR 5'/UTR 3' + ITR 5' (anotada 35T35 o 35sec35), ITR 3' + UTR 3'/UTR 3' + ITR 3' (anotada 33T33 o 33sec33), ITR 3'/UTR 3' + ITR 3' (anotada 3T33 o 3sec33).

En algunos modos de realización, la transposasa *Mos-1* suministrada en *trans* en el sistema de transposición recombinante es una transposasa hiperactiva que comprende por lo menos una mutación a nivel de por lo menos un residuo seleccionado de entre los residuos siguientes de la secuencia SEC ID nº 2: F53, Q91, E137, T216 e Y237, con la exclusión de la combinación de mutaciones Q91R+E137K+T216A o F53Y+E137K+T216A. Los resultados de los experimentos realizados por los inventores revelan que no se puede practicar cualquier mutación o combinación de mutaciones sobre la transposasa *Mos-1* para obtener una transposasa recombinante hiperactiva. En particular, se observa (véase la parte experimental a continuación) que las dos combinaciones de mutaciones excluidas en este caso conducen a una abolición de la transposición. En estas combinaciones excluidas, las mutaciones pueden ser por lo tanto consideradas como antagonistas.

Ventajosamente, la transposasa *Mos-1* hiperactiva podrá comprender por lo menos una mutación seleccionada de entre las mutaciones F53Y, Q91R, E137K, T216A, Y237C, y sus combinaciones con la exclusión de la combinación de mutaciones Q91R+E137K+T216A o F53Y+E137K+T216A.

Preferentemente, la transposasa *Mos-1* hiperactiva podrá comprender por lo menos una mutación a nivel del residuo T216. Más preferentemente, podrá comprender por lo menos la mutación T216A. Ventajosamente, dicha transposasa hiperactiva podrá comprender además por lo menos una mutación seleccionada de entre las mutaciones F53Y, Q91R, E137K, Y237C, y sus combinaciones, con la exclusión de la combinación de mutaciones Q91R+E137K+T216A o F53Y+E137K+T216A.

Tal como se desprende de la parte experimental a continuación, unas transposasas *Mos-1* hiperactivas particularmente interesantes comprenden por lo menos una de las combinaciones de mutaciones siguientes:

- T216A+Y237C; F53Y+T216A+Y237C: factor de hiperactividad por lo menos igual a 20;
- F53Y+Q91R+E137K+T216A+Y237C: factor de hiperactividad por lo menos igual a 30;
- F53Y+E137K+T216A+Y237C: factor de hiperactividad por lo menos igual a 45.

Según un modo de realización alternativo o adicional, la transposasa *Mos-1* suministrada en *trans* en los sistemas de transposición recombinantes de la presente invención es una transposasa hiperactiva que comprende por lo menos una mutación a nivel de un residuo fosforilable, siendo dicha mutación supresiva de un sitio de fosforilación en célula eucariota (por ejemplo, célula de planta, de vertebrado o de invertebrado). En este caso también, la o las mutaciones previstas, supresivas de uno o varios sitio(s) de fosforilación en célula eucariota son, evidentemente, conservadoras de la función enzimática de la proteína. Unas transposasas mutantes hiperactivas apropiadas se describen en la solicitud de patente francesa nº 03 00905 presentada el 28 de enero de 2003. En particular, estas transposasas son tales que el residuo fosforilable mutado se selecciona de entre los residuos siguientes de la secuencia SEC ID nº 2: T11, T24, S28, T42, T88, S99, S104, T135, S147, T154, S170, T181, S200, T216, T255 y S305. Preferentemente, la transposasa hiperactiva comprende por lo menos una mutación a nivel del residuo fosforilable T88. Ventajosamente, comprende además por lo menos una mutación a nivel de los residuos fosforilables T11, T24, S28, T42, S99, S104, T135, S147, T154, S170, T181, S200, T216, T255 y S305. En particular, dicho o dichos residuos fosforilables así mutados son sustituidos por uno o varios residuos no fosforilables en célula eucariota.

De manera general en el marco de la invención, la transposasa Mos-1 recombinante hiperactiva o salvaje, puede ventajosamente ser producida de manera estable en los procariotas. En este caso, un procedimiento apropiado de producción, por una célula hospedante procariota, de una transposasa Mos-1 activa (eventualmente hiperactiva) y estable comprende por lo menos las etapas siguientes:

- 5 a) la clonación de la secuencia nucleotídica que codifica la transposasa activa en un vector de expresión;
- b) la clonación de la secuencia nucleotídica que codifica la sub-unidad catalítica activa de la proteína quinasa dependiente de AMPc (pKa) en un vector de expresión;
- c) la transformación de dicha célula hospedante con dichos vectores de expresión;
- d) la expresión de dichas secuencias nucleotídicas por dicha célula hospedante; y
- 10 e) la obtención de la transposasa activa y estabilizada mediante fosforilación por la pKa.

Alternativamente, las clonaciones de las etapas a) y b) se realizan en un solo vector de expresión.

De manera interesante, dichos procedimientos comprenden además una etapa de purificación de la transposasa.

- 15 Para más detalles referentes a estos procedimientos, el experto en la materia podrá referirse a la solicitud de patente francesa FR 2 893 951 presentada el 30 de noviembre 2005.

Un primer aspecto de la presente invención se refiere a un pseudo-transposón *Mos-1* hiperactivo, en el que:

- a) por lo menos una de las dos repeticiones terminales no traducidas (UTR) y/o por lo menos una de las dos repeticiones terminales inversas (ITR) está(n) genéticamente modificada(s); y
- b) una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa Mos-1 de origen,

siendo dicho pseudo-transposón seleccionado de entre los pseudo-transposones siguientes:

- α) ITR3'-UTR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3'- (pseudo-transposón 33sec33),
- β) ITR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3' (pseudo-transposón 3sec33),
- γ) ITR3'-UTR5'-secuencia nucleotídica exógena de interés-UTR3'-ITR5' (pseudo-transposón 35sec35),
- 25 δ) los pseudo-transposones que comprenden por lo menos una ITR40 de secuencia SEC ID n° 39, y
- ε) los pseudo-transposones que comprenden por lo menos una ITR46 de secuencia SEC ID n° 38.

Otros aspectos de la invención se detallan en las reivindicaciones.

- 30 Una célula hospedante tal como se define en las reivindicaciones se selecciona de entre las células procariotas (bacterias por ejemplo, en particular *Escherichia coli*) y las células eucariotas (en particular, las células de plantas, de vertebrados y de invertebrados).

Un kit tal como se define en las reivindicaciones podrá comprender además uno o varios elementos seleccionados de entre, en particular, un tampón compatible con la o las transposasas, un tampón "stop" para detener las reacciones de transposición, uno o varios ADN de control (control de reacción), unos oligonucleótidos útiles para la secuenciación después de la reacción, unas bacterias competentes, un folleto de utilización, etc.

- 35 En un modo de realización, uno por lo menos de estos medios se utiliza para la transposición eficaz *in vitro* o *in vivo* (en particular en una célula hospedante de planta) o *ex vivo* de una secuencia nucleotídica exógena de interés en una célula eucariota o procariota, con la excepción de una célula en un organismo humano o animal.

Según otro modo de realización, por lo menos uno de estos medios se utiliza para la preparación de un medicamento destinado a permitir la transposición eficaz *in vivo* de una secuencia nucleotídica exógena de interés.

- 40 Alternativamente, por lo menos uno de estos medios se utiliza para la preparación de un medicamento que resulta de la transposición *in vitro* o *ex vivo* de una secuencia de ADN transposable de interés (secuencia nucleotídica exógena de interés) en una secuencia de ADN diana. Por ejemplo, la invención propone un procedimiento de preparación de un medicamento, que comprende por lo menos una etapa de transposición (es decir, *in vitro* o *ex vivo*) de una secuencia de ADN transposable de interés en una secuencia de ADN diana, siendo dicha transposición mediada por lo menos por uno de los medios de la invención. El medicamento puede así ser preparado *ex vivo* si la transposición se realiza *in vitro*.
- 45

5 Los medios propuestos en el marco de la presente invención pueden por ejemplo permitir modificar unas células con el fin de expresar una proteína medicamento (es decir, una proteína de interés terapéutico o profiláctico, por ejemplo la insulina, un anticuerpo particular, etc.). Estos medios pueden asimismo permitir "corregir" unas células con el fin de restaurar una función biológica deficiente. Según todavía otro modo de realización, por lo menos uno de estos medios se utiliza para la mutagénesis de inserción, o también para la secuenciación y/o la clonación de ácidos nucleicos.

10 Estas aplicaciones implican de manera general la utilización de una transposición *in vitro* o *in vivo* (en particular en unas células hospedantes de planta) en una célula eucariota o procariota con la excepción de una célula en un organismo humano o animal, lo cual se deduce de los conocimientos generales del experto en la materia en el campo de la invención (Ausubel *et al.*, 1994; Craig *et al.*, 2002). Haciendo referencia más particularmente a las transposiciones *in vivo*, la secuencia de ADN diana es típicamente el genoma del hospedante, que puede ser un organismo eucariota (por ejemplo una célula hospedante de planta) o procariota, o un tejido de un organismo, o también una célula de un organismo o de un tejido con la excepción de un organismo humano o animal o de una célula o de un tejido en un organismo humano o animal.

15 En todos los casos, las numerosas aplicaciones de los medios de la invención, para las cuales éstos resultan ser de un interés considerable, recurren a unas técnicas convencionales de biología molecular bien conocidas por el experto en la materia.

Asimismo, se describe la utilización de una transposasa Mos-1 hiperactiva que comprende por lo menos:

- una mutación a nivel de por lo menos un residuo seleccionado de entre los residuos siguientes de la secuencia SEC ID nº 2: F53, Q91 e Y237, y/o
- 20 - la mutación T216A

para mejorar, en un factor por lo menos igual a 5, preferentemente por lo menos igual a 10, la frecuencia de transposición de una secuencia nucleotídica de interés contenida por un pseudo-transposón *Mos-1* en el que dicha secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa Mos-1 de origen.

25 La utilización de dicha transposasa se realiza en particular en un sistema de transposición recombinante hiperactivo tal como se ha descrito anteriormente.

En particular, dicha transposasa Mos-1 hiperactiva comprenderá por lo menos una mutación seleccionada de entre las mutaciones F53Y, Q91R, T216A, Y237C, y sus combinaciones. Podrá comprender además una mutación a nivel del residuo E137, en particular la mutación E137K, con la exclusión de la combinación de mutaciones Q91R+E137K+T216A o F53Y+E137K+T216A.

30 En la práctica, la transposasa Mos-1 hiperactiva estará generalmente codificada por una secuencia nucleotídica dispuesta sobre un vector, bajo el control de elementos de regulación de la expresión. La expresión de la transposasa será así ventajosamente inducible.

35 De acuerdo con la descripción anterior, la transposasa Mos-1 hiperactiva forma parte preferentemente de un sistema de transposición recombinante hiperactivo, en el que se suministra en *trans* de un pseudo-transposón *Mos-1* tal como se ha descrito anteriormente. En particular, la secuencia nucleotídica exógena de interés contenida en el pseudo-transposón *Mos-1* es un gen funcional. Por otro lado, por lo menos una de las dos repeticiones terminales no traducidas (UTR) y/o por lo menos una de las dos repeticiones terminales inversas (ITR) de dicho pseudo-transposón *Mos-1* está(n) genéticamente modificada(s). En la práctica, será ventajoso utilizar un pseudo-transposón *Mos-1* contenido por un vector.

40 Las figuras siguientes se proporcionan a título puramente ilustrativo y no limitan de ninguna manera el objeto de la presente invención.

45 **Figura 1:** Secuencia nucleotídica del gen de la transposasa del transposón *Mos-1* (SEC ID nº 1) y secuencia proteica de la transposasa Mos-1 (SEC ID nº 2). Los sitios de hiperactividad (sitios diana de la mutagénesis dirigida) están localizados mediante un cuadro simple alrededor de los residuos que aparecen sobre un fondo gris. Los sitios putativos de fosforilación están localizados de la manera siguiente:

- cuadro doble: aminoácido fosforilable por las ATM quinasas;
- cuadro rayado en negrita: aminoácido fosforilable por la proteína quinasa C (pKc);
- cuadro gris claro grueso: aminoácido fosforilable por la proteína quinasa AMPc dependiente (pKa);
- 50 - cuadro negro grueso: aminoácido fosforilable por la pKa y la proteína quinasa dependiente de GMPc (pKg);
- cuadro rayado simple: aminoácido fosforilable por la caseína quinasa II (CKII).

Los residuos QTQ (posiciones 87 a 89 de la secuencia proteica SEC ID nº 2) corresponden al sitio putativo de fosforilación por la familia de las ATM quinasas. Los círculos sobre fondo punteado marcan los residuos altamente fosforilables. Los círculos sobre fondo gris identifican los residuos implicados en la triada catalítica característica de las transposasas de MLE (D,D34-35[D/E]) e implicados en la escisión del ADN.

5 Las flechas verticales indican los sitios de corte y empalme proteolítico.

Figura 2: Esquema que representa la estructura de la transposasa del elemento *Mos-1*.

N-term: dominio N-terminal responsable de la unión a las ITR;

C-term: dominio C-terminal responsable de la catálisis de la transferencia de las hebras de ADN; NLS: señal putativa de localización nuclear (o internacionalización nuclear); HTH: motivo hélice-vuelta-hélice; aa: aminoácido.

10 Los números indican las posiciones de los aminoácidos.

Se señala la triada catalítica característica [D, D34 (D/E)].

Figura 3: Representación esquemática del plásmido pBC3Neo3, construido a partir del plásmido pBC SK+ (Stratagene), y de sus derivados (A y B).

Figura 4: Representación esquemática del plásmido pCMV-Tnp y de sus derivados (A y B).

15 Figura 5: Representación esquemática del ensayo de transposición.

A): bacteria co-transformada con el vector de expresión que codifica la transposasa (Tnp) y el vector reportero de la transposición.

B): Evento de transposición después de la inducción del vector de expresión.

C): Determinación de la frecuencia de transposición.

20 Figura 6: Eficacia de la transposición de las transposasas mutantes: factor de hiperactividad y frecuencia de transposición de los mutantes.

i) Mutantes simples:

- A t = 0h después de la inducción: A) Factor de hiperactividad; B) Frecuencia de transposición;

- A t = 5h después de la inducción: C) Factor de hiperactividad; D) Frecuencia de transposición;

25 ii) Mutantes múltiples:

- A t = 0h después de la inducción: F) Factor de hiperactividad; F) Frecuencia de transposición;

- A t = 5h después de la inducción: G) Factor de hiperactividad; H) Frecuencia de transposición;

WT(53): transposasa salvaje.

Figura 7: Esquema que ilustra el principio general del método SELEX.

30 Figura 8: Esquema que ilustra el principio del método SELEX 1.

Figura 9: Esquema que presenta los ensayos de competición.

Figura 10: Esquema que representa las condiciones de realización de los ensayos de transposición *in vivo* en bacterias según el método I.

35 Figura 11: Resultados de los geles de retraso (B) realizados con las ITR presentadas en (A) (ITRSelex seleccionadas mediante los métodos SELEX realizados por los inventores).

Figura 12: Resultados de los ensayos de competición.

Figura 13: A) Resultados de los ensayos de transposición obtenidos con las ITR y B) Resultados complementarios de los ensayos de transposición en bacteria con las ITRSelex y la transposasa *Mos-1* salvaje.

40 Figura 14: Esquema que ilustra las condiciones de realización del método II de transposición *in vivo* en bacterias.

Figura 15: A) Resultados de los ensayos de transposición obtenidos con las ITR/UTR y B) Resultados complementarios de los ensayos de transposición en bacteria con las combinaciones ITR/UTR y la transposasa *Mos-1*

salvaje.

Figura 16: Representación gráfica de la cantidad de complejos [Transposasa Mos-1 salvaje + ITR] formados con unas combinaciones ITR/UTR.

5 La parte experimental siguiente, apoyada por unos ejemplos y unas figuras, ilustra de manera no limitativa la invención.

## EJEMPLOS

### PARTE I: TRANSPOSASAS MOS-1 MUTANTES HIPERACTIVAS (Tnp)

#### **I- MATERIAL Y MÉTODOS**

##### **I-A- Vectores utilizados**

##### 10 **I-A-1 Descripción de los plásmidos utilizados**

15 El vector pGEM-T-Easy (3,1 kb) (Promega Charbonnières Francia; cat nº A1360) posee los cebadores de secuenciación Pu y Prev (Ausubel *et al.*, 1004) y el gen de resistencia a la ampicilina. Se ha concebido para clonar unos productos de PCP en el gen LacZ, lo cual permite un cribado blanco/azul de las colonias bacterianas obtenidas sobre cajas LB ampicilina, en presencia de X-Gal y de IPTG. Se ha utilizado para dar el pGEM-T (Tnp) (Augé-Gouillou *et al.*, 2001). Este último sirve de matriz para la mutagénesis de la transposasa antes de su sub-clonación en los vectores pKK-233-2 y pCS2+.

20 El vector pKK-Tnp (5,6 kb) permite una expresión fuerte de la transposasa a través de un promotor *P<sub>lac</sub>* inducible con el IPTG. El promotor no es modulable y presenta una fuga de expresión natural. El pKK-Tnp contiene asimismo el gen de resistencia a la ampicilina. Este plásmido se deriva del vector pKK-233-2 (Clontech, Ozyme, Saint Quentin en Yvelines, Francia).

El vector pMalC2x-Tnp, derivado del pMalC2x (New England Biolabs, Ozyme, Saint Quentin en Yvelines, Francia), permite una expresión de la transposasa fusionada en N-terminal con la proteína MBP (Maltose binding protein). Contiene el gen de resistencia a la ampicilina.

25 El vector pCS2+-Tnp se deriva del vector pCS2+ (Turner DL *et al.*, 1994) y permite una expresión de la transposasa en células eucariotas bajo el control del promotor CMV<sub>ie</sub>. Este vector permite asimismo sintetizar unos ARN *in vitro* que corresponden al ARN mensajero que codifica para la transposasa. La transcripción se realiza bajo el control del promotor SP6 y el ARN se poliadenila.

30 El pBC 3T3 es un plásmido donante de pseudo mariner Mos-1 (Augé-Gouillou *et al.*, 2001). Contiene el gen de resistencia a la tetraciclina "OFF" (es decir sin promotor) rodeado por dos ITR 3'. Así, sólo las bacterias que han efectuado la transposición del pseudo-transposón corriente abajo de un promotor (promotor "tagging") son seleccionadas sobre cajas LB tetraciclina. El vector contiene el gen de resistencia al cloramfenicol.

35 El pBC3Neo3 (figura 3) es un plásmido donante de pseudo mariner Mos-1. Contiene el gen de resistencia a la neomicina bajo el control del promotor SV40 rodeado por dos ITR 3'. Esto permite la selección de células eucariotas en las que el gen de resistencia a la neomicina ha sido integrado en el genoma de la célula. La selección se efectúa con la ayuda de G418 (800 µg/ml) durante 2 semanas. El vector contiene el gen de resistencia al cloramfenicol.

40 Los plásmidos pBC 3T3 y pBC3Neo3 contienen por lo tanto el pseudo-transposón *Mos-1* en el que la repetición terminal inversa salvaje situada en 5' (ITR 5') ha sido mutada de tal manera que es una copia perfecta de la repetición terminal inversa salvaje situada en 3' (ITR 3'). Este pseudo-transposón está por lo tanto rodeado por 2 ITR 3' ("pseudo-transposón 2 ITR3'"). Este pseudo-transposón ha sido utilizado, en asociación con la transposasa Mos-1 salvaje, para determinar la frecuencia de transposición de referencia, a partir de la cual se han determinado los factores de hiperactividad asociados a la realización de los sistemas según la invención.

De manera general, en este tipo de construcción (pseudo-transposones), la letra "T" representa el gen reportero que confiere la resistencia a la tetraciclina.

45 El pBC KS Neo es un plásmido que contiene el gen de resistencia a la neomicina bajo el control del promotor SV40. El vector contiene el gen de resistencia al cloramfenicol.

El vector pGL3-Control (Promega; Cat. nº E1741) es un plásmido que contiene el gen que codifica para la luciferasa bajo el control del promotor SV40. Contiene asimismo el gen de resistencia a la ampicilina.

##### **I-A-2 Construcción de los vectores**

##### **I-A-2-1- Preparación del ADN de los vectores**

5 Para las diferentes construcciones, todas las eluciones de ADN a partir de un gel de agarosa han sido realizadas con el kit Wizard SV Gel and PCR Clean-Up system (Promega, Francia). Todas las mini-preparaciones de plásmidos a partir de cultivos bacterianos han sido efectuadas con el kit Wizard Plus minipréps (Promega). Las preparaciones a mayor escala de ADN han sido realizadas con el kit Pureyield plasmid midiprep system (Promega) o con los kits Midiprep o Maxiprep (Qiagen).

**I-A-2-2- Mutagénesis dirigida para la obtención de los mutantes**

10 La mutagénesis dirigida ha sido realizada según el protocolo del kit Quikchange site directed mutagenesis (Stratagene). Los oligonucleótidos que permiten introducir la mutación han sido sintetizados por MWG Biotech (Roissy CDG). Las enzimas Pfu polimerasa y Dpn1 han sido suministradas por Promega (Francia). Brevemente, la mutación a introducir está contenida por dos oligonucleótidos complementarios. La totalidad del plásmido ha sido amplificada por PCR (95°C durante 1 minuto y después 16 ciclos a 95°C durante 30 segundos, 55°C durante 1 minuto, 68°C durante 2 minutos/kb de plásmido). El plásmido que ha servido de matriz en la PCR ha sido digerido por Dpn1 (1h a 37°C, 2-3u/50 µl de PCR). Se han transformado a continuación 2 -3 µl de la PCR tratada por Dpn1 en unas bacterias XL1 Blue quimiocompetentes o JM109 electrocompetentes.

15 Las secuencias de los oligonucleótidos utilizadas para la mutagénesis dirigida se indican en la tabla 1 siguiente, y la mutación se precisa mediante los nucleótidos en negrita.

**Tabla 1**

Mutación	Cebador directo	SEC ID n°	Cebador inverso	SEC ID n°
F53Y	<b>ggtggttcaacgctacaaaagtggtgatttgacg</b>	3	cgtaaaatcaccactttgtagcgtgaaaccacc	4
Q91R	gctcaaacgaaaaac <b>g</b> actgcgagagc	5	gctctgcgagtc <b>g</b> ttttgcgttgagc	6
L92A	cgaaaaacaag <b>cc</b> gcagagcagttgg	7	ccaactgctctgc <b>g</b> gctgttttgcg	8
L92R	cgaaaaacaac <b>gc</b> gcagagcagttgg	9	ccaactgctctgc <b>gc</b> gtgttttgcg	10
Q100N	gcagttggaagtaagta <b>ac</b> caagcagttcc	11	ggaaactgct <b>gt</b> tactactccaactgc	12
Q100R	gcagttggaagtaagtc <b>g</b> acaagcagttcc	13	ggaaactgct <b>gtc</b> gactactccaactgc	14
Q100E	gcagttggaagtaagtg <b>aa</b> caagcagttcc	15	ggaaactgct <b>gtc</b> actactccaactgc	16
S104P	<b>ggaagtaagtcaacaagcagttccaatcgcttgca</b> <b>gagatgg</b>	17	<b>ccatctctcgcaagcgattgggaactgettgtgactact</b> <b>tcc</b>	18
N105A	caagcagttcc <b>gc</b> acgcttgcgag	19	ctcgaagcgt <b>gc</b> ggaactgcttg	20
E137K	ggcgaaaaacacatg <b>caaa</b> atttgccttcacg	21	cgtgaaagcaaa <b>att</b> tgcattgctttgcgcc	22
T216A	gcgaaacggtgaat <b>gc</b> ggcagctacc	23	ggtagcgtgcc <b>gc</b> attcaccgttcgc	24
Y237C	<b>gcttcagagaaaaacgaccggaatgcaaaaagac</b> <b>aacacagg</b>	25	<b>cctgtgtgtctttttgacattccggctgtttctctgaagc</b>	26
W268F	cgttgaaacactcaat <b>ttc</b> gaagtgttccgc	27	gcggaagcacttc <b>gaa</b> attgaggtttccaacg	28
W268Y	cgttgaaacactcaat <b>tac</b> gaagtgttccgc	29	gcggaagcacttc <b>gta</b> attgaggtttccaacg	30
W268A	cgttgaaacactcaat <b>gc</b> ggaagtgc	31	gcacttc <b>gc</b> attgaggtttccaacg	32

**I-A-2-3- Verificación de la secuencia de las transposasas**

20 La introducción de mutaciones en la transposasa clonada en el plásmido pGEM-T easy ha sido verificada mediante secuenciación. Para ello, se han enviado 10 microlitros de una minipreparación de ADN para su secuenciado a la compañía MWG Biotech. Los cebadores utilizados (Puniv -21 y Prev -49) han sido suministrados por esta misma compañía.

**I-A-2-4- Sub-clonación de las transposasas mutantes en el plásmido pKK-233-2.**

25 El fragmento que codifica para la transposasa salvaje o mutante (Tnp) ha sido preparado a partir del vector

5 pGEM-T (Tnp) mediante digestión Nco1/HindIII, y eluido sobre gel de agarosa al 0,8% (TAE1X: de Tris-acetato 0,04 M, EDTA 1 mM pH 8). El plásmido pKK-233-2 ha sido dirigido mediante HindIII/Nco1, depositado sobre gel de agarosa, eluido y después ligado con el fragmento que codifica para la transposasa de Mos-1 (denominada Tnp), durante una noche a 16°C. Un control de recirculación del plásmido sobre sí mismo ha sido realizado efectuando una ligadura del plásmido en ausencia de fragmento que codifica para la transposasa.

10 El producto de ligadura ha sido utilizado para transformar unas bacterias *E. coli* JM109 que han sido seleccionadas a continuación sobre cajas LB-ampicilina (100 µg/ml). Se han cultivado 4 clones resistentes a la ampicilina para una extracción del plásmido. Las minipreparaciones de ADN han sido controladas mediante digestión EcoR1/HindIII y después en electroforesis sobre gel de agarosa al 0,8% (TAE 1X) con el fin de asegurar que habían integrado el gen que codifica para la transposasa.

#### **I-A-2-5- Sub-clonación de las transposasa mutantes en el plásmido pMalC2X**

Para la sub-clonación en pMalC2X, el gen que codifica para la transposasa ha debido ser reamplificado mediante PCR con la ayuda de los cebadores MTP up y 3' HindIII:

MTP up: 5'-TACGTAATGTCGAGTTTCGTGCCG (SEC ID n°33)

15 3'HindIII: 5'-CCCAAGCTTATTCAAAGTATTTGC (SEC ID n°34). Condiciones de ciclo: 95°C durante 5 minutos y después 20 ciclos (95°C durante 30 segundos, 50°C durante 1 minuto, 72°C durante 1 minuto) y después 72°C durante 5 minutos.

20 El producto de PCR ha sido depositado a continuación sobre gel, eluido en 50 µl y clonado en el plásmido pGEM-T easy (1 µl de vector + 2 µl de producto de PCR eluido). Después de la ligadura durante una noche a 16°C, el producto de ligación ha sido transformado en células JM109 mediante electroporación, y las bacterias han sido seleccionadas sobre caja LB ampicilina (100 µg/ml) que contiene X-Gal 2% IPTG 1 mM. Dos clones blancos resistentes a la ampicilina han sido cultivados para extraer el ADN plasmídico. Después del control de la clonación mediante digestión EcoR1 y electroforesis sobre gel de agarosa al 0,8% (TAE 1X), el plásmido ha sido enviado para su secuenciación a la compañía MWG Biotech.

25 Después de la verificación de la secuencia, el fragmento que codifica para la transposasa ha sido preparado mediante digestión SnaB1/HindIII y eluido sobre gel. Se ha ligado con el plásmido pMalC2X abierto mediante Xmn1/HindIII. Unas bacterias JM109 han sido transformadas a continuación por el producto de ligación y esparcidas sobre cajas LB ampicilina (100 µg/ml).

#### **I-A-2-6- Sub-clonación de las transposasas mutantes en el plásmido pCS2+**

30 El fragmento que codifica la transposasa salvaje o mutante (Tnp) ha sido preparado a partir del vector pGEM-T (Tnp) mediante digestión EcoR1 y eluido sobre gel de agarosa al 0,08% (TAE1X; Tris-acetato 0,04 M, EDTA 1 mM pH8). El plásmido pCS2+ ha sido abierto mediante EcoR1, desfosforilado y después ligado con el fragmento que contiene la transposasa. El producto de ligadura ha sido transformado en unas bacterias JM109 y los clones han sido seleccionados sobre cajas LB ampicilina (100 µg/ml). Se han cultivado ocho clones para extraer el ADN plasmídico. La presencia del inserto y su orientación han sido evaluadas mediante digestión PvuII/BamH1 o PvuII solo y electroforesis sobre gel de agarosa al 0,8% en TAE 1X. Los clones se designan sentido+ cuando el gen está insertado en el sentido que permite la transcripción del ARNm que corresponde a la transposasa bajo el control del promotor SP6. Los clones se designan sentido- cuando el gen está insertado en el sentido contrario a la transcripción bajo el control del promotor SP6.

#### **I-B- Análisis de la actividad de las transposasas mutantes en bacterias:**

##### **40 Ensayo de transposición en bacterias**

45 Unas bacterias *Escherichia coli* JM109 han sido co-transformadas con el plásmido pBC 3T3 (portador de un pseudo mariner 2 ITR 3' reportero de la transposición y del gen de resistencia al cloramfenicol) y con el vector inducible que codifica para la transposasa (pKK-Tnp o pMal-Tnp) portador del gen de resistencia a la ampicilina. La selección de las bacterias ha sido efectuada sobre ampicilina (100 µg/ml) y cloramfenicol con el fin de verificar la presencia de los dos plásmidos.

50 Las bacterias JM109 que contienen al mismo tiempo el plásmido pBC 3T3 y el plásmido pKK-Tnp (o pMalC2X-Tnp) han sido cultivadas durante 1 hora a 37°C en 250 µl de LB. Este inóculo ha sido después vertido en 5 ml de LB suplementado con IPTG 1 mM. El cultivo ha sido titulado sobre cajas LB (100 µl de una dilución al 1/1000) y sobre cajas LB-Tet (12,5 µg/ml) (250 µl del cultivo no diluido). El cultivo inducido por IPTG ha sido cultivado a continuación durante 5 horas a 32°C (temperatura óptima de la transposasa) bajo agitación (250 rpm). La suspensión bacteriana ha sido después titulada sobre cajas LB (100 µl de una dilución al 1/250000) y sobre cajas LB -Tet (100 µl de la suspensión bacteriana pura). Las cajas han sido dispuestas a 37°C durante la noche. Al día siguiente, las colonias han sido contadas sobre las cajas LB y LB-Tet, y después se ha calculado la frecuencia de transposición (igual al número de bacterias resistentes a la tetraciclina sobre el número de bacterias para 1 ml de suspensión bacteriana pura).

**I-C- Análisis de la actividad de las transposasas mutantes en células eucariotas**

**I-C-1- Transfección celular**

5 El día anterior a la transfección, se repartieron  $2.10^5$  células HeLa por pocillo de placas de 6 pocillos. Al día siguiente, las células se transfectaron con 3  $\mu$ g de ADN (750 ng de pGL3-Control, 750 ng de pCS2/Tnp, 1500 ng de pBC3T3) y PEI (proporción 1/10) (Eurogentec). Cada condición ha sido ensayada dos veces.

10 Dos días después de la transfección, la eficacia de transfección se evaluó lisando las células y evaluando la actividad luciferasa. Las células del segundo pocillo han sido tripsinadas y cultivadas en dos cajas de cultivo de 10 cm de diámetro en presencia de un agente de selección G418 (800  $\mu$ g/ml). El medio de cultivo ha sido cambiado cada dos días y la presión de selección ha sido mantenida durante quince días. Se han aislado y amplificado 5 clones por condición durante 15 días en presión de selección (200  $\mu$ g/ml de G418).

**I-C-2- Análisis molecular de los clones**

15 El ADN genómico de los diferentes clones se extrae y después se somete a un análisis mediante transferencia Southern después de la restricción enzimática. Este método permite analizar los diferentes sitios de inserción del casete de resistencia al G418. El análisis por secuenciación permite verificar la presencia de la duplicación del dinucleótido TA que firma los eventos de transposición real y evaluar la frecuencia de los eventos de recombinación aleatoria y dependiente de transposasa.

**II- RESULTADOS**

20 Para poder analizar todas las etapas de la transposición de mariner Mos-1 en un sistema más simple que las células eucariotas, se ha realizado un modelo de estudio en bacteria. La utilización de este sistema ha permitido evaluar, en diferentes tiempos de transposición, la eficacia de la transposición de las diferentes transposasas mutantes.

25 Los resultados obtenidos para las mutaciones simples se presentan en las figuras 6A y 6B para un tiempo de transposición T0 después de la inducción. Los resultados se expresan en mediana de factor de aumento con respecto a la transposasa salvaje. Las mutaciones más interesantes son las mutaciones situadas sobre las posiciones E137, Q91, Y237, T216. El mismo análisis, pero para un tiempo de transposición T=5h después de la inducción, ha dado los resultados ilustrados por las figuras 6C y 6D.

Los mutantes múltiples han sido realizados asociando las mutaciones más prometedoras. Se han obtenido siete dobles mutantes, tres triples, dos cúdruples y 1 quintuple, de acuerdo con la tabla 2 siguiente.

Tabla 2

Dobles mutantes	F53Y + T216A F53Y + Y237C F53Y + Q91R Q91R + Y237C E137K + T216A E137K+ Y237C T216A + Y237C
Triples mutantes	Q91R + E137K + T216A F53Y + E137K + T216A F53Y + T216A + Y237C
Cúdruples mutantes	F53Y + E137K + T216A + Y237C F53Y + Q91R + T216A + Y237C
Quíntuple mutante	F53Y + Q91R + E137K + T216A + Y237C

30 Los resultados de los ensayos de transposición realizados con los mutantes múltiples se indican en las figuras 6E a 6H.

Algunas asociaciones (Q91R E137K T216A; F53Y E137K T216A) provocan una pérdida completa de transposición. Las demás combinaciones mejoran la transposición en por lo menos un factor 6 para el tiempo T0 (figura 6E). Las combinaciones más interesantes son las asociaciones F53Y Q91R E137K T216A, F53Y E137K T216A Y237C y el quintuple mutante (figuras 6E y 6F).

5 Unos resultados similares han sido obtenidos para un tiempo de transposición de 5 horas (figuras 6G y 6H).

10 Se observará que unas mutaciones, descritas por otro lado en la bibliografía como que tienen un efecto sobre algunas propiedades de la transposasa Mos-1, no son tan interesantes cuando se trata de identificar unas transposasas Mos-1 mutantes hiperactivas. Es el caso, por ejemplo, de la mutación S104P, indicada en Zhang *et al.*, (2001) como que modifica la capacidad de la transposasa Mos-1 para crear unas interacciones proteicas. En el marco de sus investigaciones, los inventores han podido observar en efecto que esta mutación provocaba una abolición total de la actividad de transposición de la transposasa Mos-1 durante la realización de los ensayos de transposición en bacterias (datos no mostrados).

## **PARTE II: PSEUDO-TRANSPOSONES *Mos-1* RECOMBINANTES HIPERACTIVOS**

### **1- Las ITR (Inverted Terminal Repeat)**

15 Brevemente, la búsqueda de secuencias de ITR optimizadas ha sido realizada mediante la técnica SELEX que consiste en obtener mediante combinación un conjunto de ITR (figura 7). Se han conservado solamente algunas posiciones conocidas por ser esenciales para el buen funcionamiento de la transposasa. La naturaleza de los demás nucleótidos ha variado por lo tanto de manera aleatoria. Las diferentes ITR han sido seleccionadas y enriquecidas por su capacidad para fijar la transposasa. Entre las ITR seleccionadas, sólo algunas eran capaces de retrasar la transposasa en gel de retraso (técnica EMSA). La transposasa asociada a las ITR modificadas ha sido ensayada en un ensayo de transposición en bacteria para evaluar el impacto del cambio de secuencia de las ITR sobre la eficacia de transposición. Para algunas configuraciones (que afectan por ejemplo a uno o dos nucleótidos con respecto a la secuencia salvaje), la eficacia de la transposición ha mejorado en un factor significativo, en particular en un factor por lo menos igual a 5.

#### ***I-1- Material y métodos***

##### **25 a) Realización del método SELEX**

La utilización de la técnica SELEX ha permitido a los inventores seleccionar unas ITR que presentan una afinidad más grande para la transposasa que las ITR salvajes, con el fin de mejorar los rendimientos del pseudo-transposón *Mos-1* recombinante y del sistema de transposición objetos de la presente invención.

30 El método SELEX, descrito en 1990 (Ellington *et al.*, 1990; Tuerk *et al.*, 1990), permite seleccionar unos ácidos nucleicos en unas mezclas que contienen más de  $10^{15}$  moléculas diferentes, en función de las propiedades particulares, por ejemplo la capacidad para unirse a una proteína. El principio general del método consiste en incubar una molécula diana particular con una mezcla de secuencias diferentes (ARN, ADN monocatenario o bicatenario). La fracción capaz de unirse a la molécula diana se aísla del resto de los ácidos nucleicos mediante columna de cromatografía, inmunoprecipitación o cualquier otra técnica de purificación apropiada. A continuación, la fracción enriquecida se amplifica mediante PCR o RT-PCR y después se utiliza para un nuevo ciclo de selección. La repetición de los ciclos de selección y de amplificación permite enriquecer la mezcla inicial en oligonucleótidos funcionales, denominados asimismo "aptámeros". Cuanto más aumenta el número de ciclos de selección y de amplificación, más aumenta la cantidad de aptámeros, hasta dominar en la población de oligonucleótidos (para revisión sobre el método SELEX, véase Klug *et al.*, 1994).

40 Dos métodos SELEX, descritos a continuación, han sido realizados por los inventores.

##### **a1) Origen de la transposasa**

45 Se ha utilizado una proteína recombinante que acopla las cualidades de fijación de las ITR por la transposasa (Tnp) y las cualidades de fijación de la maltosa por la proteína de unión a la maltosa (MBP: Maltose Binding Protein). Esta proteína recombinante, producida en bacteria y nombrada MPB-Tnp, se une a las ITR gracias al dominio de unión específico de la transposasa localizado en N-terminal y la MBP permite purificar los complejos ITR/transposasa sobre una columna de maltosa.

##### **a2) Naturaleza de las ITR de secuencias degeneradas**

50 Para realizar el SELEX, se ha sintetizado una mezcla de oligonucleótidos de 79 bases por la compañía MWG Biotech. La estructura general de estos oligonucleótidos comprende la secuencia de una ITR degenerado de 29 bases, rodeada en cada uno de sus extremos por la secuencia de los cebadores R y F de 25 bases cada uno. Estos cebadores de secuencias respectivas 5'-CAGGTCAGTTTCAGCGGATCCTGTCG-3' (SEC ID n°35) y 5'-GAGGCGAATTCAGTGCAACTGCAGC-3' (SEC ID n°36) han permitido, en las etapas posteriores de SELEX, amplificar mediante PCR las secuencias de las ITR seleccionadas.

Se han sintetizado dos mezclas distintas de oligonucleótidos: una cuya ITR de 29 bases está degenerada sobre 14 posiciones (ITR14), y una cuya ITR está degenerada sobre 21 posiciones (ITR21). Las posiciones conservadas al 100%, en todos los elementos de la sub-familia *mariner*, se han mantenido en las ITR14 e ITR21, mientras que las posiciones conservadas al 60/80% se han mantenido sólo en las ITR14. Las ITR14 estaban representadas mediante  $2,7 \times 10^8$  secuencias, las ITR21 mediante  $4,4 \times 10^{12}$  secuencias.

Con el fin de validar el método, la ITR3' de *Mos-1* ha sido utilizado como control.

Cada una de las ITR14 e ITR21 se ha vuelto bicaténaria mediante PCR antes del primer ciclo de SELEX, puesto que la transposasa se fija sobre un ADN bicatenario.

### a3) Principio del método SELEX 1

Este método se ilustra en la figura 8. Utiliza un grupo de matrices ADN monocatenarios (sb) formadas por una ITR de 29 nucleótidos (nt) degenerada sobre 12 ó 21 posiciones y por dos cebadores R y F de 25 nucleótidos que rodean los extremos de la ITR14 y de la ITR21 (a). Las matrices se vuelven bicaténarias (db) mediante PCR (b). Después se marcan radiactivamente (c) y se incuban en disolución con la resina y la proteína de fusión MBP-Tnp (anotada Tnp para simplificar la figura) o la MBP (d). La reacción de interacción se desarrolla durante 24 horas a 4°C. Después del lavado de la columna, los complejos ADN/proteína se purifican gracias a una disolución de maltosa (e). Los eluatos recuperados se denominan eluatos Tnp/ITR cuando las matrices han sido incubadas con MBP-Tnp y eluatos MBP cuando las matrices han sido incubadas con MBP. Para seguir la selección después de cada ciclo de SELEX, se deposita un alícuota de cada eluato sobre una membrana de nylon y después se cuenta (f). Las matrices seleccionadas en cada ciclo de SELEX son amplificadas mediante PCR (g). Los productos amplificados son controlados a continuación sobre gel de agarosa. Los fragmentos de interés son purificados (h) y después utilizados para un nuevo ciclo de selección.

#### \* Etapa de fijación ADN/proteína:

Con el fin de seguir la selección de las ITR a cada ciclo de SELEX, las secuencias nucleotídicas han sido radiomarcadas, o bien mediante la T4 polinucleótido quinasa, o bien mediante PCR. La selección de las secuencias dianas ha sido efectuada mediante incubación en disolución de la MBP-Tnp, de las ITR14 o de las ITR21 radiomarcadas, y de la resina de maltosa.

Paralelamente, se han realizado dos experimentos de control. Un control negativo ha permitido asegurar la especificidad de la interacción ADN/proteína, incubando la MBP que no tiene una afinidad particular para el ADN, con las ITR14 o las ITR21. El control positivo consistió en incubar la ITR3' con la MBP-Tnp por un lado y la MBP por otro lado.

#### \* Etapa de lavado y elución:

Después de la fijación de la proteína sobre su secuencia diana, se ha realizado una etapa de lavado con el fin de eliminar todos los oligonucleótidos no unidos sin disociar los complejos. La elución de los complejos considerados ha sido realizada saturando la resina mediante maltosa. Se han obtenido dos tipos de eluatos. El eluato Tnp/ITR ha sido producido mediante elución de la serie de experimentos que incuban los oligonucleótidos diana con la proteína recombinante MBP-Tnp. El eluato MBP ha sido producido mediante elución de la columna poniendo en interacción las mismas ITR diana con la MBP.

#### \* Etapa de amplificación de las secuencias seleccionadas:

La amplificación de las ITR seleccionadas ha sido efectuada directamente sobre el eluato Tnp/ITR y el eluato MBP gracias a la presencia de los cebadores R y F que encuadran las secuencias ITR (SEC ID nº 35 y 36). Si la selección fuera efectiva, se debería encontrar una señal PCR específica (de 79 pb) para la amplificación de las matrices contenidas en el eluato Tnp/ITR, pero no para las matrices del eluato MBP (puesto que la MBP no tiene afinidad para las ITR). Este fragmento positivo ha sido eluido del gel de agarosa y radiomarcado mediante la T4 polinucleótido quinasa. Para algunos ciclos de PCR, el marcado ha sido efectuado directamente durante la etapa de amplificación.

Este amplímero ha sido utilizado a continuación como diana enriquecida con secuencia afín de la Tnp en el ciclo de SELEX siguiente.

### a4) Principio del método SELEX 2

Unas investigaciones realizadas por los inventores han mostrado que la ITR3' y la ITR5' tienen la misma constante de disociación pero que su capacidad para fijar la transposasa es diferente. La cantidad de proteína activa en presencia de la ITR5' es 10 veces más baja que la observada en presencia de la ITR3'. Esto indica que la ITR3' actúa como un activador de la capacidad de la transposasa para unir una ITR. Dos informaciones están por lo tanto contenidas en una ITR. La primera tiene un efecto sobre la activación de la proteína (impacto sobre B<sub>max</sub>), la segunda modula la afinidad de la transposasa para la ITR (impacto sobre K<sub>d</sub>). Con el fin de tener en cuenta estos datos, el método SELEX 2 ha sido realizado por los inventores. El principio de este método es idéntico al del método SELEX 1. Sin embargo, las

matrices ADN han sido incubadas con la proteína durante cinco minutos a 4°C antes de realizar el enganche sobre la columna de maltosa. Este método debía permitir seleccionar así unas ITR que tienen la capacidad de activar y de unirse a la transposasa.

#### a5) Protocolos

5 Como regla general, los procedimientos experimentales realizados por los inventores se basan en unas técnicas clásicas bien conocidas por el experto en la materia (Ausubel *et al.*, 1994; Sambrook y Russel, 2001).

#### (i) Preparación de las matrices

##### \* Síntesis de la hebra complementaria

10 La transposasa se une sobre el ADN sólo en forma bicatenaria. La síntesis de la hebra complementaria de los oligonucleótidos dianas, ITR14 y ITR21, ha sido realizada mediante extensión de cebador. La reacción ha sido efectuada de forma idéntica para el método SELEX 1 y 2.

##### \* Purificación de los fragmentos de ADN

15 Los productos PCR han sido analizados sobre gel de agarosa Nusieve al 3% (FMC) en tampón TAE 1X. Los fragmentos de ADN han sido purificados a continuación a partir del gel de agarosa con el fin de eliminar cualquier traza de concatémeros.

##### \* Marcado radioactivo de las ITR14 e ITR21

20 El marcado radioactivo de las secuencias ha permitido seguir la evolución de la selección, en cada ciclo SELEX. El marcado con [ $\gamma$ <sup>32</sup>P]ATP (actividad específica superior a 4500 Ci/mmol) ha sido efectuado con la polinucleótido quinasa del fago T4 (PNK), y después se ha realizado el marcado con [ $\alpha$ <sup>32</sup>P]ATP (actividad específica superior a 3000 Ci/mmol) mediante PCR, en presencia de los cebadores R y F (SEC ID nº 35 y 36).

#### (ii) Selección de las dianas por la transposasa

##### \* Etapa de fijación ADN/proteína

25 La resina de maltosa (New-England Biolabs) ha sido equilibrada en el tampón 1 (Tris pH 9 20 mM, NaCl 50 mM, DTT 1 mM). La incubación en disolución ha sido efectuada en un volumen final de 1 ml de tampón 1, con 200 µl de resina a los que se ha añadido sucesivamente 50 µg de proteína MBP-Tnp o MBP; 200 ng de ITR14, ITR21 o ITR3'; 2 µg de ITR5' como competidor; 5 mM de MgCl<sub>2</sub> y 2 µg de ADN de esperma de salmón. La reacción de interacción del método SELEX 1 ha sido mantenida durante 24 horas a 4°C bajo agitación constante (300 rpm). En el método SELEX 2, las matrices de ADN y la proteína MBP-Tnp o MBP han sido incubadas durante 5 minutos a 4°C antes de ser pasadas sobre la columna de maltosa, siendo la reacción de interacción mantenida durante sólo 1 hora a 4°C,

##### \* Etapa de lavado de la resina y elución de los complejos

30 Al final de la incubación, la resina se lavó y se eluyeron los complejos proteína/ITR. Dos eluciones sucesivas han sido realizadas con el fin de recuperar la totalidad de los complejos.

#### (iii) Etapa de amplificación de las secuencias seleccionadas

35 Después de cada ciclo de SELEX, las matrices seleccionadas contenidas en los eluatos Tnp/ITR y MBP han sido amplificadas antes de servir para el ciclo de SELEX siguiente.

La reacción de PCR comprendía de 15 a 30 ciclos, típicamente 20 ciclos (15 ciclos en caso de amplificación de fragmentos parásitos).

40 El medio de reacción utilizado para la amplificación de las ITR14 y de las ITR21 contenía, en un volumen final de 50 µl, la matriz: 10 µl de eluato Tnp/ITR o 10 µl del eluato MBP, en presencia de 5 µl de tampón 10X; 200 µM de dNTP; 2,5 mM de MgCl<sub>2</sub>; 1 µM de cebadores R y F, y 5 unidades de Taq polimerasa. Sobre gel de agarosa, la amplificación a partir de los eluatos Tnp/ITR debía dar una banda de 79 pb y de 300 pb para ITR3' (control positivo). Los productos PCR han sido purificados a partir del gel de agarosa, y después marcados utilizando la PNK, y utilizados para un nuevo ciclo de SELEX.

45 Al quinto ciclo de SELEX, la PCR ha sido realizada en presencia de material radioactivo para marcado con [ $\alpha$ <sup>32</sup>P]ATP.

#### b) Clonación y secuenciación de las secuencias seleccionadas

Los productos de PCR purificados del ciclo número 7 del método SELEX 1 y del ciclo 8 del método SELEX 2 han sido clonados en el plásmido pGEMT-Easy (kit pGEMT-Easy Vector system, Promega) en las condiciones

preconizadas por el proveedor. Las ligaduras en el pGEMT-easy han sido realizadas con los fragmentos ITR14 del método SELEX 1, ITR14 del método SELEX 2, ITR21 del método SELEX 1 e ITR21 del método SELEX 2. Las ligaciones han sido utilizadas para transformar unas bacterias competentes DH5 $\alpha$ . El ADN plasmídico de 20 clones recombinantes de cada una de las ligaduras ha sido analizado mediante secuenciación monocatenaria. Una alineación de las secuencias ha sido realizada gracias al programa CLUSTALW accesible en la página WEB [www.infobiogen.fr](http://www.infobiogen.fr).

#### c) Cribado rápido de las secuencias clonadas

Cada una de las ITR potenciales ha sido ensayada por su capacidad para fijarse en la transposasa por retraso sobre gel. Las ITR han sido así incubadas en presencia de MBP-Tnp que debe provocar un retraso de migración si la transposasa está fijada en la ITR. En un primer tiempo, un cribado rápido de las ITR ha sido realizado incubando los ADN radiomarcados por PCR, sin purificación, en presencia de la proteína. En un segundo tiempo, se ha efectuado un retraso sobre gel con el fin de eliminar el ruido de fondo debido a la amplificación de secuencias artefactuales.

##### c1) *Marcado radioactivo de las secuencias clonadas*

###### (i) *Marcado mediante PUR*

Se han seleccionado 80 ITR. Estas ITR han sido marcadas mediante PCR, en presencia de [ $\alpha$ -<sup>32</sup>P]ATP y unos cebadores universales pU y pREV, a partir de las minipreparaciones de ADN plasmídico.

El tamaño esperado de los fragmentos amplificados era de 79 pb.

###### (ii) *Marcado mediante llenado con Klenow*

Las ITR positivas durante el cribado rápido han sido purificadas sobre gel de agarosa después de la digestión por la enzima *EcoRI* que permite eliminar los cebadores R y F. La ITR3' ha sido purificada después de la digestión del plásmido pBluescript-ITR3' por las enzimas *EcoRI* y *BamHI*. Estos fragmentos purificados han sido radiomarcados con la ayuda de [ $\alpha$ -<sup>32</sup>P]ATP mediante llenado de sitio gracias a la Klenow.

##### c2) *Formación de los complejos ADN/proteína*

###### (i) *Cribado rápido*

Los complejos ITR/proteína han sido formados con las secuencias marcadas mediante PCR al final del último ciclo de la reacción SELEX, sin purificación previa, con el fin de realizar un cribado rápido. Estas secuencias contenían una ITR encuadrada por los cebadores R y F. La reacción de interacción contenía, en un volumen final de 20  $\mu$ l: 40  $\mu$ g de proteína MBP-Tnp o MBP; 1  $\mu$ l de la reacción de PCR radioactiva; 1  $\mu$ g de ADN de esperma de salmón; 2  $\mu$ l de glicerol al 50%; 5 mM de MgCl<sub>2</sub> y 0,5  $\mu$ M de pRev. Las sondas libres han sido preparadas con 1  $\mu$ l de PCR radioactiva, 2  $\mu$ l de glicerol al 50% y 17  $\mu$ l de tampón. Las reacciones de interacción han sido mantenidas durante 15 minutos a 4°C antes de ser analizadas sobre gel de poliacrilamida.

###### (ii) *Retraso sobre gel sobre las sondas purificadas*

Las secuencias que provocan un retraso de migración después de la incubación con la Tnp (ITR positivas) han sido sometidas a un nuevo retraso sobre gel con un fragmento de ADN purificado. Los complejos ITR/proteína han sido formados en un volumen final de 20  $\mu$ l que contiene: 40  $\mu$ g de proteína MBP-Tnp, 1 nM de sonda ITR, 1  $\mu$ g de ADN de esperma de salmón, 2  $\mu$ l de glicerol al 50%, 5 mM de MgCl<sub>2</sub> y 0,5  $\mu$ M de pRev. Las ITR solas han sido utilizadas a una concentración final de 1 nM en una mezcla que contiene 2  $\mu$ l de glicerol al 50% y 17  $\mu$ l de tampón. Las reacciones de interacción han sido mantenidas durante 15 minutos a 4°C antes de ser analizadas sobre gel de poliacrilamida.

#### d) Ensayos de competición

El principio de estos ensayos se ilustra en la figura 9. Este ensayo permite demostrar las capacidades de ITRSelex para desplazar la fijación de la transposasa de ITR3' radiomarcada. Cuanto más importante es el desplazamiento, más "mejorada" estará la secuencia ITRSelex con respecto a ITR3'. En la práctica, la reacción de fijación de la transposasa se realiza en 20  $\mu$ l que contiene 10 mM de tampón Tris pH 9, 0,5 mM de DDT, MgCl<sub>2</sub> 5 mM, 5% (vol/vol) de glicerol, 1  $\mu$ g de ADN de esperma de arenque y 100 ng de BSA, en presencia de 15 nM de ITR3' radiomarcada y de ITRSelex no radiomarcada. Las concentraciones de ITRSelex no radiomarcada ensayadas eran de 0 nM, 15 nM, 75 nM, 150 nM, 300 nM, 750 nM, y 1500 nM.

#### e) Construcción de los plásmidos pBC3TSelex

Con el fin de analizar el comportamiento de las 8 ITRSelex *in vivo* en bacteria, una serie de ocho plásmidos ha sido construida a partir del plásmido pBC3T5. El plásmido pBC3T5 contiene el ORF de *Tet* (gen de resistencia a la tetraciclina) sin promotor, rodeado por las ITR3' y 5' de *Mos-1*. El gen *tet* (clonado entre los sitios de restricción *XbaI* y *HindIII*) está en orientación inversa del gen de resistencia al cloramfenicol y del gen que codifica para la proteína LacZ. La ITR3' está delimitada por el sitio de restricción *KpnI* del plásmido pBCKS+ en 5' y por el sitio de restricción *SalI* en 3'.

La ITR5' está delimitada por el sitio de restricción *SacI* del plásmido pBCKS+ en 5' y por el sitio de restricción *NotI* en 3'. La ITR5' del plásmido pBC3T5 ha sido sustituida con las ITRSelex después de una doble digestión por *NotI* y *SacI*, generando los plásmidos pBC3T5Selex. Las ITRSelex han sido sintetizadas en forma de oligonucleótidos monocatenarios por la compañía MWG Biotech (Alemania). La formación de ITRSelex bicatenaria ha sido realizada mediante hibridación de manera que se generan unos semi-sitios cohesivos *NotI* y *SacI*. Los oligonucleótidos cohesivos han sido designados de tal manera que un dinucleótido TA que rodea la ITRSelex en 5' esté dispuesto en el exterior del pseudo-elemento. Estos oligonucleótidos bicatenarios fosforilados han sido empalmados por el ADN ligasa T4 al vector digerido por las dos enzimas, con el fin de generar los plásmidos pBC3TSelex. Estos plásmidos son denominados a continuación pBC3Ts o pBC3TSelex, seguido del número de la ITRSelex.

#### 10 f) Ensayos de transposición

Una descripción detallada del protocolo experimental se proporciona en la parte I, párrafo I-B anterior.

Estos ensayos han sido realizados sobre unas bacterias *E. coli* JM109 cotransformadas por 10 ng de plásmido donante de transposasa (pKK-Tnp p pKK) y 10 ng de plásmido donante de pseudo-transposón (pBC3TSelex). Estas bacterias han sido seleccionadas sobre un medio que contiene ampicilina y cloramfenicol. Las condiciones de realización de estos ensayos ("Método I") se indican en la figura 10.

#### 15 **I-2- Resultados**

##### a) Cribado de las secuencias de ITR candidatas obtenidas por SELEX

El método SELEX realizado por los inventores utiliza una mezcla de oligonucleótidos de 79 pb formados por una ITR de 29 pb degenerada sobre 14 ó 21 posiciones (ITR14 e ITR21), y rodeada en sus extremos por unos cebadores R y F de 25 pb (SEC ID nº 35 y 36). Utiliza asimismo una proteína recombinante que fusiona la transposasa de *Mos-1* (denominada Tnp) y la MBP.

Se han desarrollado dos métodos SELEX. El principio general de estos dos métodos sigue siendo idéntico. En el método SELEX 1, los oligonucleótidos, la proteína y la resina de maltosa son incubados al mismo tiempo. En el método SELEX 2, las matrices son incubadas con la proteína durante 5 minutos a 4°C antes de ser puestas en contacto con la resina de maltosa.

Se han clonado las secuencias ITR14 e ITR21 seleccionadas en el ciclo 7 del método SELEX 1 y en el ciclo 8 del método SELEX 2. Para cada método, 20 clones que corresponden a las ITR14 y 20 clones que corresponden a las ITR21 han sido aislados y después secuenciados. Estas 80 secuencias han sido analizadas en función de su naturaleza, es decir que si se trata de una ITR14 degenerada sobre 14 posiciones o de una ITR21 degenerada sobre 21 posiciones, y el método del que proceden (método SELEX 1 o método SELEX 2). Los resultados han mostrado que el método utilizado podía tener una incidencia sobre el tipo de selección efectuada (datos no mostrados).

Las 80 secuencias han sido ensayadas en retraso sobre gel para controlar su capacidad de unión con la transposasa. Los resultados han mostrado que los clones ITR 1, 6, 9, 40, 46, 49, 60 y 69 (ITRSelex; figura 11; SEC ID nº 38 a 45) son capaces de formar un complejo con la proteína. Estos clones positivos han sido ensayados en retraso sobre gel con el fin de saber si son reconocidos por la transposasa tal como unas ITR o como unas dianas. Los resultados han mostrado que los clones 1, 40, 46, 49 y 69 son unas ITR (figura 11 y datos no mostrados). Los resultados no han permitido concluir para los clones 6 y 9 (figura 11).

##### b) Ensayo de competición

Los resultados han mostrado que, en las 8 ITR consideradas después de los experimentos de retraso sobre gel anteriores, sólo las ITR40 y 46 son capaces de inhibir la fijación de la transposasa sobre ITR3' radiomarcada.

ITR40 5'-TCAGGTGTACAAGTATGTAATGTCGTTA-3' (SEC ID nº39);

ITR46 5'- TCAGGTGTACAAGTATGAGATGTCGTTT-3' (SEC ID nº38).

Tal como se ilustra en la figura 12, sólo las ITR40 y 46 son capaces de entrar en competición con la ITR3' y desplazar la fijación de la transposasa de la ITR3' radiomarcada.

##### 45 c) Ensayos de transposición

Aunque las ITR40 y 46 aparecen como los mejores candidatos según los ensayos de competición, el comportamiento de las 8 ITR ha sido evaluado en ensayo de transposición *in vivo* en bacteria, con el fin de verificar si, y en qué medida, estas ITR tienen la capacidad de mediar el conjunto de la transposición.

Los resultados obtenidos se presentan en las figuras 13 A y B. Se observa que sólo las ITR40 y 46 permiten efectivamente mejorar la transposición en las condiciones ensayadas. Tratándose de los controles, en las condiciones experimentales del método I, la eficacia de la transposición de pBC3T3 ha aumentado en un factor 10 con respecto a la obtenida con el plásmido pBC3T5 (figura 13A).

Por último, tal como lo muestra la figura 13B, los pseudo-transposones 3T40 y 3T46 son hiperactivos.

## II- Las UTR (Untranslate Terminal Repeat)

5 La configuración mínima de los ácidos nucleicos para una transposición óptima del elemento *Mos-1* parece incluir, además de las ITR, una parte por lo menos de UTR. En efecto, *in vitro*, un marcador de resistencia únicamente rodeado por las ITR5' y 3' de *Mos-1* no transpone, mientras que la adición de los 38 primeros pb de UTR5' y de los 5 primeros pb de UTR3' a las secuencias de las ITR respectivas basta para restablecer la actividad salvaje (Tosi *et al.*, 2000).

10 La necesidad de la presencia de las UTR en la proximidad de las ITR ha sido evaluada mediante la construcción de un cierto número de configuraciones posibles: UTR en 5' o 3', UTR 5' asociada a una ITR3', o a la inversa.

En resumen, los resultados obtenidos han mostrado que la presencia de UTR favorece la transposición (en particular una mejora en un factor por lo menos igual a aproximadamente 5).

### II-1- Material y métodos

#### a) Construcción de las configuraciones ITR+UTR

##### 15 a1) Plásmidos

20 Los plásmidos ITR/UTR han sido construidos todos a partir del plásmido pBC3T5 según el mismo modo de realización. La ITR3' ha sido sustituida después de la doble digestión por las enzimas *KpnI* y *SacI*. La ITR5' ha sido sustituida mediante doble digestión *NotI* y *SacI*. Las diferentes secuencias ITR/UTR 33 y 55 han sido sintetizadas y clonadas en pCR4-TOPO (Invitrogen) por la compañía ATG biosynthetics (Alemania). La secuencia ITR/UTR35 - MCS - UTR/ITR35, es decir ITR3'/UTR5' - sitio de clonación múltiple (MCS: "Multiple Cloning Site") - UTR3'/ITR5' ha sido sintetizada y clonada en pCR-Script AmpSK(+) (Stratagene) por la compañía Intelechon (Alemania). Esta secuencia ha sido introducida en pBC mediante doble digestión *KpnI* y *SacI*. Estas secuencias han sido designadas de tal manera que un dinucleótido TA que rodea ITR/UTR en 5' esté dispuesto en el exterior del pseudo-elemento. Una quincena de construcciones han sido realizadas y evaluadas en los ensayos de transposición en bacteria según el método 2. Los resultados se presentan para los plásmidos siguientes con la anotación pBC ITR/UTR-T-UTR/ITR: pBC33T33, pBC33T55 y pBC35T35.

30 El plásmido donante de transposasa pKKTnp es un derivado del plásmido pKK233-2 (Clontech; Amp<sup>r</sup>) en el que se ha clonado, en el sitio *NcoI*, el ORF de la transposasa *Mos-1* anotado Tnp. Su expresión está bajo el control del promotor inducible con el IPTG *P<sub>trc</sub>* (este promotor posee sin embargo una actividad de transcripción de base en ausencia de inductor; datos no mostrados). Este plásmido pKK233-2 será simplemente anotado pKK en la continuación de la descripción cuando el ORF de la transposasa está ausente.

##### a2) *Transformación de las cepas E. coli JM109 para ensayos de transposición*

35 Las bacterias JM109 competentes han sido cotransformadas con un plásmido donante de transposón cuyos plásmidos pBC33T33, pBC3T5, pBC3T3, pBC33T55, pBC35T35, pBC3T33, y el plásmido donante de Tnp pKKTnp. Unas cepas de control han sido cotransformadas con los mismos plásmidos donantes de transposón y el plásmido de control pKK.

#### b) Ensayos de transposición

Una descripción detallada del protocolo experimental se proporciona en la parte I, párrafo I-B anterior.

40 La quincena de configuraciones ha sido ensayada en transposición *in vivo* en bacterias según el método II, ilustrado en la figura 14.

### II-2- Resultados

45 La eficacia de transposición ha sido calculada dividiendo el número de clones Tet<sup>R</sup> aparecidos por el número de bacterias analizadas en presencia del plásmido donante de transposasa pKK-Tnp, al que se le ha retirado el ruido de fondo del experimento obtenido en presencia del plásmido de control pKK. Los resultados más significativos han sido obtenidos para las construcciones 3T33, 33T33 y 35T35, en comparación con las construcciones de control 3T3, 3T5 y 33T55. La eficacia de la transposición ha aumentado en un factor de 5 y 20 para las construcciones pBC33T33 y pBC3T33, con respecto a la construcción de control pBC3T3. Estos resultados muestran que la presencia de la secuencia UTR es extremadamente importante para la reacción de transposición puesto que se observa una eficacia de transposición, con la construcción pBC33T33, 300 veces superior a la obtenida para el plásmido pBC3T5 como para el plásmido pBC33T55. Los mejores resultados han sido obtenidos para la construcción pBC35T35, que proporciona un aumento de la eficacia de la transposición en un factor de 54000 con respecto a pBC3T5 y en un factor de 1000 con respecto a pBC3T3.

Según la figura 15A, las construcciones interesantes son pBC35T35, pBC3T33 y pBC33T33.

La figura 15B muestra que los pseudo-transposones 3T33, 33T33 y 35T35 son hiperactivos.

5 Los inventores han ensayado además las construcciones 53T35, 53T33, 35T33, 55T35, 53T55, 55T55, 5T35, 5T33, 3T55, 3T53. De esta forma, han podido observar que estas construcciones no provocaban ninguna hiperactividad; tenían una eficacia o bien equivalente a la de 3T5 o de 3T3, o bien más baja o nula (datos no mostrados).

**PARTE III: SISTEMAS DE TRANSPOSICIÓN RECOMBINANTES HIPERACTIVOS QUE COMPREDEN UNA TRANSPOSASA MOS-1 MUTANTE HIPERACTIVA (Tnp) Y UN PSEUDO-TRANSPOSÓN *Mos-1* RECOMBINANTE HIPERACTIVO**

10 Para evaluar la eficacia de la transposición de sistemas que asocian una transposasa hiperactiva y un pseudo-transposón asimismo hiperactivo, se han ensayado diversas combinaciones en el ensayo de transposición en bacteria descrito en la parte I, párrafo I-B anterior.

En este ensayo, el plásmido pBC3T3 ha sido sustituido por los pseudo-transposones hiperactivos pBC3T33, pBC3T40, y pBC3T46. El plásmido pKK-Tnp salvaje ha sido sustituido por unos vectores de tipo pKK que expresan cada uno una transposasa mutante hiperactiva particular.

15 La tabla 3 siguiente proporciona los resultados obtenidos combinando unas transposasas hiperactivas (FETY, FQETY, FTY, FT, TY, ET, FQ, FQET, QY) y unos pseudo-transposones 3T33, 3T40, 3T46, 33T55 (siendo este último utilizado como control).

Tabla 3

	Factor de amplificación con respecto a 3T3JWT		Factor de amplificación con respecto a 3T3/WT		
	T0	T5	T0	T5	
3T40/			33T55/		
FETY	4	0,1	FETY	1,1	1,5
FQETY	2,5	0,8	FQETY	0,4	1,6
FTY	3	2,4	FTY	2,1	4,6
FT	0	2	FT	0	0,1
TY	1,4	36,2	TY	0	8,4
ET	4,1	8,9	ET	0	7,1
FQ	0,7	2,5	FQ	0	0,1
FQET	1,8	3,7	FQET	0	0,2
QY	0,3	1,3	QY	0	0
3T46/			3T33/		
FETY	6,5	2	FETY	24,6	51,4
FQETY	3,6	2,3	FQETY	22,4	59,7
FTY	6,7	19,4	FTY	44	56
FT	0,1	3,7	FT	42,9	60
TY	2,3	61,8	TY	68,4	66,5
ET	15,6	6,7	ET	37,6	218,5
FQ	1,4	3,3	FQ	4,3	109,3
FQET	3,3	12,3	FQET	36	63,5
QY	1,7	1,7	QY	5,2	28

WT: transposasa Mos-1 salvaje

F: transposasa Mos-1 con la mutación F53Y

E: transposasa Mos-1 con la mutación E137K

T: transposasa Mos-1 con la mutación T216A

Y: transposasa Mos-1 con la mutación Y237C

Q: transposasa Mos-1 con la mutación Q91 R

De manera completamente sorprendente e imprevisible, los resultados obtenidos sobre las combinaciones ensayadas revelan que todas las combinaciones de una transposasa Mos-1 hiperactiva con un pseudo-transposón *Mos-1* hiperactivo no son necesariamente hiperactivas.

5 Los resultados así obtenidos permiten seleccionar, como combinaciones interesantes para los fines de la presente invención, las asociaciones:

- pseudo-transposón 3T40 + transposasa TY,
- pseudo-transposón 3T46 + transposasa TY o ET o FTY,
- pseudo-transposón 3T33 + transposasa TY o ET o FQ o FQET.

10 En las condiciones de los ensayos de transposición en bacteria indicados en este caso, la combinación pseudo-transposón 3T33 + transposasa ET es la más interesante puesto que es respectivamente más eficaz que la asociación pseudo-transposón 3T3 + transposasa WT (200 veces), pseudo-transposón 3T3 + transposasa FETY (3,5 veces) y pseudo-transposón 3T33 + transposasa WT (10 veces).

15 Un análisis bioquímico en gel de retraso ha sido efectuado siguiendo los procedimientos descritos anteriormente [en particular en la solicitud internacional WO 2004/078981 publicada el 16 de septiembre de 2004 y en Augé-Gouillou *et al.*, (2001 b)] con el fin de determinar la estabilidad de los complejos transposasa + ITR o transposasa + ITR/UTR.

Estos trabajos han permitido mostrar que los fragmentos de ADN que asocian ITR3' a UTR3' e ITR3' a UTR5' son mucho más estables (4 veces) que los formados con las ITR solas u otras combinaciones (figura 16). Esta mayor estabilidad de los complejos podría ser el origen de la hiperactividad observada.

20

**REFERENCIAS**

- Lampe DJ. *et al.* (1996) EMBO J. 15: 5470-5479
- Plasterk RHA. *et al.* (1999) Trends in genetics 15: 326-332
- Renault S. *et al.* (1997) Virologie 1: 133-144
- 5 Jacobson y Hartl (1985) Genetics 111: 57-65
- Craig *et al.* (2002) Mobile ADN II. ASM Press. Washington. USA
- Jeong *et al.* (2002) PNAS 99: 1076-1081
- Martienssen y Colot (2001) Science 293: 1070-1074
- Ketting *et al.* (1999) Cell 99: 133-141
- 10 Tabara *et al.* (1999) Cell 99: 123-132
- Ausubel *et al.* (1994) En Janssen, K. (Ed) Current Protocols in Molecular Biology. J. Wiley & Sons, Inc. Massachusetts General Hospital. Harvard Medical School
- Augé-Gouillou *et al.* (2001) Mol. Genet. Genomics 265: 58-65
- Augé-Gouillou *et al.* (2001b) Mol. Genet. Genomics 265: 51-57
- 15 Turner DL *et al.* (1994) Gènes Dev. 8: 1434-1447
- Sambrook y Russel (2001) Molecular Cloning: a laboratory manual (3<sup>a</sup> Ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Tosi *et al.* (2000) Nucleic Acids Res. 28: 784-790
- Ellington *et al.* (1990) Nature 346: 818-822
- 20 Tuerk *et al.* (1990) Science 249: 505-510
- Klug *et al.* (1994) Mol. Biol. Rep. 20: 97-107
- Zhang *et al.* (2001) Nucleic Acids Res. 29: 3566-3575

**LISTADO DE SECUENCIAS**

- <110> Centre National de la Recherche Scientifique (CNRS)  
Universidad François Rabelais de Tours
- 5 <120> SISTEMAS DE TRANSPOSICIÓN RECOMBINANTES HIPERACTIVOS DERIVADOS DEL TRANSPOSÓN Mos-1
- <130> 351306 - D23688
- <150> FR0604285
- <151> 15/05/2006
- <160> 45
- 10 <170> PatentIn versión 3.3
- <210> 1
- <211> 1038
- <212> ADN
- <213> *drosophila mauritiana*
- 15 <220>
- <223> Secuencia del gen que codifica la transposasa del transposón Mos-1
- <400> 1

```

atgtcgagtt tcgtgccgaa taaagagcaa acgcgacag tattaatfff ctgttttcat    60
ttgaagaaaa cagctgcgga atcgcaccga atgcttgttg aagcctttgg cgaacaagta    120
ccaactgtga aaacgtgtga acgggtggtt caacgcttca aaagtgttga ttttgacgtc    180
gacgacaaa agcagcgaaa accgcacaaa aggtacgaag acgcccgaact gcaagcatta    240
ttgatgaa agcagtttca aacgcaaaaa caactcgac agcagttgga agtaagtcaa    300
caagcagttt ccaatcgctt gcgagagatg ggaaagattc agaaggtcgg tagatgggtg    360
ccacatgagt tgaacgagag gcagatggag aggcgcaaaa acacatgcga aatfttgctt    420
tcacgataca aaaggaagtc gtttttgcac cgtatcgtta ctggcgatga aaaatggatc    480
ttttttgtta atcctaaacg taaaaagtca tacgttgatc ctggacaacc gtccacatcg    540
actgctcgac cgaatcgctt tggcaagaag acgatgctct gtgtttggtg ggatcagagc    600
ggtgtcattt actatgagct cttgaaaccc ggcgaaacgg tgaatacggc acgctaccaa    660
caacaattga tcaatttgaa ccgtgcgctt cagagaaaac gaccggaata tcaaaaaaga    720
caacacaggg tcatttttct ccatgacaac gctccatcac atacggcaag agcggttcgc    780
gacacgttgg aaacactcaa ttgggaagtg ctcccgatg cggcttactc accagacctg    840
gccccatccg attaccacct attcgtctcg atgggacacg cactcgttga gcagcgttc    900
gattcttacg aaagtgtgaa aaaatggctc gatgaatggt tcgcccgaac agacgatgag    960
ttctactggc gtggaatcca caaattgccc gagagatggg aaaaatgtgt agctagcgac   1020
ggcaataact ttgaataa                                     1038
    
```

- <210> 2
  - 20 <211> 345
  - <212> PRT
  - <213> *drosophila mauritiana*
  - <400> 2
- ```

Met Ser Ser Phe Val Pro Asn Lys Glu Gln Thr Arg Thr Val Leu Ile
1          5          10          15

Phe Cys Phe His Leu Lys Lys Thr Ala Ala Glu Ser His Arg Met Leu
20          25          30

Val Glu Ala Phe Gly Glu Gln Val Pro Thr Val Lys Lys Cys Glu Arg
35          40          45
    
```

ES 2 374 418 T3

Trp Phe Gln Arg Phe Lys Ser Gly Asp Phe Asp Val Asp Asp Lys Glu  
 50 55 60

His Gly Lys Pro Pro Lys Arg Tyr Glu Asp Ala Glu Leu Gln Ala Leu  
 65 70 75 80

Leu Asp Glu Asp Asp Ala Gln Thr Gln Lys Gln Leu Ala Glu Gln Leu  
 85 90 95

Glu Val Ser Gln Gln Ala Val Ser Asn Arg Leu Arg Glu Met Gly Lys  
 100 105 110

Ile Gln Lys Val Gly Arg Trp Val Pro His Glu Leu Asn Glu Arg Gln  
 115 120 125

Met Glu Arg Arg Lys Asn Thr Cys Glu Ile Leu Leu Ser Arg Tyr Lys  
 130 135 140

Arg Lys Ser Phe Leu His Arg Ile Val Thr Gly Asp Glu Lys Trp Ile  
 145 150 155 160

Phe Phe Val Ser Pro Lys Arg Lys Lys Ser Tyr Val Asp Pro Gly Gln  
 165 170 175

Pro Ala Thr Ser Thr Ala Arg Pro Asn Arg Phe Gly Lys Lys Thr Met  
 180 185 190

Leu Cys Val Trp Trp Asp Gln Ser Gly Val Ile Tyr Tyr Glu Leu Leu  
 195 200 205

Lys Arg Gly Glu Thr Val Asn Thr Ala Arg Tyr Gln Gln Gln Leu Ile  
 210 215 220

Asn Leu Asn Arg Ala Leu Gln Arg Lys Arg Pro Glu Tyr Gln Lys Arg  
 225 230 235 240

Gln His Arg Val Ile Phe Leu His Asp Asn Ala Pro Ser His Thr Ala  
 245 250 255

Arg Ala Val Arg Asp Thr Leu Glu Thr Leu Asn Trp Glu Val Leu Pro  
 260 265 270

His Ala Ala Tyr Ser Pro Asp Leu Ala Pro Ser Asp Tyr His Leu Phe  
 275 280 285

Ala Ser Met Gly His Ala Leu Ala Glu Gln Arg Phe Asp Ser Tyr Glu  
 290 295 300

Ser Val Lys Lys Trp Leu Asp Glu Trp Phe Ala Ala Lys Asp Asp Glu  
 305 310 315 320

Phe Tyr Trp Arg Gly Ile His Lys Leu Pro Glu Arg Trp Glu Lys Cys  
 325 330 335

Val Ala Ser Asp Gly Lys Tyr Leu Glu  
 340 345

<210> 3

<211> 36

<212> ADN

5 <213> Secuencia artificial

<220>

ES 2 374 418 T3

|    |                                         |    |
|----|-----------------------------------------|----|
|    | <223> Cebador directo F53Y              |    |
|    | <400> 3                                 |    |
|    | ggtggttca acgctacaaa agtgggtgatt ttgacg | 36 |
|    | <210> 4                                 |    |
| 5  | <211> 36                                |    |
|    | <212> ADN                               |    |
|    | <213> Secuencia artificial              |    |
|    | <220>                                   |    |
|    | <223> Cebador inverso F53Y              |    |
| 10 | <400> 4                                 |    |
|    | cgtaaaaatc accacttttg tagcgttgaa accacc | 36 |
|    | <210> 5                                 |    |
|    | <211> 28                                |    |
|    | <212> ADN                               |    |
| 15 | <213> Secuencia artificial              |    |
|    | <220>                                   |    |
|    | <223> Cebador directo Q91R              |    |
|    | <400> 5                                 |    |
|    | gctcaaacgc aaaaacgact cgcagagc          | 28 |
| 20 | <210> 6                                 |    |
|    | <211> 28                                |    |
|    | <212> ADN                               |    |
|    | <213> Secuencia artificial              |    |
|    | <220>                                   |    |
| 25 | <223> Cebador inverso Q91R              |    |
|    | <400> 6                                 |    |
|    | gctctgagcgag tcgttttgc gttgagc          | 28 |
|    | <210> 7                                 |    |
|    | <211> 27                                |    |
| 30 | <212> ADN                               |    |
|    | <213> Secuencia artificial              |    |
|    | <220>                                   |    |
|    | <223> Cebador directo L92A              |    |
|    | <400> 7                                 |    |
| 35 | cgcaaaaaca agccgagag cagttgg            | 27 |
|    | <210> 8                                 |    |

|    |                                    |    |
|----|------------------------------------|----|
|    | <211> 27                           |    |
|    | <212> ADN                          |    |
|    | <213> Secuencia artificial         |    |
|    | <220>                              |    |
| 5  | <223> Cebador inverso L92A         |    |
|    | <400> 8                            |    |
|    | ccaactgctc tgcggcttgt tttgcg       | 27 |
|    | <210> 9                            |    |
|    | <211> 27                           |    |
| 10 | <212> ADN                          |    |
|    | <213> Secuencia artificial         |    |
|    | <220>                              |    |
|    | <223> Cebador directo L92R         |    |
|    | <400> 9                            |    |
| 15 | cgcaaaaaca acgcgcagag cagttgg      | 27 |
|    | <210> 10                           |    |
|    | <211> 27                           |    |
|    | <212> ADN                          |    |
|    | <213> Secuencia artificial         |    |
| 20 | <220>                              |    |
|    | <223> Cebador inverso L92R         |    |
|    | <400> 10                           |    |
|    | ccaactgctc tgcgcgttgt tttgcg       | 27 |
|    | <210> 11                           |    |
| 25 | <211> 31                           |    |
|    | <212> ADN                          |    |
|    | <213> Secuencia artificial         |    |
|    | <220>                              |    |
|    | <223> Cebador directo Q100N        |    |
| 30 | <400> 11                           |    |
|    | gcagttggaa gtaagtaacc aagcagtttc c | 31 |
|    | <210>12                            |    |
|    | <211> 31                           |    |
|    | <212> ADN                          |    |
| 35 | <213> Secuencia artificial         |    |
|    | <220>                              |    |

# ES 2 374 418 T3

|    |                                   |    |
|----|-----------------------------------|----|
|    | <223> Cebador inverso Q100N       |    |
|    | <400> 12                          |    |
|    | ggaaactgct tggttactta ctccaactg c | 31 |
|    | <210> 13                          |    |
| 5  | <211> 31                          |    |
|    | <212> ADN                         |    |
|    | <213> Secuencia artificial        |    |
|    | <220>                             |    |
|    | <223> Cebador directo Q100R       |    |
| 10 | <400> 13                          |    |
|    | gcagttggaa gtaagtegac aagcagttc c | 31 |
|    | <210> 14                          |    |
|    | <211> 31                          |    |
|    | <212> ADN                         |    |
| 15 | <213> Secuencia artificial        |    |
|    | <220>                             |    |
|    | <223> Cebador inverso Q100R       |    |
|    | <400> 14                          |    |
|    | ggaaactgct tgcgactta ctccaactg c  | 31 |
| 20 | <210> 15                          |    |
|    | <211> 31                          |    |
|    | <212> ADN                         |    |
|    | <213> Secuencia artificial        |    |
|    | <220>                             |    |
| 25 | <223> Cebador directo Q100E       |    |
|    | <400> 15                          |    |
|    | gcagttggaa gtaagtgaac aagcagttc c | 31 |
|    | <210> 16                          |    |
|    | <211> 31                          |    |
| 30 | <212> ADN                         |    |
|    | <213> Secuencia artificial        |    |
|    | <220>                             |    |
|    | <223> Cebador inverso Q100E       |    |
|    | <400> 16                          |    |
| 35 | ggaaactgct tgtcactta ctccaactg c  | 31 |
|    | <210> 17                          |    |

ES 2 374 418 T3

|    |                                                 |    |
|----|-------------------------------------------------|----|
|    | <211> 44                                        |    |
|    | <212> ADN                                       |    |
|    | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |
| 5  | <223> Cebador directo S104P                     |    |
|    | <400> 17                                        |    |
|    | ggaagtaagt caacaagcag ttccaatcg ctgcgagag atgg  | 44 |
|    | <210> 18                                        |    |
|    | <211> 44                                        |    |
| 10 | <212> ADN                                       |    |
|    | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |
|    | <223> Cebador inverso S104P                     |    |
|    | <400> 18                                        |    |
| 15 | ccatctctcg caagcgattg ggaactgct gttgacttac ttcc | 44 |
|    | <210> 19                                        |    |
|    | <211> 25                                        |    |
|    | <212> ADN                                       |    |
|    | <213> Secuencia artificial                      |    |
| 20 | <220>                                           |    |
|    | <223> Cebador directo N105A                     |    |
|    | <400> 19                                        |    |
|    | caagcagttt ccgcacgctt gcgag                     | 25 |
|    | <210> 20                                        |    |
| 25 | <211> 25                                        |    |
|    | <212> ADN                                       |    |
|    | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |
|    | <223> Cebador inverso N105A                     |    |
| 30 | <400> 20                                        |    |
|    | ctcgcaagcg tgcggaaact gcttg                     | 25 |
|    | <210> 21                                        |    |
|    | <211> 34                                        |    |
|    | <212> ADN                                       |    |
| 35 | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |

# ES 2 374 418 T3

|    |                                                 |    |
|----|-------------------------------------------------|----|
|    | <223> Cebador directo E137K                     |    |
|    | <400> 21                                        |    |
|    | ggcgcaaaaa cacatgcaaa atttgctt cacg             | 34 |
|    | <210> 22                                        |    |
| 5  | <211> 34                                        |    |
|    | <212> ADN                                       |    |
|    | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |
|    | <223> Cebador inverso E137K                     |    |
| 10 | <400> 22                                        |    |
|    | cgtgaaagca aaatttgca tgtgtttg cgcc              | 34 |
|    | <210> 23                                        |    |
|    | <211> 27                                        |    |
|    | <212> ADN                                       |    |
| 15 | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |
|    | <223> Cebador directo T216A                     |    |
|    | <400> 23                                        |    |
|    | gcgaaacggt gaatgcqgca cgctacc                   | 27 |
| 20 | <210> 24                                        |    |
|    | <211> 27                                        |    |
|    | <212> ADN                                       |    |
|    | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |
| 25 | <223> Cebador inverso T216A                     |    |
|    | <400> 24                                        |    |
|    | ggtagcgtgc cgattcacc gttcgc                     | 27 |
|    | <210> 25                                        |    |
|    | <211> 43                                        |    |
| 30 | <212> ADN                                       |    |
|    | <213> Secuencia artificial                      |    |
|    | <220>                                           |    |
|    | <223> Cebador directo Y237C                     |    |
|    | <400> 25                                        |    |
| 35 | gcttcagaga aaacgaccgg aatgtcaaaa aagacaacac agg | 43 |
|    | <210> 26                                        |    |

<211> 43  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 5 <223> Cebador inverso Y237C  
 <400> 26  
 cctgtgtgtg ctttttgac attccggtcg ttttctctga agc 43  
 <210> 27  
 <211> 33  
 10 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador directo W268F  
 <400> 27  
 15 cgttgaaac actcaattc gaagtgttc cgc 33  
 <210> 28  
 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 20 <220>  
 <223> Cebador inverso W268F  
 <400> 28  
 gcggaagcac ttcgaaattg agtgtttcca acg 33  
 <210> 29  
 25 <211> 33  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador directo W268Y  
 30 <400> 29  
 cgttgaaac actcaattac gaagtgttc cgc 33  
 <210> 30  
 <211> 33  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>

# ES 2 374 418 T3

|    |                                      |    |
|----|--------------------------------------|----|
|    | <223> Cebador inverso W268Y          |    |
|    | <400> 30                             |    |
|    | gcggaagcac ttcgtaattg agtgtttcca acg | 33 |
|    | <210> 31                             |    |
| 5  | <211> 27                             |    |
|    | <212> ADN                            |    |
|    | <213> Secuencia artificial           |    |
|    | <220>                                |    |
|    | <223> Cebador directo W268A          |    |
| 10 | <400> 31                             |    |
|    | cgttgaaac actcaatgcg gaagtgc         | 27 |
|    | <210> 32                             |    |
|    | <211> 27                             |    |
|    | <212> ADN                            |    |
| 15 | <213> Secuencia artificial           |    |
|    | <220>                                |    |
|    | <223> Cebador inverso W268A          |    |
|    | <400> 32                             |    |
|    | gcactccgc attgagtgtt tccaacg         | 27 |
| 20 | <210> 33                             |    |
|    | <211> 24                             |    |
|    | <212> ADN                            |    |
|    | <213> Secuencia artificial           |    |
|    | <220>                                |    |
| 25 | <223> Cebador PCR MTP up             |    |
|    | <400> 33                             |    |
|    | tacgtaatgt cgagttcgt gccg            | 24 |
|    | <210> 34                             |    |
|    | <211> 24                             |    |
| 30 | <212> ADN                            |    |
|    | <213> Secuencia artificial           |    |
|    | <220>                                |    |
|    | <223> Amorce PCR 3'HindIII           |    |
|    | <400> 34                             |    |
| 35 | ccaagcta tcaaagtat ttgc              | 24 |
|    | <210> 35                             |    |

|    |                                      |    |
|----|--------------------------------------|----|
|    | <211> 25                             |    |
|    | <212> ADN                            |    |
|    | <213> Secuencia artificial           |    |
|    | <220>                                |    |
| 5  | <223> Cebador R para el método Selex |    |
|    | <400> 35                             |    |
|    | caggtcagtt cagcggatcc tgtcg          | 25 |
|    | <210> 36                             |    |
|    | <211> 25                             |    |
| 10 | <212> ADN                            |    |
|    | <213> Secuencia artificial           |    |
|    | <220>                                |    |
|    | <223> Cebador F para el método Selex |    |
|    | <400> 36                             |    |
| 15 | gaggcgaatt cagtgcaact gcagc          | 25 |
|    | <210> 37                             |    |
|    | <211> 28                             |    |
|    | <212> ADN                            |    |
|    | <213> <i>drosophila mauritiana</i>   |    |
| 20 | <220>                                |    |
|    | <223> Secuencia ITR3'                |    |
|    | <400> 37                             |    |
|    | tcaggtgtac aagtatgaaa tgtcgttt       | 28 |
|    | <210> 38                             |    |
| 25 | <211> 28                             |    |
|    | <212> ADN                            |    |
|    | <213> Secuencia artificial           |    |
|    | <220>                                |    |
|    | <223> Secuencia selex46              |    |
| 30 | <400> 38                             |    |
|    | tcaggtgtac aagtatgaga tgtcgttt       | 28 |
|    | <210> 39                             |    |
|    | <211> 28                             |    |
|    | <212> ADN                            |    |
| 35 | <213> Secuencia artificial           |    |
|    | <220>                                |    |

# ES 2 374 418 T3

|    |                                |    |
|----|--------------------------------|----|
|    | <223> Secuencia selex40        |    |
|    | <400> 39                       |    |
|    | tcagggttac aagtatgtaa tgcggtta | 28 |
|    | <210> 40                       |    |
| 5  | <211> 28                       |    |
|    | <212> ADN                      |    |
|    | <213> Secuencia artificial     |    |
|    | <220>                          |    |
|    | <223> Secuencia selex49        |    |
| 10 | <400> 40                       |    |
|    | gcgggtgtga aaatagaaa tattcact  | 28 |
|    | <210> 41                       |    |
|    | <211> 28                       |    |
|    | <212> ADN                      |    |
| 15 | <213> Secuencia artificial     |    |
|    | <220>                          |    |
|    | <223> Secuencia selexl         |    |
|    | <400> 41                       |    |
|    | gaagagttat aactcaagaa aactgaat | 28 |
| 20 | <210> 42                       |    |
|    | <211> 25                       |    |
|    | <212> ADN                      |    |
|    | <213> Secuencia artificial     |    |
|    | <220>                          |    |
| 25 | <223> Secuencia selex69        |    |
|    | <400> 42                       |    |
|    | cgattacaat tatgaaaact ttct     | 25 |
|    | <210> 43                       |    |
|    | <211> 22                       |    |
| 30 | <212> ADN                      |    |
|    | <213> Secuencia artificial     |    |
|    | <220>                          |    |
|    | <223> Secuencia selex6         |    |
|    | <400> 43                       |    |
| 35 | ggagaaatgg tagaatagtt ta       | 22 |
|    | <210> 44                       |    |

# ES 2 374 418 T3

|    |                                     |    |
|----|-------------------------------------|----|
|    | <211> 19                            |    |
|    | <212> ADN                           |    |
|    | <213> Secuencia artificial          |    |
|    | <220>                               |    |
| 5  | <223> Secuencia selex9              |    |
|    | <400> 44                            |    |
|    | atagtgttta aagcttact                | 19 |
|    | <210> 45                            |    |
|    | <211> 34                            |    |
| 10 | <212> ADN                           |    |
|    | <213> Secuencia artificial          |    |
|    | <220>                               |    |
|    | <223> Secuencia selex60             |    |
|    | <400> 45                            |    |
| 15 | gtcagtgatg tcaccaacgt atggaatc gtat | 34 |

**REIVINDICACIONES**

1. Pseudo-transposón *Mos-1* hiperactivo, en el que
  - a) por lo menos una de las dos repeticiones terminales no traducidas (UTR) y/o por lo menos una de las dos repeticiones terminales inversas (ITR) está(n) genéticamente modificada(s); y
  - 5 b) una secuencia nucleotídica exógena de interés sustituye a la secuencia nucleotídica que codifica la transposasa MOS-1 de origen, caracterizado porque se selecciona de entre los pseudo-transposones siguientes:
    - α) ITR3'-UTR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3'- (pseudo-transposón 33sec33),
    - β) ITR3'-secuencia nucleotídica exógena de interés-UTR3'-ITR3' (pseudo-transposón 2sec33),
    - 10 γ) ITR3'-UTR5'-secuencia nucleotídica exógena de interés-UTR3'-ITR5' (pseudo-transposón 35sec35),
    - δ) los pseudo-transposones que comprenden por lo menos una ITR40 de secuencia SEC ID nº 39, y
    - ε) los pseudo-transposones que comprenden por lo menos una ITR46 de secuencia SEC ID nº 38, y en el que ITR3', UTR3', UTR5' e ITR5' designan respectivamente la secuencia de la repetición terminal inversa 3' del transposón *Mos-1*, la secuencia de la repetición terminal no traducida 3' del transposón *Mos-1*, la secuencia de la repetición terminal no traducida 5' del transposón *Mos-1* y la secuencia de la repetición terminal inversa 5' del transposón *Mos-1*.
    - 15
2. Sistema de transposición recombinante hiperactivo derivado del transposón *Mos-1*, que comprende por lo menos las dos parejas siguientes:
  - a) un pseudo-transposón *Mos-1* hiperactivo según la reivindicación 1; y
  - 20 b) una transposasa *Mos-1* suministrada en trans de dicho pseudo-transposón, en el que la frecuencia de transposición de dicha secuencia nucleotídica exógena de interés está mejorada en un factor por lo menos igual a 5, preferentemente por lo menos igual a 10, con respecto a la frecuencia de transposición de dicha secuencia nucleotídica exógena de interés contenida por un pseudo-transposón *Mos-1* rodeado de dos ITR3' (3sec3) en presencia de la transposasa *Mos-1* suministrada en trans.
- 25 3. Sistema según la reivindicación 2, caracterizado porque dicha transposasa *Mos-1* suministrada en *trans* es una transposasa mutante.
4. Sistema según la reivindicación 3, caracterizado porque dicha transposasa *Mos-1* mutante es hiperactiva.
5. Sistema según la reivindicación 4, caracterizado porque dicha transposasa *Mos-1* mutante hiperactiva comprende por lo menos una mutación a nivel de por lo menos un residuo seleccionado de entre los residuos siguientes de la secuencia SEC ID nº 2: F53, Q91, E137, T216 e Y237.
- 30 6. Sistema según la reivindicación 5, caracterizado porque comprende por lo menos las dos parejas siguientes:
  - a) un pseudo-transposón *Mos-1* hiperactivo que comprende por lo menos una ITR40 de secuencia SEC ID nº 39 y una transposasa *Mos-1* mutante hiperactiva que comprende las mutaciones T216A e Y237C;
  - 35 b) un pseudo-transposón *Mos-1* hiperactivo que comprende por lo menos una ITR46 de secuencia SEC ID nº 38 y una transposasa *Mos-1* mutante hiperactiva que comprende las mutaciones T216A e Y237C, o E137K y T216A, o F53Y e T216A e Y237C;
  - c) un pseudo-transposón *Mos-1* hiperactivo 3sec33 y una transposasa *Mos-1* mutante hiperactiva que comprende las mutaciones T216A e Y237C, o E137K y T216A, o también F53Y e Q91R, o F53Y y Q91R y E137K y T216A.
- 40 7. Pseudo-transposón según la reivindicación 1 o sistema según cualquiera de las reivindicaciones 2 a 6, caracterizado porque dicha secuencia nucleotídica exógena de interés es un gen funcional.
8. Vector que comprende por lo menos un pseudo-transposón según la reivindicación 1.
9. Célula hospedante que comprende por lo menos:
  - a) un pseudo-transposón según la reivindicación 1 ó 7, o
  - b) un sistema según cualquiera de las reivindicaciones 2 a 7; o
  - 45 c) un vector según la reivindicación 8; o

- d) una combinación de éstos.
10. Kit que comprende por lo menos:
- a) un pseudo-transposón según la reivindicación 1 ó 7; o
- b) un sistema según cualquiera de las reivindicaciones 2 a 7; o
- 5 c) un vector según la reivindicación 8; o
- d) una célula hospedante según la reivindicación 9; o
- e) una combinación de éstos.
11. Utilización de por lo menos:
- a) un pseudo-transposón según la reivindicación 1 ó 7; o
- 10 b) un sistema según cualquiera de las reivindicaciones 2 a 7; o
- c) un vector según la reivindicación 8; o
- d) una célula hospedante según la reivindicación 9; o
- e) un kit según la reivindicación 10; o
- f) una combinación de éstos.
- 15 para la transposición eficaz *in vitro* o *ex vivo* de una secuencia nucleotídica exógena de interés.
12. Utilización de por lo menos:
- a) un pseudo-transposón según la reivindicación 1 ó 7; o
- b) un sistema según cualquiera de las reivindicaciones 2 a 7; o
- c) un vector según la reivindicación 8; o
- 20 d) una célula hospedante según la reivindicación 9; o
- e) un kit según la reivindicación 10; o
- f) una combinación de éstos.
- para la mutagénesis de inserción *in vitro* o *ex vivo*.
13. Utilización de por lo menos:
- 25 a) un pseudo-transposón según la reivindicación 1 ó 7; o
- b) un sistema según cualquiera de las reivindicaciones 2 a 7; o
- c) un vector según la reivindicación 8; o
- d) una célula hospedante según la reivindicación 9; o
- e) un kit según la reivindicación 10; o
- 30 f) una combinación de éstos.
- para la secuenciación y/o la clonación de ácidos nucleicos.
14. Procedimiento de preparación de un medicamento, que comprende por lo menos una etapa de transposición *in vitro* o *ex vivo* de una secuencia de ADN transposable de interés en una secuencia de ADN diana, siendo dicha transposición mediada por lo menos por uno de los medios siguientes:
- 35 a) un pseudo-transposón según la reivindicación 1 ó 7; o
- b) un sistema según cualquiera de las reivindicaciones 2 a 7; o
- c) un vector según la reivindicación 8; o

- d) una célula hospedante según la reivindicación 9; o
- e) un kit según la reivindicación 10; o
- f) una combinación de éstos.

15. Utilización de por lo menos:

5

- a) un pseudo-transposón según la reivindicación 1 ó 7; o
- b) un sistema según cualquiera de las reivindicaciones 2 a 7; o
- c) un vector según la reivindicación 8; o
- d) una célula hospedante según la reivindicación 9; o
- e) un kit según la reivindicación 10; o

10

- f) una combinación de éstos.

para la preparación de un medicamento destinado a permitir la transposición eficaz *in vivo* de una secuencia nucleotídica exógena de interés.