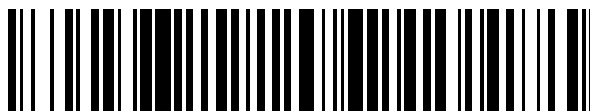


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 701 883**

51 Int. Cl.:

C12N 15/10 (2006.01)

C12P 19/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **31.07.2008 E 16154206 (3)**

97 Fecha y número de publicación de la concesión europea: **19.09.2018 EP 3031919**

54 Título: **Ensamblaje de combinaciones de múltiples sitios a medida**

30 Prioridad:

31.07.2007 US 953171 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

26.02.2019

73 Titular/es:

**BASF ENZYMES LLC (100.0%)
3550 John Hopkins Court
San Diego, CA 92121, US**

72 Inventor/es:

TAN, XUQIU

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 701 883 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

Ensamblaje de combinaciones de múltiples sitios a medida

Referencia a solicitudes relacionadas

La presente solicitud reivindica la prioridad a la Solicitud Provisional de EE. UU. Número de Serie 60/953.171, presentada el 31 de julio de 2007.

Antecedentes de la invención**Campo de la invención**

La presente invención se refiere en general a un procedimiento de ensamblaje de combinaciones de múltiples sitios a medida ("TMCA") como un procedimiento de producción de una pluralidad de progenie de polinucleótidos y producción de cambios en un gen y reagrupamiento de mutaciones o cambios en múltiples sitios. Las mutaciones o cambios se diseñan y sintetizan en oligonucleótidos cortos. Los oligonucleótidos se hibridan con una matriz de ADN que comprende el gen de tipo silvestre. Se utiliza una ADN polimerasa para amplificar el ADN completo. El ADN amplificado resultante se recupera de un huésped. Las ventajas de este procedimiento son la velocidad, simplicidad técnica, y capacidad para controlar el ensamblaje.

Descripción de los antecedentes

Los procedimientos publicados para hacer cambios en un gen utilizan, por ejemplo, una PCR con tendencia a errores, el Kit Gene Tailor site-directed Mutagénesis™ de Invitrogen, el Kit QuickChange Mutagénesis™ de Stratagene, PCR solapadas y PCR basada en ligado/recombinación. Una supervisión de los procedimientos conocidos revela que estos procedimientos tienden a hacer frente a una dificultad primaria para la generación de una mutación y/o modificación en un único sitio/región vecinos y/o cuesta mucho trabajo hacer modificaciones en múltiples regiones.

La patente de EE. UU. 7.202.086 ("la patente '086") reivindica un procedimiento de mutagénesis que utiliza al menos 5 oligonucleótidos que se solapan o no y un dsADN (plasmídico) para generar una biblioteca de genes mutados, en la que cada mutación está presente como media en menos de 1/5 de los genes de la biblioteca. La patente '086 describe que la invención desvelada es diferente de la de la técnica anterior debido a que la patente '086 necesita el control de la frecuencia de mutaciones para evitar un "exceso de mutaciones" en una molécula de ADN (col. 5, líneas 28-45). Se desea obtener mutantes que contengan cada uno, una mutación. Con el fin de conseguir este objetivo, la relación entre la cantidad de cada oligonucleótido mutante y la cantidad de matriz debe ser entre 0,01-100 (col 5, líneas 28-45). Esta característica la distingue de la técnica anterior, en la que utilizando varios oligonucleótidos simultáneamente da lugar al nivel de incorporación de cada cebador de más del 75 % (col 5, líneas 28-45). La patente '086 necesita el control de la frecuencia de mutaciones para evitar un "exceso de mutaciones" en una molécula de ADN y para generar mutantes que contengan cada uno, una mutación.

La patente de EE. UU. 7.132.265 ("la patente '265") y la publicación de patente de EE. UU. 2003/0064516 reclaman un procedimiento de introducción de mutaciones en una molécula de ADN de cadena sencilla ("ssADN") que comprende la hibridación de un cebador, la síntesis de una cadena de ADN, y la digestión de la molécula de ADN. El procedimiento TMCA utiliza un ADN de doble cadena ("dsADN") como matriz. La patente '265 hace una clara distinción entre la utilización de ssADN y dsADN como sustrato de partícula para los protocolos de mutagénesis (véase, por ejemplo, la col. 6. Líneas 45-55).

La solicitud publicada de EE. UU. 2003/0194807 se refiere a una biblioteca en la que los mutantes de una proteína comprenden un único aminoácido predeterminado en una o más posiciones de una región definida, en el que la región definida es de al menos tres aminoácidos. Solamente se permite un único cambio en una localización, es decir, se excluyen cambios degenerativos en una posición de aminoácido determinada.

Cada una de las solicitudes publicadas de EE. UU. 2006/0051748; 2006/0134624; 2004/0248131; y 2002/0083488; y patentes de EE. UU. 6.673.610; 6.335.160; y 5.354.670 necesitan ligar el ADN sintetizado para producir una progenie de ADN circular con las mutaciones.

Además, la solicitud publicada de EE. UU. 2006/0051748 necesita el uso de una endonucleasa flap y la hibridación de todos los cebadores en la misma cadena de ADN. La solicitud publicada de EE. UU. 2006/0134624 necesita el uso de dos cebadores consecutivamente (es decir, no en una reacción). En la solicitud publicada de EE. UU. 2004/0248131, los cebadores se hibridan a las dos cadenas, en la que los cebadores no tienen que comprender 2-4 pares de bases complementarias. La patente de EE. UU. 6.673.610 y solicitud publicada de EE. UU. 2002/0083488 necesitan el uso de fragmentos producidos por digestión de la cadena de ADN parental como un megacebador para la obtención de un ADN circular que es utiliza en la transformación. La patente de EE. UU. 6.335.160 se refiere al ensamblaje genético de fragmentos solapados y la generación de una biblioteca recombinante. Finalmente, la patente de EE. UU. 5.354.670 necesita dos etapas de transformación y un tratamiento intermedio con una restricción.

Cada una de las patentes de EE. UU. 7.176.004; 6.713.285; 6.391.548; y 5.932.419; y solicitudes publicadas de EE. UU. 20040253729 y 20030032037 necesitan hibridar dos cebadores a dos cadenas diferentes para iniciar la amplificación en direcciones contrarias (es decir, cebadores directo e inverso) y tener regiones complementarias. Las patentes de EE. UU. 7.078.389 y 5.935.830 necesitan que un cebador comprenda un mutágeno (por ejemplo, psoraleno) que interactúa con una matriz de manera que se forma una molécula de triple cadena.

La solicitud publicada de EE. UU. 2006/0228786 necesita llevar a cabo la polimerización de dos cadenas utilizando dos cebadores diferentes en dos reacciones diferentes seguido por hibridación de las moléculas de ssADN sintetizadas. La solicitud publicada de EE. UU. 2003/0077613 se refiere a un procedimiento para el ensamblaje genético y la creación de una biblioteca, en la que un gen ensamblado (ssADN) se hibrida con un armazón de ADN para llenar huecos y generar un dsADN que se subclona en un vector.

La solicitud publicada de EE. UU. 2004/0002057 describe un procedimiento para la detección de un ligando en una muestra que no comprende mutagénesis. La solicitud publicada de EE. UU. 2004/0002057 describe un procedimiento de establecimiento de una cepa mutante de *E. coli* utilizando un mutágeno en células cultivadas. La solicitud publicada de EE. UU. 2006/0199222 describe un procedimiento genérico de evolución dirigida en el que el ADN mutado se transforma en una cepa de *Bacillus* particular.

Sigue existiendo la necesidad de un procedimiento mejor y más eficaz de generación de variantes genéticas específicas y una biblioteca genética de combinaciones de manera eficaz y rápida.

Sumario de la invención

A menos de que se defina específicamente, todos los términos científicos utilizados en el presente documento tienen el significado que entendería comúnmente, una vez visto en contexto, un experto en la técnica que proporciona el contexto, por ejemplo, química, bioquímica, biología celular, biología molecular, o ciencias médicas.

En consecuencia, un objetivo de la presente invención es proporcionar un procedimiento de producción de una pluralidad de polinucleótidos modificados que tienen diferentes combinaciones de distintas mutaciones en múltiples sitios mediante un ensamblaje de combinaciones de sitios múltiples a medida. La presente invención permite hacer cambios específicos en un gen y reagrupar las mutaciones o cambios de múltiples sitios del gen. Estos y otros objetivos de la presente invención, que se volverán más evidentes en conjunción con la siguiente descripción detallada de las realizaciones preferidas, sea solas o en combinación de las mismas, han sido satisfechas por el descubrimiento de un procedimiento de producción de una pluralidad de polinucleótidos modificados que tienen diferentes combinaciones de distintas mutaciones en múltiples sitios, en los que las mutaciones son cambios de uno o más nucleótidos en la secuencia de una secuencia de ácido nucleico parental, comprendiendo dicho procedimiento:

(a) añadir al menos tres cebadores a un polinucleótido matriz de doble cadena en una única mezcla de reacción, en el que los al menos tres cebadores no se solapan, y en el que cada uno de los tres cebadores comprende al menos una mutación diferentes de los otros cebadores, en el que al menos un cebador es un cebador directo que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz, y

(b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de los al menos tres cebadores.

En otra realización, un procedimiento de producción una pluralidad de polinucleótidos modificados que tienen diferentes combinaciones de distintas mutaciones en múltiples sitios y comprenden las mutaciones de interés, en el que las mutaciones son cambios de uno o más nucleótidos en la secuencia de una secuencia de ácido nucleico parental, comprendiendo dicho procedimiento:

(a) añadir al menos dos cebadores a un polinucleótido matriz en una única mezcla de reacción, en el que los al menos dos cebadores no se solapan, y en el que cada uno de los al menos dos cebadores comprenden al menos una mutación diferentes de la del otro cebador(es), en el que al menos un cebador es un cebador directo que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz,

(b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de al menos dos cebadores,

(c) tratar la pluralidad de polinucleótidos modificados extendidos con una enzima, destruyendo de esta manera el polinucleótido matriz,

(d) transformar los polinucleótidos modificados extendidos tratados que no se habían tratado con una ligasa en una célula,

(e) recuperar la pluralidad de polinucleótidos modificados extendidos de la célula, y

(f) seleccionar la pluralidad de polinucleótidos modificados extendidos que comprenden las mutaciones de interés.

Los aspectos y realizaciones de la presente divulgación también se pueden proporcionar de acuerdo con los siguientes párrafos numerados:

- 5 1. Un procedimiento de producción de una pluralidad de polinucleótidos modificados, que comprende:
 - (a) añadir al menos tres cebadores a un polinucleótido matriz de doble cadena en una única mezcla de reacción, en el que los al menos tres cebadores no se solapan, y en el que cada uno de los al menos tres cebadores comprenden al menos una mutación diferentes de la de los otros cebadores, en el que al menos un cebador es un cebador directo que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz, y
 - (b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de los al menos tres cebadores.
- 10 2. El procedimiento del párrafo 1, que comprende adicionalmente la transformación de una célula con la pluralidad de productos extendidos que no se han tratado con una ligasa.
- 15 3. El procedimiento del párrafo 2, que comprende adicionalmente la recuperación de la pluralidad de polinucleótidos modificados extendidos de la célula.
4. El procedimiento del párrafo 3, que comprende adicionalmente el análisis de la pluralidad de polinucleótidos modificados extendidos.
- 20 5. El procedimiento del párrafo 4, en el que el análisis comprende la expresión de al menos uno de la pluralidad de polinucleótidos modificados extendidos y el análisis del polipéptido que se expresa a partir de estos.
6. El procedimiento del párrafo 5, que comprende adicionalmente la selección de la pluralidad de polinucleótidos modificados extendidos que comprende las mutaciones de interés.
- 25 7. El procedimiento del párrafo 1, que comprende adicionalmente antes de la etapa (a) la obtención de la información de secuencia del polinucleótido matriz, y la identificación de tres o más mutaciones de interés a lo largo del polinucleótido matriz.
8. El procedimiento del párrafo 1, que comprende adicionalmente el análisis de la pluralidad de polinucleótidos modificados extendidos producidos por la extensión con la polimerasa.
- 30 9. El procedimiento del párrafo 1, que comprende adicionalmente el tratamiento de la pluralidad de polinucleótidos modificados extendidos con una enzima, destruyendo de esta manera el polinucleótido matriz, la transformación de los polinucleótidos modificados extendidos tratados en una célula, la recuperación de la pluralidad de polinucleótidos modificados extendidos de una célula, y la selección de la pluralidad de polinucleótidos modificados extendidos que comprende las mutaciones de interés.
10. El procedimiento del párrafo 9, en el que la célula es una célula de *E. coli*.
11. El procedimiento del párrafo 9, en el que la enzima es una enzima de restricción.
- 35 12. El procedimiento del párrafo 11, en el que la enzima de restricción es la enzima de restricción *DpnI* y la célula es una célula de *E. coli*.
13. El procedimiento del párrafo 1, en el que se añaden al menos cuatro cebadores.
14. El procedimiento del párrafo 1, en el que se añaden al menos cinco cebadores.
15. El procedimiento del párrafo 1, en el que se añaden al menos seis cebadores.
- 40 16. El procedimiento del párrafo 1, en que se añaden al menos ocho cebadores.
17. El procedimiento del párrafo 1, en el que se añaden al menos doce cebadores.
18. El procedimiento del párrafo 1, en el que cada cebador comprende una única mutación puntual.
19. El procedimiento del párrafo 1, en el que al menos dos cebadores directos comprenden un cambio diferente en la misma posición del polinucleótido matriz.
- 45 20. El procedimiento del párrafo 1, en el que al menos dos cebadores inversos comprenden un cambio diferente en la misma posición del polinucleótido matriz.

21. El procedimiento del párrafo 1, en el que al menos un cebador comprende al menos dos cambios en posiciones diferentes del polinucleótido matriz.
22. El procedimiento del párrafo 1, en el que al menos un cebador comprende al menos dos cambios en posiciones diferentes y al menos dos cebadores directos o dos inversos comprenden un cambio diferente en la misma posición del polinucleótido matriz.
23. El procedimiento de la reivindicación 1, en el que al menos una mutación se selecciona de entre el grupo que consiste en un cambio de uno o más nucleótidos o secuencias de aminoácidos codificadas, una inserción y una eliminación.
24. El procedimiento del párrafo 1, en el que el polinucleótido matriz es una ADN circular de doble cadena.
25. El procedimiento del párrafo 1, en el que al menos un cebador es un conjunto de cebadores degenerados que comprenden cada uno una posición degenerada, en el que la mutación de interés es un intervalo de diferentes nucleótidos en la posición degenerada.
26. El procedimiento del párrafo 1, en el que al menos un cebador es un conjunto de cebadores degenerados que comprende al menos un codón degenerado correspondiente a al menos un codón del polinucleótido matriz y al menos una secuencia adyacente que es homólogo de una secuencia adyacente al codón del polinucleótido matriz.
27. El procedimiento del párrafo 26, en el que el codón degenerado es N,N,N que codifica un aminoácido de origen natural.
28. El procedimiento del párrafo 26, en el que el codón degenerado puede codificar menos de los 20 aminoácidos de origen natural.
29. El procedimiento del párrafo 1, en el que los cebadores directos se agrupan en un grupo directo y los cebadores inversos se agrupan en una grupo inverso, y los cebadores del grupo directo y los cebadores del grupo inverso, se normalizan independientemente entre ellos para tener una concentración igual en el grupo correspondiente independientemente de las posiciones en el polinucleótido matriz, y en el que después de la normalización se añade una cantidad igual de cebadores directos e inversos se añade a la reacción.
30. El procedimiento del párrafo 1, que comprende adicionalmente antes de la etapa (b):
 la organización de los cebadores en múltiples grupos dependiendo de su localización en el polinucleótido matriz, en el que los cebadores que cubren la misma región seleccionada en la matriz están en un grupo, la normalización de los cebadores agrupados en cada grupo para que tengan la misma concentración, el agrupamiento de los cebadores directos en un grupo del grupo directo y la normalización de la concentración entre cada grupo de los cebadores directos para que sean iguales, el agrupamiento de los cebadores inversos en un grupo en el grupo inverso y la normalización de la concentración entre cada grupo de los cebadores inversos para que sean iguales, y la adición de una cantidad igual de los cebadores directos e inversos agrupados en la reacción.
31. El procedimiento del párrafo 1, que comprende adicionalmente llevar a cabo dos rondas de las etapas (a) a (b), y la utilización de los polinucleótidos producidos en la primera ronda como polinucleótido matriz en la segunda ronda.
32. Un procedimiento de producción de una pluralidad de polinucleótidos modificados que comprende las mutaciones de interés, que comprenden:
 (a) añadir al menos dos cebadores a un polinucleótido matriz de doble cadena en una única mezcla de reacción, en el que al menos dos cebadores no se solapan, y en el que cada uno de los al menos dos cebadores comprenden al menos una mutación diferentes de la del otro cebador, en el que al menos un cebador es un cebador directo que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz,
 (b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de al menos dos cebadores,
 (c) tratar la pluralidad de polinucleótidos modificados extendidos con una enzima, destruyendo de esta manera el polinucleótido matriz,
 (d) transformar los polinucleótidos modificados extendidos tratados que no se han tratado con una ligasa en una célula,
 (e) recuperar la pluralidad de polinucleótidos modificados extendidos de la célula, y

(f) seleccionar la pluralidad de polinucleótidos modificados extendidos que comprenden las mutaciones de interés.

Los detalles de una o más realizaciones de la invención se exponen en la descripción adjunta posteriormente.

5 Los procedimientos y materiales preferidos se describen ahora. Otras características, objetos, y ventajas de la invención serán evidentes a partir de la descripción y de las reivindicaciones. En la memoria descriptiva y las reivindicaciones adjuntas, las formas singulares incluyen los referentes en plural a menos de que el contexto dicte claramente otra cosa. A menos de que se establezca expresamente otra cosa, las técnicas empeladas o contempladas en el presente documento son metodologías convencionales bien conocidas por un experto habituado en la técnica. Los ejemplos de realizaciones son solo con fines ilustrativos.

10 **Breve descripción de los dibujos**

Una apreciación más completa de la invención y muchas de las ventajas relacionadas de la misma se obtendrán fácilmente cuando la misma se entienda mejor en referencia a la siguiente descripción detallada cuando se considera en conexión con los dibujos adjuntos, en los que:

- 15 Fig. 1. Una representación esquemática de GSSMSM.
 Fig. 2. Una representación esquemática de un flujo del procedimiento de evolución - GSSMSM.
 Fig. 3. Una representación esquemática de un ensamblaje de combinaciones de sitios múltiples a medida.
 Fig. 4A-D. combinaciones de cebadores en la reacción de TMCA.
 Fig. 5. Un mapa de cebadores que se hibridan en un conjunto de seis mutaciones.
 20 Fig. 6. Distribución de posibles combinaciones con seis sitios de mutación. Referencia: distribución de variantes calculada en una situación ideal; 1A: condición de reacción 1 con la cepa XL1-Blue de *E. coli*; 7A: condición de reacción 2 con la cepa XL1-Blue de *E. coli*; 13A: condición de reacción 3 con la cepa XL1-Blue de *E. coli*; Total: datos combinados de 1A, 7A y 13A; 0x: sin mutación; 1x: mutación única; 2x: dos mutaciones; 4x: cuatro mutaciones; 5x: cinco mutaciones; 6x: seis mutaciones.
 Fig. 7. Cálculos estadísticos vs datos experimentales en el conjunto de seis mutaciones.
 25 Fig. 8. Un mapa de hibridación de cebadores en un conjunto de cuatro mutaciones.
 Fig. 9. Distribución de las combinaciones posibles con cuatro sitios de mutación. Referencia: distribución de variantes calculada en una situación ideal; 2A: condición de reacción 1 con la cepa XL1-Blue de *E. coli*; 8A: condición de reacción 2 con la cepa XL1-Blue de *E. coli*; 14A: condición de reacción 3 con la cepa XL1-Blue de *E. coli*; Total: datos combinados de 2A, 8A y 14A; 0x: sin mutación; 1x: mutación única; 2x: dos mutaciones; 4x: cuatro mutaciones.
 30 Fig. 10. Distribución de las combinaciones posibles con cuatro sitios de mutación. Referencia: distribución de variantes calculada en una situación ideal; 2B: condición de reacción 1 con la cepa Stb12 de *E. coli*; 8B: condición de reacción 2 con la cepa Stb12 de *E. coli*; 14B: condición de reacción 3 con la cepa Stb12 de *E. coli*; Total: datos combinados de 2B, 8B y 14B; 0x: sin mutación; 1x: mutación única; 2x: dos mutaciones; 4x: cuatro mutaciones.
 35 Fig. 11. Un mapa de la hibridación de cebadores en un conjunto de tres mutaciones.
 Fig. 12. Distribución de las combinaciones posibles con tres sitios de mutación. Referencia: distribución de variantes calculada en una situación ideal; 15A: condición de reacción 1 con la cepa XL1-Blue de *E. coli*; 9B: condición de reacción 2 con la cepa Stb12 de *E. coli*; 15B: condición de reacción 3 con la cepa Stb12 de *E. coli*; Total: datos combinados de 15A, 9B y 15B; 0x: sin mutación; 1x: mutación única; 2x: dos mutaciones.
 40 Fig. 13. Un mapa de la hibridación de cebadores con 5 sitios de mutación y 13 mutantes.
 Fig. 14A, B, C. Combinaciones de variantes únicas en cinco sitios de mutación y ensamblaje de 13 cebadores. Desglose de las variantes en la ronda I de TMCA (A), después la ronda II (B), un mapa de la hibridación de cebadores (C).

Descripción detallada de las realizaciones preferidas

45 El procedimiento TMCA puede generar una variante genética específica que comprende múltiples cambios o una biblioteca genética de combinaciones eficaz y rápidamente; necesita el mínimo coste y esfuerzo; y se puede hacer a medida para hacer una biblioteca de combinaciones con una tendencia según las "necesidades". El procedimiento TMCA se puede llevar a cabo sin emplear una etapa de ligado y, por lo tanto, se simplifica el procedimiento de generación de múltiples mutaciones. Las "necesidades" de una biblioteca en particular varían por los experimentos.
 50 Los sitios potenciales de mutación – "las necesidades" – por ejemplo, pueden ser 1) diseñar racionalmente cambios de aminoácidos o 2) alteraciones de aminoácidos individuales determinadas empíricamente para producir un efecto deseado en una enzima (determinado por GSSMSM y esfuerzo de exploración). Cada biblioteca se crea con un número específico de sitios potenciales de mutación. Puede preferirse crear una biblioteca con tendencia hacia la progenie con más o menos mutaciones en los sitios potenciales de mutación. Al igual, puede ser preferible crear una
 55 biblioteca en la que exista una tendencia hacia o contra una mutación particular o sitio de mutación.

En la presente solicitud, los presentes inventores diseñaron un procedimiento de ensamblaje de combinaciones de sitios múltiples a medida que se ilustra en general en la Fig. 3. Por comparación, la Fig. 1 y Fig. 2 ilustran la evolución - Mutagénesis de Sitio Genético Saturado ("GSSM") en la que cada posición de mutación puede contener dos o más mutaciones para diferentes aminoácidos. La evolución - GSSMSM se puede utilizar para introducir

cambios de nucleótidos en un gen específico y para mutar cada codón de una fase de lectura abierta de todos los demás aminoácidos de un resto o más al mismo tiempo. Por lo tanto, se crea una biblioteca GSSMSM en la que un único clon comprende un ADN que tiene un cambio, mientras que los polipéptidos de la progenie en una biblioteca creada por el procedimiento TMCA pueden comprender múltiples mutaciones, preferentemente dos o más, más preferentemente tres o más, más preferentemente cuatro o más, cinco o más, seis o más, ocho o más, diez o más, y más preferentemente doce o más mutaciones.

Utilizando la técnica GSSMSM, se cambia un resto cada vez para cubrir los 20 aminoácidos. La biblioteca se explora y se identifican los mutantes. La reacción TMCA se diseña para hacer mutaciones en múltiples sitios de una molécula. La reacción TMCA se puede utilizar para combinar los mutantes identificados a partir de la biblioteca GSSMSM. En las condiciones de la reacción TMCA, se esperaría la formación de múltiples productos de PCR. No se espera que los productos de la PCR se transformen en las células y se amplifiquen.

En el contexto de la presente invención, el término “aminoácido”, como se utiliza en el presente documento se refiere a cualquier compuesto orgánico que contiene un grupo amino (-NH₂) y un grupo carboxilo (-COOH); preferentemente como grupos libres o de manera alternativa tras la condensación como parte de enlaces peptídicos. Se entiende en la técnica que los “veinte aminoácidos de origen natural” y se refieren a: alanina (ala o A), arginina (arg o R), asparagina (asn o N), ácido aspártico (asp o D), cisteína (cys o C), ácido glutámico (glu o E), glutamina (gln o Q), glicina (gly o G), histidina (his o H), isoleucina (ile o I), leucina (leu o L), lisina (lys o K), metionina (met o M), fenilalanina (phe o F), prolina (por o P), serina (ser o S), treonina (thr o T), triptófano (trp o W), tirosina (tyr o Y) y valina (val o V).

El término “amplificación” (“una reacción de extensión por la polimerasa”) significa que el número de copias de un polinucleótido está aumentado.

La expresión “se corresponde a” se utiliza en el presente documento para significar que una secuencia de polinucleótido es homóloga (es decir, es idéntica, no relacionada estrictamente por la evolución) con toda o una parte de una secuencia de polinucleótido de referencia, o que una secuencia de polipéptido es idéntica a una secuencia de polipéptido de referencia. Por el contrario, el término “complementario a” se utiliza en el presente documento para significar que la secuencia complementaria es homóloga a toda o parte de una secuencia de polinucleótido. A modo de ilustración, la secuencia de nucleótidos “TATAC” se corresponde con una “TATAC” de referencia y es complementaria a una secuencia “GTATA” de referencia.

Un “cebador” se define en el presente documento como una cadena de ácido nucleico que se puede hibridar con un ácido nucleico matriz y sirve como punto de partida para la amplificación de ADN. El cebador puede ser completa o parcialmente complementario de una región específica del polinucleótido matriz. Un nucleótido no complementario se define en el presente documento como una falta de coincidencia. Una falta de coincidencia se puede localizar dentro del cebador o en cualquier extremo del cebador. Preferentemente, se localizan en el cebador la falta de coincidencia de un único nucleótido, más preferentemente dos, y más preferentemente, tres o más faltas de coincidencias consecutivas o no consecutivas de nucleótidos. El cebador tiene de 6 a 200 nucleótidos, preferentemente, de 20 a 80 nucleótidos, y más preferentemente, de 43 a 65 nucleótidos. Más preferentemente, el cebador tiene 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, o 190 nucleótidos. Un “cebador directo” como se define en el presente documento es un cebador que es complementario a una cadena menos del polinucleótido matriz. Un “cebador inverso” como se define en el presente documento es un cebador complementario a una cadena más del polinucleótido matriz. Preferentemente, los cebadores directo e inverso no comprenden secuencias de nucleótidos solapadas. “No comprende secuencias de nucleótidos solapadas” como se define en el presente documento significa que un cebador directo e inverso no se hibridan a una región de las cadenas menos y más, respectivamente, del polinucleótido matriz en el que las cadenas más y menos son complementarias entre ellas. Con respecto a la hibridación de cebadores a la misma cadena del polinucleótido matriz “no comprende secuencias de nucleótidos solapadas” significa que los cebadores no comprenden secuencias complementarias a la misma región de la misma cadena del polinucleótido matriz.

La cadena más es la misma que la cadena en sentido y también se puede llamar la cadena codificante o no matriz. Esta es la cadena que tiene la misma secuencia que el ARNm (excepto que tiene T en vez de U). La otra cadena, llamada la matriz, menos o antisentido, es complementaria del ARNm.

Los “cebadores que cubren la misma región seleccionada del polinucleótido matriz” se definen en el presente documento como un conjunto de cebadores degenerados que comprende cada uno al menos una posición degenerada, en el que la mutación de interés es un intervalo de diferentes nucleótidos en la posición degenerada, o un conjunto de cebadores degenerados que comprende al menos un codón degenerado correspondiente a al menos un codón del polinucleótido matriz, o una combinación de los mismos. Por ejemplo, un conjunto de cebadores para todas las posibles combinaciones para las mutaciones de tres codones Y276F/S282L, H, P, R, o C/L284F (véase, por ejemplo, la Fig. 4, 15, o 16) son los cebadores que cubren la misma región de la matriz. “Cebadores que cubren la misma región seleccionada del polinucleótido matriz” también pueden ser, por ejemplo, una combinación de cebadores específicos.

“Digestión” de ADN se refiere a la escisión catalítica del ADN con una enzima de restricción que actúa solo en ciertas secuencias del ADN. Las distintas enzimas de restricción que se utilizan en el presente documento están disponibles en el mercado y sus condiciones de reacción, cofactores y otras necesidades se utilizaron como se conoce por el experto habituado en la técnica. Con fines analíticos, normalmente se utiliza 1 µg de plásmido o fragmento de ADN con aproximadamente 2 unidades de enzima en aproximadamente 20 µl de solución tampón. Con el fin de aislar los fragmentos de ADN para la construcción de plásmidos, normalmente se digirieron 5 a 50 µg de ADN con 20 a 250 unidades de enzima en un volumen más grande. Las cantidades apropiadas de tampones y de sustrato para las enzimas de restricción particulares se especifican por el fabricante. Se utilizan habitualmente tiempos de incubación de aproximadamente 1 hora a 37 °C, pero puede variar de acuerdo con las instrucciones del suministrador. Después de la digestión, la reacción puede someterse a electroforesis en gel.

Enzimas “recombinantes” se refiere a enzimas producidas por técnicas de ADN recombinante, es decir, producidas a partir de células transformadas por una construcción de ADN exógena que codifica a enzima deseada. Enzimas “sintéticas” son las que se preparan por síntesis química.

La expresión “sitio de restricción” se refiere a una secuencia de reconocimiento de secuencia que es necesario para la manifestación de la acción de una enzima de restricción, e incluye un sitio de escisión catalítica. Se aprecia que un sitio de escisión puede estar contenido o no en una parte de un sitio de restricción que comprende una secuencia de baja ambigüedad (es decir, una secuencia que contiene el determinante principal de la frecuencia de la existencia del sitio de restricción). Por lo tanto, en muchos casos, los sitios de restricción relevantes contienen solo una secuencia de baja ambigüedad con un sitio de escisión interno (por ejemplo, G/AATTC en el sitio *EcoRI*) o un sitio de escisión inmediatamente adyacente (por ejemplo, /CCWGG en el sitio *EcoRII*). En otros casos, las enzimas de restricción relevantes (por ejemplo, el sitio *Eco57I* o CTGAAG (16/14)) que contienen una secuencia de baja ambigüedad (por ejemplo, la secuencia CTGAAG en el sitio *Eco57I*) con un sitio de escisión externa (por ejemplo, en la parte N₁₆ del sitio *Eco57I*). Cuando se dice que una enzima (por ejemplo, una enzima de restricción) “escinde” un polinucleótido, se entiende que significa que la enzima de restricción cataliza o facilita una escisión de un polinucleótido.

Un “requisito de base ambigua” en un sitio de restricción se refiere a un requisito de base de nucleótido que no se especifica en su extensión más completa, es decir, que no es una base específica (tal como, en una ejemplificación no limitante, una base específica seleccionada de entre A, C, G y T), sino más bien puede ser una cualquiera de al menos dos o más bases. Las abreviaturas comúnmente aceptadas que se utilizan en la técnica así como en el presente documento para representar la ambigüedad en las bases incluyen las siguientes: R = G o A; M = A o C; K = G o T; S = G o C; W = A o T; H = A o C o T; B = G o T o C; V = G o C o A; D = G o A o T; N = A o C o G o T.

Una “secuencia de referencia” es una secuencia definida utilizada como base para una comparación de secuencias; una secuencia de referencia puede ser un subconjunto de una secuencia más grande, por ejemplo, como un segmento de un ADNc de longitud completa o secuencia genética determinada de una lista de secuencias, o puede comprender un ADNc completo o secuencia genética. En general, una secuencia de referencia tiene al menos 20 nucleótidos de longitud, frecuentemente al menos 25 nucleótidos de longitud, y a menudo al menos 50 nucleótidos de longitud. Como dos polinucleótidos pueden (1) comprender cada uno una secuencia (es decir, una parte de la secuencia de polinucleótido completa) que sea similar entre los dos polinucleótidos y (2) puede comprender adicionalmente una secuencia que sea divergente entre los dos polinucleótidos, normalmente se llevan a cabo comparaciones de secuencia entre dos (o más) polinucleótidos comparando las secuencias de los dos polinucleótidos sobre una “ventana de comparación” para identificar y comparar regiones locales de similitud de secuencia.

Una “ventana de comparación”, como se utiliza en el presente documento, se refiere a un segmento conceptual de al menos 20 posiciones de nucleótido contiguos en el que una secuencia de polinucleótido se puede comparar con una secuencia de referencia de al menos 20 nucleótidos contiguos y en el que la parte de la secuencia de polinucleótido en la ventana de comparación puede comprender adiciones o eliminaciones (es decir, huecos) del 20 por ciento o menos en comparación con la secuencia de referencia (que no comprende adiciones o eliminaciones) para el alineamiento óptimo de las dos secuencias. El alineamiento óptimo de secuencias por alineamiento en una ventana de comparación se puede llevar a cabo por el algoritmo de homología local de Smith (Smith y Waterman, *Adv Appl Math*, 1981; Smith y Waterman, *J Teor Biol*, 1981; Smith y Waterman, *J Mol Biol*, 1981; Smith y col, *J Mol Evol*, 1981), por el algoritmo de homología de alineamiento de Needleman (Needleman y Wuncsch, 1970), por la búsqueda del procedimiento de similitud de Pearson (Pearson y Lipman, 1988), por implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA del Paquete de Software de Genética Wisconsin, lanzamiento 7.0, Genetics Computer Group, 575 Science Dr., Madison, Wis.), o por inspección, y se selecciona el mejor alineamiento (es decir, que resulte en el porcentaje más alto de homología en la ventana de comparación) generados por los distintos procedimientos.

“Sustituciones de aminoácidos conservadoras” se refiere a la capacidad de intercambio entre restos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácido que tiene cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tiene cadenas laterales alifáticas-hidroxilo es de serina y treonina; un grupo de aminoácidos que tiene cadenas laterales que contienen amida es de asparagina y glutamina; un grupo de aminoácidos que tiene cadenas laterales aromáticas es de fenilalanina, tirosina, y triptófano;

un grupo de aminoácidos que tiene cadenas laterales básicas es de cisteína e histidina; y un grupo de aminoácidos que tiene cadenas laterales que contienen azufre es de cisteína y metionina. Los grupos de sustitución conservadora de aminoácidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina, y asparagina-glutamina.

5 Los términos “fragmento”, “derivado” y “análogo” cuando se refieren a un polipéptido de referencia comprende un polipéptido que mantiene al menos una función o actividad biológica que es al menos esencialmente la misma que la del polipéptido de referencia. Además, los términos “fragmento”, “derivado” o “análogo” se ejemplifica por una molécula “pro-forma”, tal como una proproteína de baja actividad que se puede modificar por escisión para producir una enzima madura con significativamente mayor actividad.

10 El término “gen” significa el segmento de ADN implicado en la producción de una cadena de polipéptido; incluye regiones precedentes y siguientes a la región codificante (líder y remolque) así como secuencias de intervención (intrones) entre segmentos codificantes individuales (exones).

15 El término “heterólogo” significa que una secuencia de ácido nucleico de cadena sencilla es incapaz de hibridarse con otra secuencia de ácido nucleico de cadena sencilla o su complementaria. Por lo tanto, áreas de heterología significa que hay áreas de polinucleótidos o polinucleótidos que tienen áreas o regiones en su secuencia que son incapaces de hibridarse con otro ácido nucleico o polinucleótido. Dichas regiones o áreas son, por ejemplo, áreas de mutaciones.

20 El término “homólogo” u “homeólogo” significa que una secuencia de ácido nucleico de cadena sencilla puede hibridarse con una secuencia de ácido nucleico de cadena sencilla complementaria. El grado de hibridación puede depender de varios factores que incluyen la cantidad de identidad entre las secuencias y las condiciones de hibridación tal como a temperatura y la concentración de sales como se expone anteriormente. Preferentemente, la región de identidad es mayor de aproximadamente 5 pb, más preferentemente la región de identidad es mayor de 10 pb.

25 El término “idéntico” o “identidad” significa que dos secuencias de ácido nucleico tienen la misma secuencia o una secuencia complementaria. Por lo tanto, “áreas de identidad” significa que regiones o áreas de un polinucleótido o el polinucleótido completo son idénticas o complementarias con áreas de otro polinucleótido o el polinucleótido.

30 El término “aislado” significa que el material se retira de su entorno original (por ejemplo, el entorno natural si es de origen natural). Por ejemplo, un polinucleótido de origen natural o enzima presente en un animal vivo no está aislado, pero el mismo polinucleótido o enzima, separado de alguno o todos los materiales coexistentes en el sistema natural, está aislado. Dichos polinucleótidos pueden ser parte de un vector y/o dichos polinucleótidos o enzimas pueden ser parte de una composición, y aun estar aislados ya que ese dicho vector o composición no es parte de su entorno natural.

35 Por “ácido nucleico aislado” se entiende un ácido nucleico, por ejemplo, una molécula de ADN o ARN, que no está inmediatamente contigua a las secuencias 5' y 3' flanqueantes con las que está normalmente inmediatamente contigua cuando está presente en el genoma de origen natural del organismo del que se deriva. La expresión por lo tanto describe, por ejemplo, un ácido nucleico que se incorpora en un vector, tal como un plásmido o vector vírico; un ácido nucleico que se incorpora en una genoma de una célula heteróloga (o el genoma de una célula homóloga, pero en un sitio diferente de su origen natural); y un ácido nucleico que existe como una molécula separada, por ejemplo, un fragmento de ADN producido por una amplificación por PCR o una digestión por una enzima de restricción, o una molécula de ARN producida por una transcripción *in vitro*. La expresión también describe un ácido nucleico recombinante que forma parte de un gen híbrido que codifica secuencias polipeptídicas adicionales que se pueden utilizar, por ejemplo, en la producción de una proteína de fusión.

45 “Ligado” se refiere al proceso de formación de enlaces fosfodiéster entre las cadenas de ácidos nucleicos (Sambrook y col, 1982, p. 146; Sambrook, 1989). La ADN ligasa puede unir juntas dos cadenas de ADN que tienen roturas de cadena sencilla (una rotura en ambas cadenas complementarias de ADN). La alternativa, una rotura de doble cadena, se fija por un tipo diferente de ADN ligasa que utiliza la cadena complementaria como matriz, pero sigue necesitando una ADN ligasa para crear el enlace fosfodiéster final para reparar completamente en ADN. A menos de que se proporcione otra cosa, el ligado se puede conseguir utilizando tampones y condiciones conocidos con 10 unidades de ADN ligasa T4 (“ligasa”) por 0,5 mg de cantidades aproximadamente equimolares de fragmentos de ADN que se van a ligar. Los “productos que no están ligados” se refiere a que no se forman enlaces fosfodiéster entre los extremos de un ácido nucleico obtenido por amplificación del polinucleótido matriz de doble cadena circular completo utilizando los cebadores.

55 El término “mutaciones” se define como cambios en la secuencia de una secuencia de ácido nucleico de tipo silvestre o parental o cambios en la secuencia de un péptido. Dichas mutaciones pueden ser mutaciones puntuales tales como transiciones o transversiones. Una mutación puede ser un cambio de uno o más nucleótidos o de secuencias de aminoácidos codificadas. Las mutaciones pueden ser eliminaciones, inserciones o duplicaciones.

Como se utiliza en el presente documento, la secuencia de nucleótido “N,N,N” degenerada representa tripletes, en los que “N” puede ser A, C, G o T.

La expresión “de origen natural” como se utiliza en el presente documento al aplicarse al objeto se refiere al hecho de que un objeto se puede encontrar en la naturaleza. Por ejemplo, una secuencia de polipéptido o polinucleótido que está presente en un organismo (incluyendo virus) que se puede aislar de una fuente en la naturaleza y que no se ha modificado intencionadamente por el hombre en el laboratorio es de origen natural. En general, la expresión de origen natural se refiere a un objeto que está presente en un individuo no patológico (no enfermo), tal como sería típico en la especie.

Como se utiliza en el presente documento, una “molécula de ácido nucleico” está compuesta por al menos una base o un par de bases, dependiendo de si es de cadena sencilla o de doble cadena, respectivamente. Además, una molécula de ácido nucleico puede pertenecer exclusiva o químicamente a cualquier grupo de moléculas que contiene nucleótidos, como se ejemplifica por, pero sin limitarse a los siguientes grupos de moléculas de ácido nucleico: ARN, ADN, ácidos nucleicos genómicos, ácidos nucleicos no genómicos, ácidos nucleicos de origen natural y no natural, y ácidos nucleicos sintéticos. Esto incluye, a modo de ejemplo no limitante, los ácidos nucleicos asociados a cualquier orgánulo, tal como la mitocondria, el ARN ribosómico, y moléculas de ácido nucleico comprendidas químicamente en uno o más componentes que no son de origen natural junto con componentes de origen natural. Adicionalmente, una “molécula de ácido nucleico” puede contener en parte uno o más componentes no basados en nucleótidos como se ejemplifica, pero no se limita a, aminoácidos y azúcares. Por lo tanto, a modo de ejemplo, pero sin limitación, una ribozima que está basada en parte en nucleótidos y en parte basada en proteínas se considera una “molécula de ácido nucleico”. Además, a modo de ejemplo, pero sin limitación, una molécula de ácido nucleico que está marcada con un resto detectable, tal como un marcador radioactivo o alternativa no radioactivo, se considera igualmente una “molécula de ácido nucleico”.

Las expresiones “secuencia de ácido nucleico que codifica” o una “secuencia de ADN codificante de” o una “secuencia de nucleótidos que codifica” una proteína o polipéptido particular se refiere a una secuencia de ADN que se transcribe y traduce en una proteína o polipéptido cuando se sitúa bajo el control de secuencias reguladoras apropiadas. Una “secuencia promotora” es una región reguladora de ADN capaz de unir la ARN polimerasa en una célula e iniciar la transcripción de una secuencia codificante corriente abajo (en dirección 3’). El promotor es parte de la secuencia de ADN. Esta región de secuencia tiene un codón de inicio en su extremo 3’. La secuencia promotora incluye el número mínimo de bases con los elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Sin embargo, después de que la ARN polimerasa se une a la secuencia y la transcripción se inicia en el codón de inicio (en el extremo 3’ con el promotor), la transcripción progresa corriente abajo en la dirección 3’. En la secuencia promotora se encontrará el sitio de inicio de la transcripción (convenientemente definido por mapeo con nucleasa S1) así como los dominios de unión proteica (secuencias de consenso) responsables de la unión de la ARN polimerasa.

Las expresiones “ácido nucleico que codifica una proteína o péptido” o “ADN que codifica una proteína o péptido” o “polinucleótido que codifica una proteína o péptido” y otras expresiones sinónimas engloban un polinucleótido que incluye solamente la secuencia codificante para la proteína o péptido, así como un polinucleótido que incluye la secuencia codificante y/o no codificante adicional.

En consecuencia, en una realización no limitante, una “biblioteca de ácidos nucleicos” está comprendida por una colección basada en vectores de una o más moléculas de ácido nucleico. En otra realización preferida, una “biblioteca de ácidos nucleicos” está comprendida por una colección combinada de moléculas de ácido nucleico que está en parte basada en vectores y en parte basada en no vectores. Preferentemente, la colección de moléculas que comprenden una biblioteca se puede buscar y separar de acuerdo con especies de moléculas de ácidos nucleicos individuales.

Un “oligonucleótido” (o su sinónimo “oligo”) se refiere a cualquier polidesoxinucleótido de cadena sencilla o dos cadenas complementarias de polidesoxinucleótido que se puede sintetizar químicamente. Dichos oligonucleótidos sintéticos pueden tener o no tener un fosfato 5’. Los que no, no se ligarán a otro oligonucleótido sin la adición de un fosfato con ATP en presencia de una cinasa. Un oligonucleótido sintético se ligará a un fragmento que no se ha desfosforilado.

Como se utiliza en el presente documento, la expresión “conjunto de polinucleótidos parentales” es un conjunto compuesto por una o más especies de polinucleótido distintas. Habitualmente esta expresión se utiliza en referencia a un conjunto de progenie de polinucleótidos que se obtiene preferentemente por mutagenización del conjunto de parentales, en cuyo caso los términos “parental”, “comienzo” y “matriz” se utilizan de manera intercambiable.

Como se utiliza en el presente documento la expresión “condiciones fisiológicas” se refiere a la temperatura, pH, fuerza iónica, viscosidad, y parámetros bioquímicos similares que son compatibles con un organismo viable, y/o que existen normalmente intracelularmente en una célula de levadura o célula de mamífero viables cultivadas. Por ejemplo, las condiciones intracelulares en un cultivo celular de levaduras en condiciones típicas de cultivo son condiciones fisiológicas. Las condiciones de reacción *in vitro* adecuadas para los cocteles de transcripción *in vitro* son en general condiciones fisiológicas. En general, las condiciones fisiológicas *in vitro* comprenden 50-200 mM de NaCl o KCl, pH 6,5-8,5, 20-45 °C y 0,001-10 mM de un catión divalente (por ejemplo, Mg⁺⁺, Ca⁺⁺); preferentemente aproximadamente 150 mM de NaCl o KCl, pH 7,2-7,6, 5 mM de un catión divalente, y a menudo incluyen un 0,01-1,0 por ciento de proteína no específica (por ejemplo, BSA). A menudo puede estar presente un detergente no iónico

(Tween, NP-40, Triton X-100), habitualmente en aproximadamente un 0,001 a 2 %, normalmente un 0,05-0,2 % (v/v). Se pueden seleccionar condiciones acuosas particulares por el facultativo de acuerdo con procedimientos convencionales. Para una guía general, se pueden aplicar las siguientes condiciones acuosas tampón: 10-250 mM de NaCl, 5-50 mM de Tris-HCl, pH 5-8, con adición opcional de cationes divalentes y/o quelantes metálicos y/o detergentes no iónicos y/o fracciones de membrana y/o agentes antiespumantes y/o centelleantes.

Se utilizó en el presente documento la convención de referencia (5' a 3') para describir la secuencia de polinucleótidos de doble cadena.

La expresión "polinucleótidos relacionados" significa que las regiones o áreas de los polinucleótidos son idénticas y las regiones o áreas de los polinucleótidos son heterólogas.

"Hibridación específica" se define en el presente documento como la formación de híbridos entre un primer polinucleótido y un segundo polinucleótido (por ejemplo, un polinucleótido que tiene una secuencia distinta pero sustancialmente idéntica al primer polinucleótido), en el que secuencias de polinucleótido sustancialmente no relacionadas no forman híbridos en la mezcla.

"Condiciones de hibridación rigurosas" significa que la hibridación se producirá solo si hay al menos un 90 % de identidad, preferentemente al menos un 95 % de identidad y más preferentemente al menos un 97 % de identidad entre las secuencias. Véase, Sambrook y col, 1989.

La expresión "de tipo silvestre" significa que el polinucleótido no comprende ninguna mutación. Una proteína "de tipo silvestre" significa que la proteína será activa a un nivel de actividad que se encuentra en la naturaleza y comprenderá la secuencia de aminoácidos que se encuentra en la naturaleza.

Las fuentes de polinucleótidos originales se pueden aislar de organismos individuales ("aislados"), colecciones de organismos que se han cultivado en un medio definido ("enriquecimiento de cultivos"), o más preferentemente, organismos no cultivados ("muestras ambientales") El uso de una estrategia independiente del cultivo para derivar polinucleótidos que codifican nuevas bio-actividades de muestras ambientales es más preferible ya que permite el acceso a recursos de biodiversidad no protegidos.

Los microorganismos de los que se puede preparar el polinucleótido incluyen microorganismos procariotas, tales como Eubacterias y Arqueobacterias, y microorganismos eucariotas inferiores tales como hongos, algunas algas y protozoos. Los polinucleótidos se pueden aislar de muestras ambientales en cuyo caso el ácido nucleico se puede recuperar sin el cultivo de un organismo o recuperarse de uno o más organismos cultivados. En un aspecto, dichos microorganismos pueden ser extremófilos, tales como hipertermófilos, psicrófilos, psicrotrofos, halofitos, barófilos o acidófilos.

Los polinucleótidos seleccionados y aislados como se ha descrito anteriormente en el presente documento se introducen en una célula huésped adecuada. Los polinucleótidos seleccionados están preferentemente ya en un vector que incluye las secuencias de control apropiadas. La célula huésped puede ser una célula eucariota superior, tal como una célula de mamífero o una célula eucariota inferior, tal como una célula de levadura, o preferentemente, la célula huésped puede ser una célula procariota, tal como una célula bacteriana. La introducción de la construcción en la célula huésped se puede efectuar por transfección en fosfato cálcico, transfección mediada por DEAE-dextrano, o electroporación (Davis y col, 1986).

Como ejemplos representativos de huéspedes apropiados, hay que mencionar: células bacterianas, tales como *E. coli* y *Pseudomonas fluorescens*; bacteriófagos; células fúngicas tales como levaduras, *Pichia pastoris* y *Aspergillus niger*, células de insecto tales como *Drosophila* S2 y *Spodoptera* Sf9; células animales tales como CHO, COS, o melanoma de *Bowes*, adenovirus; y células vegetales. La biblioteca TMCA se puede hacer, por ejemplo, en células de *E. coli* en forma de plásmido, después la biblioteca se puede introducir adicionalmente en otros huéspedes. La selección de un huésped apropiado se considera que está en el ámbito de los expertos en la técnica a partir de las enseñanzas del presente documento.

En referencia particular a distintos sistemas de cultivo celular que se pueden emplear para expresar una proteína recombinante, los ejemplos de sistemas de expresión en mamíferos incluyen las líneas COS-7 de fibroblastos de riñón de mono, descritos en "SV40-transformed simian cells support the replication of early SV40 mutants" (Gluzman, 1981), y otras líneas celulares capaces de expresar un vector compatible, por ejemplo, las líneas celulares C127, 3T3, CHO, HeLa y BHK. Los vectores de expresión en mamíferos comprenderán un origen de replicación, un promotor adecuado y un amplificador, y también cualquier sitio de unión al ribosoma necesario, sitio de poliadenilación, sitios donante y receptor de corte y empalme, secuencias de terminación de la transcripción, y secuencias no transcritas flanqueantes 5'. Las secuencias de ADN derivadas del corte y empalme en SV40, y los sitios de poliadenilación se pueden utilizar para proporcionar los elementos genéticos no transcritos necesarios.

Las células huésped que contienen los polinucleótidos de interés se pueden cultivar en medios nutritivos convencionales modificados según sea apropiado por activación de promotores, selección de transformantes o amplificación de genes. Las condiciones de cultivo, tales como la temperatura, pH y similares, son los que se utilizan previamente con la célula huésped seleccionada para la expresión, y serán evidentes para el experto habitado.

Como ejemplos representativos de vectores de expresión que se pueden utilizar se pueden mencionar partículas víricas, baculovirus, fagos, plásmidos, fagémidos, cósmidos, fósmidos, cromosomas bacterianos artificiales, ADN vírico (por ejemplo, vaccinia, adenovirus, virus de viruela aviar, pseudorrabia y derivados de SV40), cromosomas artificiales basados en P1, plásmidos de levadura, cromosomas artificiales de levadura, y cualquier otro vector específico para los huéspedes específicos de interés (tales como *Bacillus*, *Aspergillus* y levaduras). Por lo tanto, por ejemplo, el ADN puede incluirse en uno cualquiera de una variedad de vectores de expresión para la expresión de un polipéptido. Dichos vectores incluyen secuencias cromosómicas, no cromosómicas y sintéticas de ADN. Se conoce una gran cantidad de vectores adecuados por los expertos en la técnica, y están disponibles en el mercado. Los siguientes vectores se proporcionan a modo de ejemplo; Bacterianos: vectores pQE (Qiagen), plásmidos pBluescript, vectores pNH, (vectores lambda-ZAP (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eucariotas: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Sin embargo, cualquier otro plásmido u otro vector se puede utilizar a condición de que se puedan replicar y sean viables en el huésped. Se pueden emplear vectores de bajo número de copias o alto número de copias en la presente invención.

Un tipo preferido de vector para su uso en la presente invención contiene un origen de replicación factor-f. El factor-f (o factor de fertilidad) en *E. coli* es un plásmido que efectúa una transferencia de alta frecuencia de sí mismo durante la conjugación y menos frecuentemente del propio cromosoma bacteriano. Una realización particularmente preferida es el uso de vectores de clonación, a los que se hace referencia como "fósmidos" o vectores de cromosomas bacterianos artificiales (BAC). Estos se derivan del factor-f de *E. coli* que es capaz de integrar establemente grandes segmentos de ADN genómico. Cuando se integra en el ADN a partir de una muestra ambiental no cultivada, hace posible conseguir grandes fragmentos genómicos en forma de una "biblioteca de ADN ambiental".

Otro tipo de vector preferido para su uso en la presente invención es un vector cósmido. Los vectores cósmidos se diseñaron originalmente para clonar y propagar grandes segmentos de ADN genómico. La clonación en vectores cósmidos se describen en detalle en "Molecular Cloning: A laboratory Manual" (Sambrook y col, 1989).

La secuencia de ADN en el vector de expresión está unida operativamente a una secuencia(s) de control de la expresión (promotor) adecuada para dirigir la síntesis de ARN. Los promotores bacterianos particulares nominados incluyen lacI, lacZ, T3, T7, gpt, lambda P_R, P_L y trp. Los promotores de eucariotas incluyen el CMV inmediato temprano, HSV de timidina cinasa, SV40 temprano y tardío, LTR de retrovirus y metalotioneína-I de ratón. La selección del vector y promotor adecuados está a nivel del experto en la técnica. El vector de expresión también contiene un sitio de unión al ribosoma para el inicio de la traducción y un terminador de la transcripción. El vector puede incluir también secuencias apropiadas para la amplificación de la expresión. Las regiones promotoras se pueden seleccionar de cualquier gen deseado utilizando vectores CAT (cloranfenicol transferasa) u otros vectores con indicadores genéticos.

Además, los vectores de expresión contienen preferentemente uno o más genes marcadores indicadores para proporcionar un rasgo fenotípico para la selección de las células huésped transformadas tales como la dihidrofolato reductasa o resistencia a neomicina para los cultivos celulares de eucariotas, o tal como la resistencia a tetraciclina o ampicilina en *E. coli*.

En general, los vectores de expresión recombinante incluirán orígenes de replicación e indicadores genéticos que permiten la transformación de la célula huésped, por ejemplo, el gen de resistencia de *E. coli* y *S. cerevisiae*, el gen TRP 1, y un promotor derivado de un gen altamente expresado para dirigir la transcripción de una secuencia estructural corriente abajo. Dichos promotores se pueden derivar de operones que codifican enzima glicolíticas tales como 2-fosfoglicerato cinasa (PGK), factor-alfa, fosfatasa ácida, o proteínas de choque térmico entre otras. La secuencia estructural heteróloga se ensambla en una fase apropiada con secuencias de inicio y terminación de la traducción, y preferentemente, una secuencia líder capaz de dirigir la secreción de la proteína traducida en el espacio periplásmico o el medio extracelular.

La estrategia de clonación permite la expresión mediante ambos promotores el dirigido por el vector y el endógeno; la promoción del vector puede ser importante en la expresión de genes cuyo promotor endógeno no funciona en *E. coli*.

El ADN aislado o derivado de microorganismos se puede insertar preferentemente en un vector o un plásmido antes de sondear el ADN seleccionado. Dichos vectores o plásmidos son preferentemente los que contienen secuencias reguladoras de la expresión, que incluyen promotores, amplificadores y similares. Dichos polinucleótidos pueden ser parte de un vector y/o una composición y estar aún aislados, si dicho vector o composición no es parte de su entorno natural. Los fagos y plásmidos preferidos particularmente y los procedimientos para la introducción empaquetamiento en ellos se describen en detalle en el protocolo expuesto en el presente documento.

Se puede utilizar cualquier fuente de ácido nucleico como ácido nucleico de partida (también definido como "un polinucleótido matriz"). Por lo tanto, el procedimiento puede emplear ADN o ARN que incluye ARN mensajero, cuyo ADN o ARN puede ser de cadena sencilla, y preferentemente de cadena doble. Además, se puede utilizar un híbrido de ADN-ARN que contenga una cadena de cada. La secuencia de ácido nucleico puede tener distintas longitudes dependiendo del tamaño de la secuencia de ácido nucleico que se va a mutar. Preferentemente la secuencia de ácido nucleico específica tiene de 50 a 50000 pares de bases, y más preferentemente de 50 a 11000 pares de

bases.

5 El ácido nucleico puede obtenerse de cualquier fuente, por ejemplo, de plásmidos tales como pBR322, de ADN o ARN clonado de ADN o ARN natural de cualquier fuente incluyendo bacterias, levaduras, virus y organismos superiores tales como plantas o animales. El ADN o ARN se puede extraer de la sangre o material tisular. El polinucleótido matriz puede obtenerse por amplificación utilizando la reacción en cadena de polinucleótidos (PCR, véase la Pat. de EE. UU. N.º 4.683.202 y Pat. de EE. UU. N.º 4.683.195). De manera alternativa, el polinucleótido puede estar presente en un vector presente en una célula y se puede obtener suficiente ácido nucleico cultivando la célula y extrayendo el ácido nucleico de la célula por procedimientos conocidos en la técnica.

10 La pequeña población inicial de las secuencias de ácido nucleico específicas que tienen mutaciones se puede crear por varios procedimientos diferentes. Las mutaciones se pueden crear por PCR con tendencia a errores. La PCR con tendencia a errores utiliza condiciones de polimerización de baja fiabilidad para introducir un bajo nivel de mutaciones puntuales aleatoriamente en una secuencia larga. De manera alternativa, se pueden introducir mutaciones en el polinucleótido matriz por mutagénesis dirigida a oligonucleótido. En la mutagénesis dirigida a oligonucleótido, se retira una corta secuencia de polinucleótido del polinucleótido utilizando digestión por enzimas de restricción y se sustituye con un polinucleótido sintético en el que se han alterado algunas bases de la secuencia original. La secuencia de polinucleótido se puede alterar también por mutagénesis química. Los mutágenos químicos incluyen, por ejemplo, bisulfito sódico, ácido nítrico, hidroxilamina, hidracina o ácido fórmico. Otros agentes que son análogos de precursores de nucleótidos incluyen nitrosoguanidina, 5-bromouracilo, 2-aminopurina, o acridina. En general, estos agentes se añaden a la reacción de PCR en lugar del precursor de nucleótido mutando de esta manera la secuencia. También se pueden utilizar agentes intercaladores tales como la proflavina, acriflavina, quinacrina y similares. La mutagénesis aleatoria de la secuencia de polinucleótidos también se puede conseguir por radiación con rayos X o luz ultravioleta. En general, los polinucleótidos plasmídicos mutagenizados de esta manera se introducen en *E. coli* y se propagan como un agrupamiento o biblioteca de plásmidos híbridos.

25 De manera alternativa la pequeña población mixta de ácidos nucleicos específicos se puede encontrar en la naturaleza y puede consistir en diferentes alelos del mismo gen o el mismo gen de diferentes especies relacionadas (es decir, genes equivalentes). De manera alternativa, pueden ser secuencias de ADN relacionadas que se encuentran en una especie, por ejemplo, genes de inmunoglobulina.

Una vez que se genera la población mixta de secuencias de ácido nucleico, los polinucleótidos se pueden utilizar directamente o insertarse en un vector de clonación adecuado, utilizando técnicas bien conocidas en la técnica.

30 La elección del vector depende del tamaño de la secuencia de polinucleótido y la célula huésped que es va a emplear en los procedimientos de la presente invención. Las matrices de la presente invención pueden ser plásmidos, fagos, cósmidos, fagémidos, virus (por ejemplo, retrovirus, virus de parainfluenza, herpesvirus, reovirus, paramixovirus, y similares), o partes seleccionadas de los mismos (por ejemplo, proteína de revestimiento, glicoproteína fija, proteína de la cápside). Por ejemplo, los cósmidos y fagémidos son preferidos cuando la secuencia de ácido nucleico específica que se va a mutar es más grande debido a que estos vectores son capaces de propagar establemente grandes polinucleótidos.

Por simplicidad, el procedimiento TMCA de la invención se explicará con el intento de ensamblar seis mutaciones puntuales en seis sitios diferentes.

40 Primero, se diseñan y se sintetizan seis cebadores. Cada cebador contiene una mutación puntual en comparación con la secuencia de tipo silvestre. Se diseñan tres oligos como cebadores directos y se diseñan tres como cebadores inversos para hibridarse al gen (Figura 5). Los seis oligos se mezclan juntos y se utilizan para llevar a cabo las reacciones de TMCA en las condiciones detalladas en los Ejemplos. Después, las reacciones de TMCA terminadas se verifican mediante geles de agarosa para determinar si las reacciones son satisfactorias. La enzima de restricción *DpnI* se añade a las reacciones de TMCA para destruir el ADN matriz circular. Con el fin de que la *DpnI* trabaje, el ADN matriz tiene que ser de la *E. Coli* huésped que puede metilar el ADN. Las reacciones tratadas

45 El procedimiento de la invención no se limita a seis sitios. Se puede ensamblar un número mayor o menor de posiciones por este procedimiento. También no se limita a un único cambio en una posición. Se pueden diseñar múltiples cebadores para cubrir diferentes cambios en la misma posición, con un único cambio en cada cebador. Se utilizó *E. Coli* para la demostración; sin embargo, otros huéspedes bacterianos servirían para este procedimiento. El procedimiento de la invención no solo puede introducir una mutación puntual, también puede hacer eliminaciones o inserciones o múltiples mutaciones con cebadores degenerados.

Las reacciones de TMCA se pueden alterar con la concentración de cebadores, la T_m del cebador (temperatura de hibridación a una matriz), ADN polimerasa, concentración de matriz, combinación de los cebadores y diferentes huéspedes para controlar cómo se ensamblan los cambios en los diferentes sitios.

55 El ensamblaje puede producirse *in vitro* o *in vivo* o una combinación de ambos. La Fig. 5 ilustra un mapa de los cebadores hibridados en cada gen. El uso principal del procedimiento de la invención es para el reagrupamiento de mutantes en GSSMSM. Sin embargo, el procedimiento de la invención también puede ser útil, por ejemplo, para otras aplicaciones enumeradas posteriormente.

1. La TMCA se puede utilizar para hacer cambios específicos en un gen, incluyendo mutación, eliminación e inserción.
2. La TMCA se puede utilizar para hacer una variante genética específica basada en un gen de tipo silvestre.
3. La TMCA puede utilizarse para combinar una mutación, eliminación o inserción.
- 5 4. La TMCA se puede utilizar para hacer una biblioteca de combinaciones de mutación, eliminación o inserción de una manera controlable.
5. La TMCA se puede utilizar para hacer una biblioteca de combinaciones GSSMSM de sitios múltiples.

En general, la presente invención proporciona un procedimiento para la producción de una pluralidad de polinucleótidos de una progenie que tiene diferentes combinaciones de distintas mutaciones en múltiples sitios en el que las mutaciones son cambios de uno o más nucleótidos de la secuencia de una secuencia de ácido nucleico parental. El procedimiento se puede llevar a cabo en parte por una combinación de al menos una o más de las siguientes etapas:

Obtención de información de secuencia de un polinucleótido (“primero” o “matriz”). Por ejemplo, la secuencia puede ser una secuencia de tipo silvestre, de tipo silvestre mutada, o de origen no natural. La información de secuencia puede ser del polinucleótido completo o de regiones particulares de interés, tal como una secuencia que codifica un sitio para la unión, especificidad de unión, catálisis, o especificidad de sustrato. El polinucleótido puede comprender una secuencia tal como la fase de lectura abierta, un gen, una secuencia que codifica un polipéptido, o una secuencia que codifica una enzima, con o sin secuencia de señal o secreción.

Identificación de tres o más mutaciones de interés a lo largo de la secuencia del polinucleótido, tal como mutaciones en las posiciones 3, 4, 5, 6, 8, 10, 12, 20 o más. Las mutaciones pueden estar a nivel de la secuencia de los polinucleótidos o de la secuencia de aminoácidos codificada por la secuencia de polinucleótido, por ejemplo, los codones. Las posiciones se pueden determinar por la posición absoluta o por el contexto de restos circundantes u homología. Las secuencias que flanquean las posiciones de mutación a cada lado son preferentemente conocidas. Cada posición de mutación puede contener dos o más mutaciones, tal como por diferentes aminoácidos. Dichas mutaciones se pueden identificar utilizando la Mutagénesis de Saturación del Sitio Genético (GSSM), como se ha descrito anteriormente, y en las Patentes de EE. UU. N.º 6.171.820, N.º 6.562.594, o N.º 6.764.835.

Provisión de cebadores que comprenden las mutaciones de interés con respecto a la secuencia matriz. Los cebadores pueden ser oligonucleótidos sintéticos. Preferentemente, se proporciona un cebador para cada mutación de interés. Las mutaciones pueden ser cambios en uno o más nucleótidos o secuencias de aminoácidos codificadas, inserciones o eliminaciones. Por lo tanto, una posición que tiene 3 mutaciones de interés puede utilizar 3 cebadores en esa posición. El cebador también se puede proporcionar como un agrupamiento de cebadores que contienen una posición degenerada de manera que la mutación de interés está en el intervalo de cualquier nucleótido o aminoácido de origen natural, o un subconjunto de ese intervalo. Por ejemplo, puede proporcionarse un agrupamiento de cebadores que favorezca las mutaciones de restos de aminoácidos alifáticos.

Los cebadores se pueden preparar como cebadores directos o inversos, preferentemente al menos un cebador directo y al menos un cebador inverso, y más preferentemente un número de cada uno relativamente equilibrado (por ejemplo, 3 directos y 4 inversos). Los tres cebadores directos se pueden seleccionar por los relativamente adyacentes, con cebadores inversos adyacentes similarmente, por ejemplo, 1F, 2F, 3F, 4R, 5R, 6R, 7R. Cuando las mutaciones se posicionan estrechamente juntas, puede ser conveniente el uso de cebadores que contienen mutaciones para más de una posición o combinaciones diferentes de mutaciones en múltiples posiciones.

Provisión de un polinucleótido que contiene el polinucleótido matriz. El polinucleótido es preferentemente circular, más preferentemente superenrollado, tal como un plásmido o vector para la clonación, secuenciación o expresión. El polinucleótido puede ser de cadena sencilla (“ssADN”), y preferentemente de doble cadena (“dsADN”). Por ejemplo, el procedimiento TCMA somete el dsADN (“sc”) superenrollado matriz a una etapa de calentamiento a 95 °C durante 1 min, la matriz no se convierte en un ssADN (véase, Levy, NAR, 28(12):e57(i-vii) (2000), muestra que el calentamiento dl dsADN sc a 95 °C durante 5 min no produce moléculas de ssADN y es reversible si las moléculas se enfrían después del calentamiento (páginas ii-iii, fig. 2)).

Adición de los cebadores al polinucleótido matriz en una mezcla de reacción en condiciones que permitan a los cebadores hibridarse con el polinucleótido. Preferentemente, los cebadores se añaden al polinucleótido en una única mezcla de reacción, pero se pueden añadir en múltiples reacciones de acuerdo con un diseño experimental.

Realización de una extensión por polimerasa de los cebadores, preferentemente permitiendo que la extensión progrese completamente alrededor de una molécula matriz circular. Los productos de interés (como se define en el presente documento, “progenie” o “polinucleótido modificado extendido”) se pueden amplificar por medios convencionales.

Los productos se pueden analizar en cuanto a longitud, secuencia, propiedades de ácido nucleico deseadas, o que se expresen como polinucleótidos y/o polipéptidos. Otros procedimientos de análisis incluyen la hibridación in situ, exploración de secuencia o exploración de la expresión. El análisis puede incluir una o más rondas de exploración y selección de una propiedad deseada.

Los productos también se pueden transformar en una célula u otro sistema de expresión, tal como un sistema Libre de células. El sistema libre de células puede contener enzimas relacionadas con la replicación, reparación, recombinación, transcripción de ADN o para la traducción. Los huéspedes ejemplares incluyen células y líneas celulares bacterianas, de levadura, plantas y animales, e incluyen *E. coli*, *Pseudomonas fluorescens*, *Pichia pastoris* y *Aspergillus niger*. Por ejemplo, se pueden utilizar las cepas XL1-Blue o Stb12 de *E. coli* como huéspedes. Cuando se utiliza la *E. coli* con *DpnI* (que se puede utilizar para retirar la matriz no deseada después de la reacción), la matriz de ADN puede ser de una *E. coli* huésped que puede metilar el ADN. Las células se pueden utilizar para la expresión de los polinucleótidos de la progenie.

La expresión de los productos polinucleotídicos y polipeptídicos se puede conseguir de las células y se analiza en cuanto a su longitud, secuencia, propiedades de ácido nucleico deseadas, o que se expresan como polipéptidos. El análisis puede incluir una o más rondas de exploración y selección de una propiedad deseada.

El procedimiento de la invención puede utilizarse con los mismos cebadores o diferentes en condiciones de reacción diferentes para promover productos que tengan diferentes combinación o número de mutaciones, tal como en las condiciones 1A, 7A y 13A ilustradas en los ejemplos.

Llevando a cabo el procedimiento descrito anteriormente, la divulgación también proporciona uno o más polinucleótidos producidos por el procedimiento, que se pueden explorar o seleccionar en cuanto a una propiedad deseada. Se pueden expresar uno o más de los polinucleótidos de la progenie como polipéptidos, y opcionalmente explorarse o seleccionarse por una propiedad deseada. Por lo tanto, la divulgación proporciona polinucleótidos y polipéptidos producidos por el procedimiento de la invención, así como bibliotecas de dichos polinucleótidos y polipéptidos. La divulgación proporciona además la exploración de bibliotecas explorando o seleccionando la biblioteca para obtener uno o más polinucleótidos o polipéptidos.

En un aspecto de la invención, un procedimiento preferido de producción de una pluralidad de polinucleótidos modificados, que tienen combinaciones diferentes de distintas mutaciones en múltiples sitios, en el que las mutaciones son cambios de uno o más nucleótidos de la secuencia de una secuencia de ácido nucleico parental, dicho procedimiento comprende:

(a) añadir al menos tres cebadores de un polinucleótido matriz de doble cadena en una única mezcla de reacción, en el que los al menos tres cebadores no están solapados, y en el que cada uno de los al menos tres cebadores comprende al menos una mutación diferentes de los otros cebadores, en el que al menos un cebador es un cebador directo y que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz, y

(b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de los al menos tres cebadores.

En otro aspecto de la invención, se transforma una célula con la pluralidad de productos extendidos que no se han tratado con una ligasa. En otro aspecto de la invención, la pluralidad de polinucleótidos modificados extendidos se recuperan de la célula. En otra realización, la pluralidad de polinucleótidos modificados extendidos recuperada se analiza, por ejemplo, expresando al menos uno de los polinucleótidos modificados extendidos de la pluralidad y analizando el polipéptido expresado del mismo. En otra realización, se selecciona la pluralidad de polinucleótidos modificados extendidos comprende las mutaciones de interés.

En una realización, el polinucleótido matriz es un ADN circular, por ejemplo, un ADN plasmídico o vector. En otra realización, el ADN circular es un ADN superenrollado.

En otro aspecto, se puede obtener la información de secuencia con respecto al polinucleótido matriz, y se pueden identificar tres o más mutaciones de interés a lo largo del polinucleótido matriz. En otra realización, los productos obtenidos por la extensión con la polimerasa se pueden analizar antes de transformar una célula con la pluralidad de productos modificados extendidos.

En otro aspecto de la invención, los productos obtenidos por extensión con la polimerasa se tratan con una enzima, preferentemente una enzima de restricción, y más preferentemente una enzima de restricción *DpnI*, destruyendo de esta manera la secuencia de polinucleótido matriz. Los productos tratados se transformaron en una célula, preferentemente una célula de *E. coli*.

En otra realización, se pueden utilizar al menos dos, preferentemente al menos tres, más preferentemente al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, al menos diez, al menos once, al menos doce o más cebadores. En una realización, cada cebador comprende una única mutación puntual (Fig. 4A). En otra realización, dos cebadores directos o dos inversos comprenden un cambio diferente en la misma posición del polinucleótido matriz (Fig. 4B). en otro aspecto de la invención, al menos un cebador comprende al menos dos cambios en diferentes posiciones del polinucleótido matriz (Fig. 4C). En otra realización más, al menos un cebador comprende al menos dos cambios en diferentes posiciones y al menos dos cebadores directos o dos inversos comprenden un cambio diferente en la misma posición del polinucleótido matriz (Fig. 4D).

En una realización, los cebadores directos se agrupan en un grupo directo y los cebadores inversos se agrupan en un grupo inverso, y los cebadores del grupo directo y los cebadores del grupo inverso, independientemente entre ellos, se normalizan para tener una concentración igual en su grupo correspondiente independientemente de las posiciones en el polinucleótido matriz, y en el que después de la normalización se añade una cantidad igual de cebadores directo o inverso a la reacción. En este procedimiento de normalización, se puede desviar una combinación de algunas posiciones. La desviación puede deberse a, por ejemplo, una concentración relativamente baja de cebador en una posición que contiene un único cebador en comparación con una posición que contiene múltiples cebadores. "Tendencia posicional" se refiere a los polinucleótidos resultantes que muestran una fuerte preferencia por la incorporación de cebadores en una única posición con respecto a las otras posiciones en su grupo de cebadores directos o inversos. Esto resulta en una combinación de polinucleótidos modificados que puede tener un alto porcentaje de mutaciones en una única posición de cebador, pero con un bajo porcentaje de mutaciones en otra posición de su grupo de cebadores directo e inverso. Esta desviación no es favorable cuando el objetivo de TMCA es generar unos polinucleótidos de la progenie que comprendan todas las combinaciones o cambios posibles de la matriz. La desviación se puede corregir, por ejemplo, normalizando los cebadores como un agrupamiento en cada posición para que sea igual. Llevando a cabo dos rondas del procedimiento TMCA se puede aumentar el rendimiento de una progenie de polinucleótidos deseada que comprende múltiples cambios en la matriz, en el que la ronda II utiliza algunas de las variantes obtenidas en la ronda I.

En otra realización, la normalización del cebador se lleva a cabo organizando los cebadores en múltiples grupos dependiendo de su localización en el polinucleótido matriz, en el que los cebadores que cubren la misma región seleccionada en la matriz están en un grupo; normalizando los cebadores agrupados en cada grupo para que tengan igual concentración; agrupando de cebadores directos de un grupo en un grupo directo y normalizando la concentración entre cada grupo de cebadores inversos para que sea igual; y añadiendo una cantidad igual de los cebadores directos e inversos agrupados en la reacción. No se ha observado una tendencia en la combinación de posiciones.

En una realización, se proporciona un conjunto de cebadores degenerados que comprende cada uno una posición degenerada, en el que la mutación de interés está en un intervalo de diferentes nucleótidos en la posición degenerada. En otra realización, se proporciona un conjunto de cebadores degenerados que comprende al menos un codón degenerado que se corresponde con al menos un codón del polinucleótido matriz, y al menos una secuencia adyacente que es homóloga a una secuencia adyacente del codón de la secuencia de polinucleótido matriz. En otra realización, el codón degenerado es N,N,N y codifica cualquiera de los 20 aminoácidos de origen natural. En otra realización, el codón degenerado codifica menos de los 20 aminoácidos naturales.

En una realización diferente, un procedimiento de producción de una pluralidad de polinucleótidos modificados que tienen diferentes combinaciones de distintas mutaciones en múltiples sitios y comprenden las mutaciones de interés, en el que las mutaciones son cambios de uno o más nucleótidos en la secuencia de una secuencia de ácido nucleico parental, comprendiendo dicho procedimiento:

(a) añadir al menos dos cebadores a un polinucleótido matriz de doble cadena en una única mezcla de reacción, en el que los al menos dos cebadores no se solapan, y en el que cada uno de los al menos dos cebadores comprenden al menos una mutación diferente de la del otro cebador, en el que al menos un cebador es un cebador directo que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz,

(b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de los al menos dos cebadores;

(c) tratar la pluralidad de polinucleótidos modificados extendidos con una enzima, destruyendo de esta manera el polinucleótido matriz,

(d) transformar los polinucleótidos modificados extendidos tratados que no se han tratado con una ligasa en una célula,

(e) recuperar la pluralidad de polinucleótidos modificados extendidos de la célula, y

(f) seleccionar la pluralidad de polinucleótidos modificados extendidos que comprenden la mutación de interés.

Los siguientes Ejemplos muestran que mutaciones únicas en múltiples sitios de una matriz o gen se pueden combinar satisfactoriamente en una única mezcla de reacción simple, que se espera que se base en los procedimientos conocidos de generación de mutaciones. La distribución de todas las combinaciones posibles del experimento refleja estrechamente el patrón de distribución del cálculo estadístico. Las reacciones se pueden hacer a medida para crear combinaciones con tendencias basándose en las necesidades.

Con la tecnología GSSMSM, la técnica TMCA no emplea cebadores complementarios que se hibridan a ambas cadenas positiva y negativa del polinucleótido matriz. La expectativa razonable de que la extensión por ciclado térmico contenía cebadores como se describe para esta invención de TMCA (grupos directo e inverso en una única reacción de ciclado térmico) sería una colección exclusiva de polinucleótidos lineales definidos por pares

- 5 individuales de los cebadores directo e inverso. La condición de ejecución del TMCA es casi idéntico a las condiciones de PCR convencional. Se debería esperar que se produzcan múltiples productos de PCR en la TMCA, en el que cada producto tiene menos del conjunto completo de mutaciones englobadas por un conjunto de cebadores utilizados en una única reacción, por ejemplo, menos de 6 mutaciones cuando se utilizan 6 cebadores y cada cebador comprende 1 mutación. Además, no se espera que los productos de PCR se transformen en las células y se amplifiquen. Sorprendentemente, el procedimiento TMCA puede generar una variante genética específica que comprende múltiples cambios en una molécula y que se puede llevar a cabo sin emplear una etapa de unión, y, por lo tanto, simplifica el procedimiento de generar múltiples mutaciones.

Ejemplos

10 **Ejemplo 1**

El protocolo ejemplar del procedimiento de TMCA se muestra a continuación:



1. Preparación de las reacciones de TMCA

Condición 1

10x de tampón Pfu	2,5 µl
DMSO	2,5 µl
dNTP (10 mM)	0,5 µl
ADN matriz (25 ng/µl)	1 µl
PfuTurbo	0,5 µl
Agua	14 µl
Cebador directo (5 µM)	2 µl
Cebador inverso (5 µM)	2 µl
Total	25 µl

Condición 2

Tampón Pfx Accu	5 µl
ADN matriz (25 g/µl)	1 µl
Pfx Accuprime	0,4 µl
Agua	37,6 µl
Cebador directo (5 µM)	3 µl
Cebador inverso (5 µM)	3 µl
Total	50 µl

Condición 3

Tampón Pfx Accu	2,5 µl
ADN matriz (25 g/µl)	1 µl
Pfx Accuprime	0,2 µl
Agua	17,3 µl
Cebador directo (5 µM)	2 µl
Cebador inverso (5 µM)	2 µl
Total	25 µl

	Ciclado		
	Robocycler	Perkin-Elmer	
Desnaturalización inicial	95 °C; 1 min	95 °C; 3 min	} 20 ciclos
Desnaturalización	95 °C; 45 s	95 °C; 45 s	
Hibridación	50 °C; 1 min	50 °C; 45 s	
Extensión	68 °C; 2 min/kb	68 °C; 2 min/kb	
Pulido	68 °C; 5 min	68 °C; 5 min	
	4 °C; Indefinidamente	4 °C; Indefinidamente	

2. Ejecución de 50 reacciones de TMCA sobre gel de agarosa para determinar si las reacciones eran satisfactorias.

5 3. Dilución de 10 de enzima de restricción *DpnI* en 3 µl de agua y 1 µl de tampón 4 (New England Biolab). Adición de 5 µl de enzima diluida en cada reacción TMCA. Incubación a 37 grados durante 4-8 horas.

4. Transformación de las reacciones tratadas con *DpnI* en células de *E. coli* mediante un protocolo de transformación convencional.

5. Las colonias resultantes se exploraron mediante secuenciación o un ensayo deseado.

10 **Ejemplo 2**

En el primer experimento, seis sitios de un gen se escogieron para ser combinados (Fig. 5). Las reacciones se prepararon utilizando cebadores directos en tres sitios y tres cebadores inversos en los otros tres sitios. Las variantes de las reacciones se identificaron y secuenciaron. Había sesenta y cuatro combinaciones posibles diferentes. En la condición 1, se prefirieron más variantes con un número menor de sitios de mutación (Fig. 6). En las condiciones 2 y 3, se prefirieron más variantes con un número mayor de sitios de mutación (Fig. 6). La distribución de todas las combinaciones posibles de los datos combinados (Total) era similar al patrón de distribución del cálculo estadístico (Fig. 6). Una curva de la Fig. 7 muestra la cobertura esperada (%) de variantes y otra curva muestra la probabilidad de cobertura completa cuando se exploran de cero a seiscientos clones. Los círculos de la curva de cobertura esperada (%) (es decir, el 78 %, 95 %, 99, y 100 %) muestra la cobertura esperada cuando se exploran 96, 192, 288 o 384 clones. Los cuadrados debajo en la curva de cobertura esperada (%) (es decir, el 70 %, 91 %, 95 %, y 98 %) muestra la cobertura real de los datos experimentales. Los datos muestran una coincidencia casi perfecta de la cobertura esperada vs la cobertura real.

Tabla 1. Conjunto de seis mutaciones

Condiciones	Exploración	Clones únicos	Clones únicos (-1x)
1A	1,5x	69 % (44/64)	69 % (44/64)
1A	3x	80 % (51/64)	80 % (51/64)
7A	1,5x	63 % (40/64)	75 % (42/64)
13A	1,5x	61 % (39/64)	67 % (43/64)
13A	3x	76 % (49/64)	80 % (51/64)
1A + 7A	3x	91 % (58/64)	91 % (58/64)
1A + 7A	4,5x	94 % (60/64)	94 % (60/64)
1A + 13A	3x	91 % (58/64)	91 % (58/64)
1A + 13A	4,5x	95 % (61/64)	95 % (61/64)
1A + 13A	6x	98 % (63/64)	98 % (63/64)

25 **Ejemplo 3**

En el segundo experimento, se escogieron cuatro sitios en un gen para combinarlos (Fig. 8). Las reacciones se prepararon utilizando cebadores directos en dos sitios y dos cebadores inversos en los otros dos sitios. Las variantes de las reacciones se identificaron por secuenciación. Había dieciséis combinaciones posibles. De manera

similar al primer experimento, la condición 1 generaba más variantes con un número menor de sitios de mutación y las condiciones 2 y 3 generaban más variantes con un número mayor de sitios de mutación (Fig. 9 y 10). La distribución de todas las combinaciones posibles de los datos combinados (Total) era similar al patrón de distribución del cálculo estadístico (Fig. 9 y 10).

5

Tabla 2. Conjunto de cuatro mutaciones

Condiciones	Exploración	Clones únicos	Clones únicos (-1x)
2A	2x	75 % (12/16)	81 % (13/16)
8A	2x	62 % (10/16)	87 % (14/16)
14A	2x	68 % (11/16)	75 % (12/16)
2A + 8A	4x	93 % (15/16)	106 % (16/16)
2A + 14A	4x	93 % (15/16)	93 % (15/16)
2B	2x	75 % (12/16)	75 % (12/16)
8B	2x	50 % (8/16)	68 % (11/16)
14B	2x	75 % (12/16)	87 % (14/16)
2B + 8B	4x	81 % (13/16)	61 % (13/16)
2B + 14B	4x	87 % (14/16)	87 % (14/16)

Ejemplo 4

En el tercer experimento, tres sitios, se escogen tres sitios en un gen para combinarse (Fig. 11). Las reacciones se prepararon utilizando cebadores directos en dos sitios y un cebador inverso en el tercer sitio. Las variantes de las reacciones se identificaron mediante secuenciación. En este caso, había ocho combinaciones posibles diferentes. En la condición 9B, las 8 variantes se recuperaron mediante secuenciación en 24 clones. Véase la Fig. 12.

Ejemplo 5

Se seleccionaron 13 mutantes de la GSSMSM (5 sitios) para mejorar la estabilidad térmica y aumentar la actividad específica de la lipasa (Fig. 13). Se agruparon tres sitios (N168S, N171E, y M176W) y estaban contenidos en un cebador. El tamaño de la biblioteca era $6 \times 6 \times 2 \times 2 \times 2 = 288$. Las reacciones se prepararon mediante el siguiente procedimiento: los cebadores directos se agruparon en un grupo directo y los cebadores inversos se agruparon en un grupo inverso, y los cebadores del grupo directo y los cebadores del grupo inverso, independientemente entre ellos, se normalizaron para que tuvieran una concentración igual en el grupo correspondiente independientemente de las posiciones en el polinucleótido matriz, y en los que después de la normalización se añadió una cantidad igual de cebadores directos y cebadores inversos a la reacción. La combinación de las posiciones 1 y 2 tenía una tendencia (Fig. 13 y 14A, B, C). Un porcentaje menor de las variantes únicas posibles se consiguieron para las combinaciones de las posiciones 1 y 2 en comparación con las combinaciones de las posiciones 1 y 3 o las posiciones 2 y 3. Se llevaron a cabo dos rondas de reacción TMCA. En la Ronda II, se utilizaron algunas de las variantes obtenidas en la Ronda I. Después de la secuenciación se obtuvieron 720 clones (2,5x de cobertura de la biblioteca), el 46 % de las 288 variantes únicas en la Ronda I. Una cobertura de 1x de secuenciación significaría que el número de variantes (la progenie) secuenciadas igualaba el número de variantes únicas posibles, por lo tanto, una cobertura de 2,5x indica que el número de variantes (la progenie) secuenciadas igualaba 2,5 veces el número posible de variantes únicas (288). Se llevaron a cabo dos rondas de reacción TMCA. En la Ronda II, algunas de las variantes obtenidas en la Ronda I se utilizaron como polinucleótidos matriz. Los cebadores utilizados para cada reacción TMCA en la Ronda II se hacían a medida para obtener variantes no conseguidas en la Ronda I. Después de la Ronda II, se obtuvo el 95,5 % de las 288 variantes únicas. Se obtuvieron diez mutantes de esta biblioteca después de la exploración (Fig. 14A, B, C).

Los ejemplos demuestran que el procedimiento TMCA permite hacer bibliotecas de combinación eficazmente. Se definen algunas limitaciones por las posiciones mutantes, pero se pueden diseñar alternativas para superar estas limitaciones. La nueva modificación mejorada reduce significativamente la tendencia. Se pueden llevar a cabo múltiples rondas de reacción TMCA para superar alguna de las tendencias. El procedimiento TMCA ha demostrado ser eficaz para múltiples enzimas y sistemas de vectores. Una limitación del tamaño del vector puede ser de hasta 11 kb.

REIVINDICACIONES

1. Un procedimiento de producción de una pluralidad de polinucleótidos modificados que tienen diferentes combinaciones de mutaciones distintas en múltiples sitios, en el que las mutaciones son cambios de uno o más nucleótidos de la secuencia de una secuencia de ácido nucleico parental, comprendiendo dicho procedimiento:
 - 5 (a) la adición de al menos tres cebadores a un polinucleótido matriz de doble cadena en una única mezcla de reacción, en la que los al menos tres cebadores no se solapan, y en la que cada uno de los al menos tres cebadores comprende al menos una mutación diferente de la de los otros cebadores, en la que al menos un cebador es un cebador directo que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz, y
 - 10 (b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de los al menos tres cebadores.
2. El procedimiento de la reivindicación 1, que comprende adicionalmente la transformación de una célula con la pluralidad de productos extendidos que no se han tratado con una ligasa; y
 - 15 opcionalmente que comprende adicionalmente la recuperación de la pluralidad de polinucleótidos modificados extendidos de la célula; y
 - opcionalmente que comprende adicionalmente el análisis de la pluralidad de polinucleótidos modificados extendidos; y
 - opcionalmente en el que el análisis comprende la expresión de al menos uno de la pluralidad de polinucleótidos modificados extendidos y el análisis del polipéptido que se expresa a partir del mismo; y
 - 20 opcionalmente que comprende adicionalmente la selección de la pluralidad de polinucleótidos modificados extendidos que comprende las mutaciones de interés.
3. El procedimiento de la reivindicación 1, que comprende adicionalmente antes de la etapa (a) la obtención de la información de secuencia del polinucleótido matriz, y la identificación de tres o más mutaciones de interés a lo largo del polinucleótido matriz.
- 25 4. El procedimiento de la reivindicación 1, que comprende adicionalmente el análisis de la pluralidad de polinucleótidos modificados extendidos producidos por extensión con la polimerasa.
5. El procedimiento de la reivindicación 1, que comprende adicionalmente el tratamiento de la pluralidad de polinucleótidos modificados extendidos con una enzima, destruyendo de esta manera el polinucleótido matriz, la transformación de los polinucleótidos modificados extendidos tratados en una células, la recuperación de la
 - 30 pluralidad de polinucleótidos modificados extendidos de la célula, y la selección de la pluralidad de los polinucleótidos modificados extendidos que comprende las mutaciones de interés.
6. El procedimiento de la reivindicación 5, en el que la célula es una célula de *E. coli*.
7. El procedimiento de la reivindicación 5, en el que la enzima es una enzima de restricción; y opcionalmente en el que la enzima de restricción es la enzima de restricción *DpnI* y la célula es una célula de *E. coli*.
- 35 8. El procedimiento de la reivindicación 1, en el que se añaden al menos cuatro cebadores; o en el que se añaden al menos cinco cebadores; o en el que se añaden al menos seis cebadores; o en el que se añaden al menos ocho cebadores; o en el que se añaden al menos doce cebadores; o
- 40 en el que cada cebador comprende una única mutación puntual; o en el que al menos dos cebadores directos comprenden un cambio diferente en la misma posición del polinucleótido matriz; o en el que al menos dos cebadores inversos comprenden un cambio diferente en la misma posición del polinucleótido matriz; o
- 45 en el que al menos un cebador comprende al menos dos cambios en diferentes posiciones del polinucleótido matriz; o en el que al menos un cebador comprende al menos dos cambios en diferentes posiciones y al menos dos cebadores directos o dos cebadores inversos comprenden un cambio diferente en la misma posición del polinucleótido matriz; o
- 50 en el que la al menos una mutación se selecciona de entre el grupo que consiste en un cambio en uno o más secuencias de nucleótidos o aminoácidos codificados, una inserción, y una eliminación.
9. El procedimiento de la reivindicación 1, en el que el polinucleótido matriz es un ADN de doble cadena circular.
10. El procedimiento de la reivindicación 1, en el que al menos un cebador es un conjunto de cebadores degenerados que comprende cada uno una posición degenerada, en el que la mutación de interés es un intervalo de
 - 55 diferentes nucleótidos en la posición degenerada.

11. El procedimiento de la reivindicación 1, en el que al menos un cebador es un conjunto de cebadores degenerados que comprende al menos un codón degenerado que se corresponde con al menos un codón del polinucleótido matriz y al menos una secuencia adyacente que es homóloga de una secuencia adyacente del codón del polinucleótido matriz; y opcionalmente
 5 en el que el codón degenerado es N,N,N que codifica un aminoácido de origen natural; o en el que el codón degenerado puede codificar menos de los 20 aminoácidos de origen natural.
12. El procedimiento de la reivindicación 1, en el que los cebadores directos se agrupan en un grupo directo y los cebadores inversos se agrupan en un grupo inverso, y los cebadores del grupo directo y los cebadores del grupo inverso, se normalizan, independientemente entre ellos, para tener una concentración igual en el grupo correspondiente independientemente de las posiciones en el polinucleótido matriz, y en el que después de la normalización se añade a la reacción una cantidad igual de los cebadores directos e inversos.
 10
13. El procedimiento de la reivindicación 1, que comprende adicionalmente antes de la etapa (b).
 la organización de los cebadores en múltiples grupos dependiendo de su localización en el polinucleótido matriz; en el que los cebadores que cubren la misma región seleccionada en la matriz están en un grupo,
 15 la normalización de los cebadores agrupados en cada grupo para que tengan una concentración igual, el agrupamiento de los cebadores directos de un grupo en un grupo directo y normalizando la concentración entre cada grupo de los cebadores directos para que sea igual, el agrupamiento de los cebadores inversos de un grupo en un grupo inverso y normalizando la concentración entre cada grupo de los cebadores inversos para que sea igual, y
 20 la adición de una cantidad igual de los cebadores directo e inverso en la reacción.
14. El procedimiento de la reivindicación 1, que comprende adicionalmente llevar acabo dos rondas de las tapas (a) a (b), y utilizar los polinucleótidos producidos en la primera ronda como polinucleótido matriz en la segunda ronda.
15. Un procedimiento de producción de una pluralidad de polinucleótidos modificados que tienen diferentes combinaciones de mutaciones distintas en múltiples sitios y que comprenden las mutaciones de interés, en el que las mutaciones son cambios de uno o más nucleótidos de la secuencia de una secuencia de ácido nucleico parental, comprendiendo dicho procedimiento:
 25
- (a) la adición de al menos dos cebadores a un polinucleótido matriz de doble cadena en una única mezcla de reacción, en la que los al menos dos cebadores no se solapan, y en la que cada uno de los al menos dos cebadores comprende al menos una mutación diferente de la del otro cebador o cebadores, en la que al menos un cebador es un cebador directo que se puede hibridar con una cadena menos de la matriz y al menos un cebador es un cebador inverso que se puede hibridar con una cadena más de la matriz,
 - (b) someter la mezcla de reacción a una reacción de extensión con una polimerasa para dar lugar a una pluralidad de polinucleótidos modificados extendidos a partir de los al menos dos cebadores,
 - (c) el tratamiento de la pluralidad de polinucleótidos modificados extendidos con una enzima, destruyendo de esta manera el polinucleótido matriz,
 - (d) la transformación de los polinucleótidos modificados extendidos tratados que no se han tratado con una ligasa en una célula,
 - (e) la recuperación de una pluralidad de polinucleótidos modificados extendidos de la célula, y
 - (f) la selección de la pluralidad de polinucleótidos modificados extendidos que comprenden la mutación de interés.
 30
 35
 40

Evolución – GSSM™ (1)

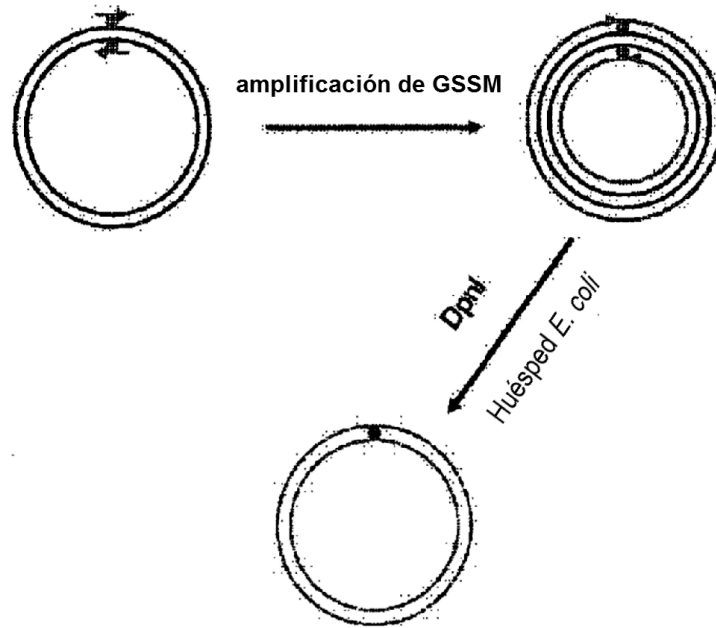


Figura 1.

Evolución – GSSM™ (2)

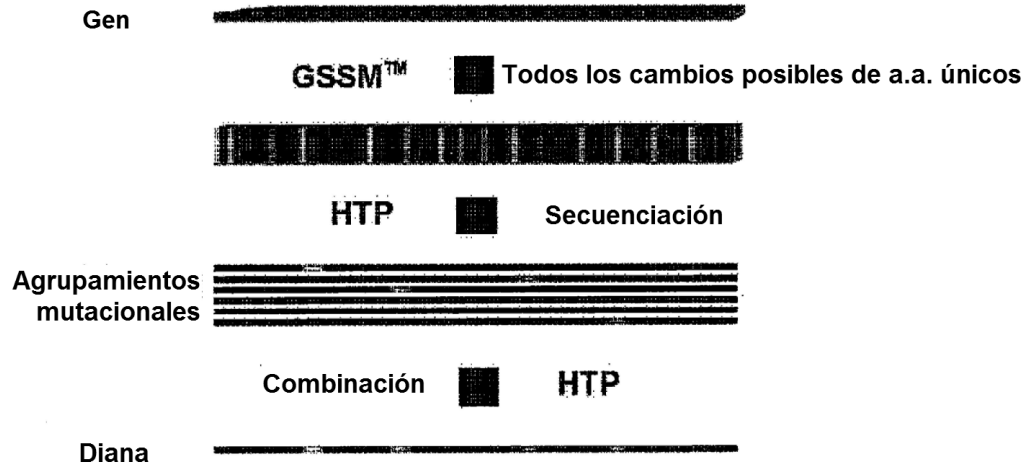


Figura 2.

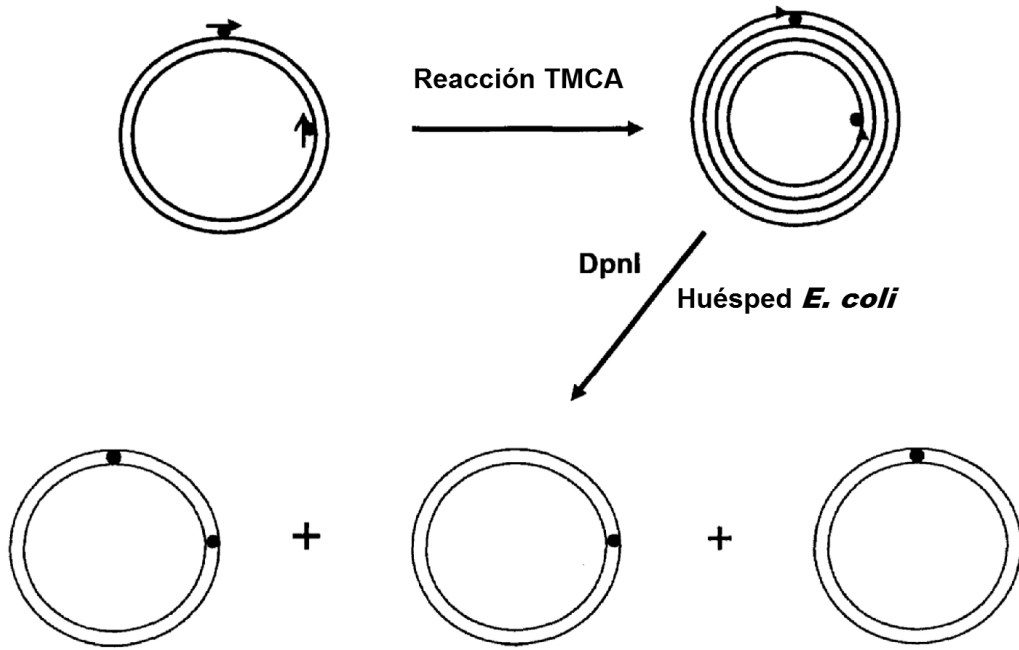


Figura 3.

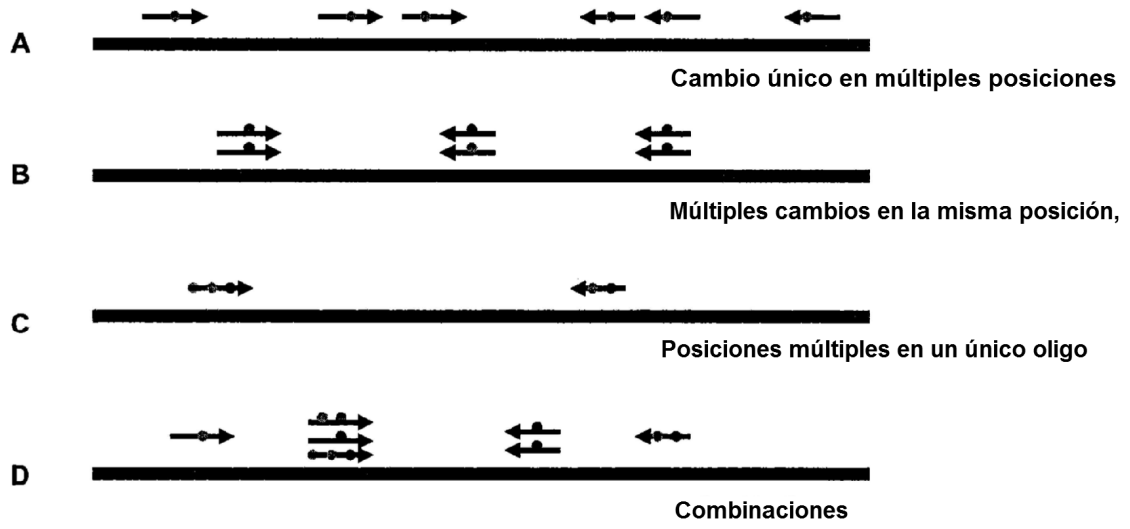


Figura 4.

Conjunto de seis mutaciones

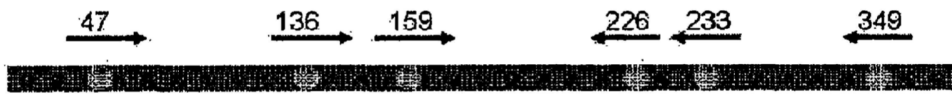


Figura 5.

Conjunto de seis mutaciones

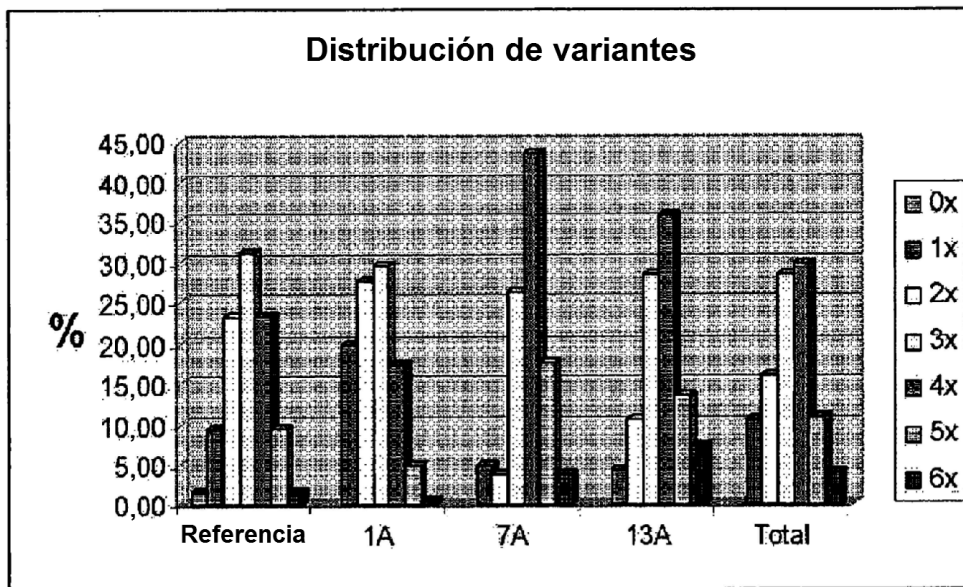


Figura 6.

Conjunto de seis mutaciones

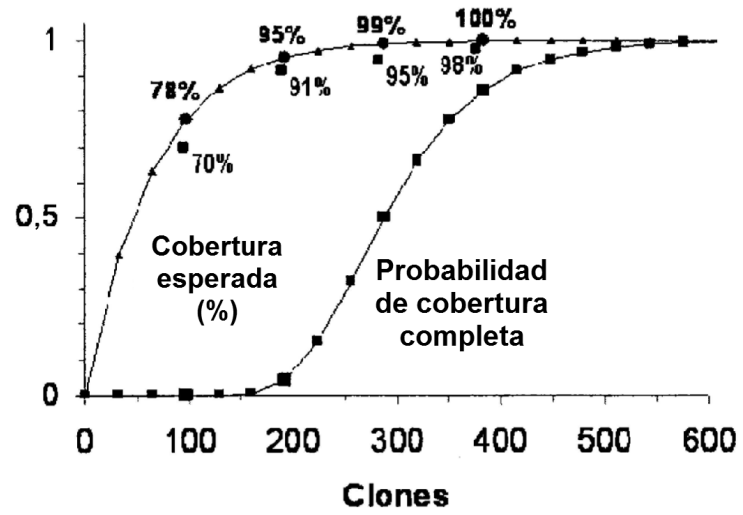


Figura 7.

Conjunto de Cuatro Mutaciones



Figura 8.

Conjunto de Cuatro Mutaciones

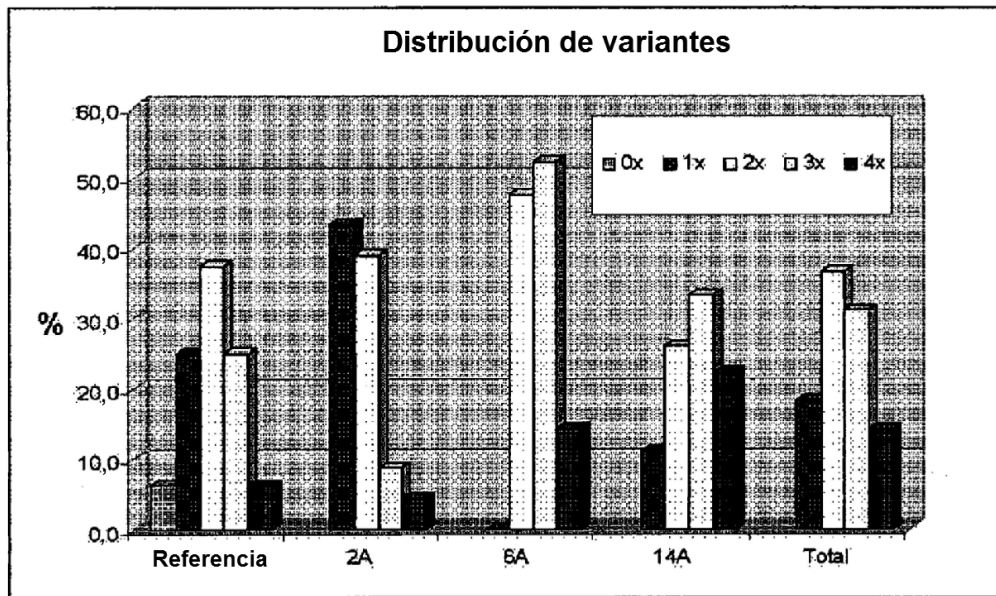


Figura 9.

Conjunto de Cuatro Mutaciones

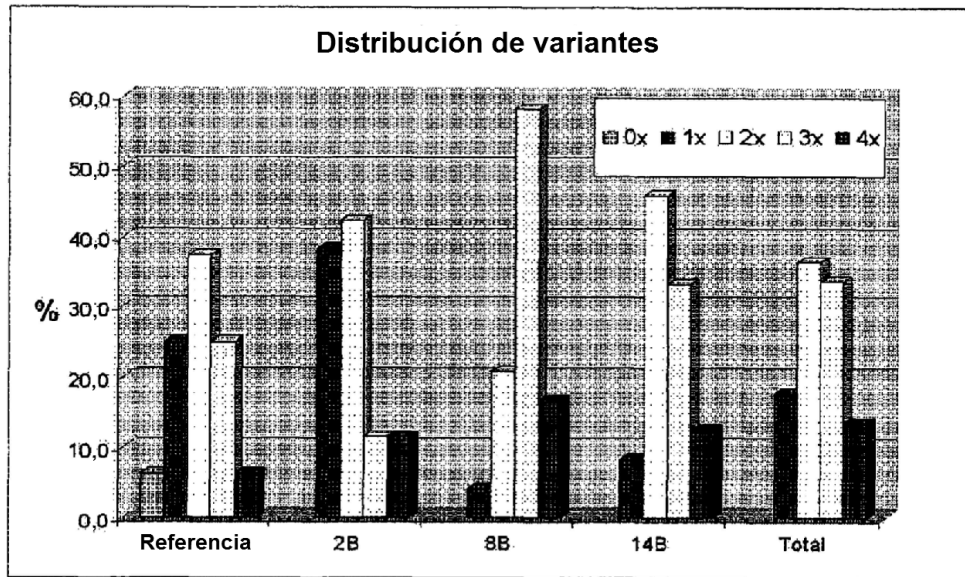


Figura 10.

Conjunto de Tres Mutaciones



Figura 11.

Conjunto de Tres Mutaciones

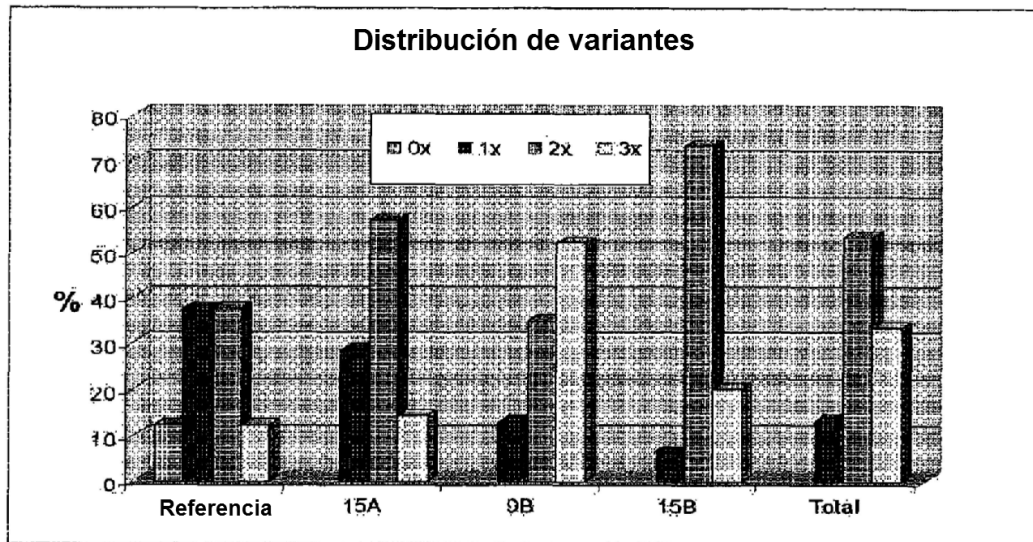
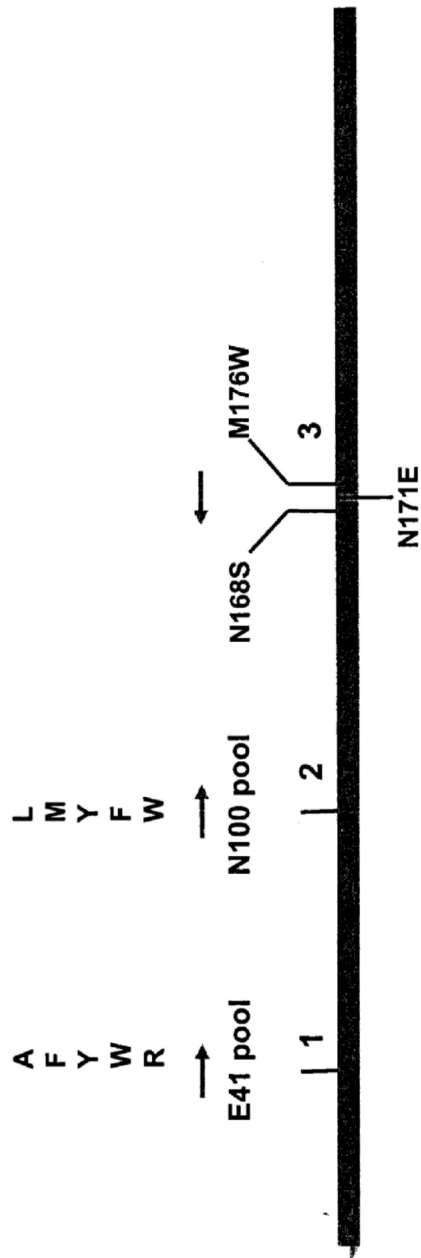


Figura 12.



Tamaño de la biblioteca $6 \times 6 \times 2 \times 2 \times 2 = 288$

Figura 13.

Desglose de las Variantes – ronda I de TMCA

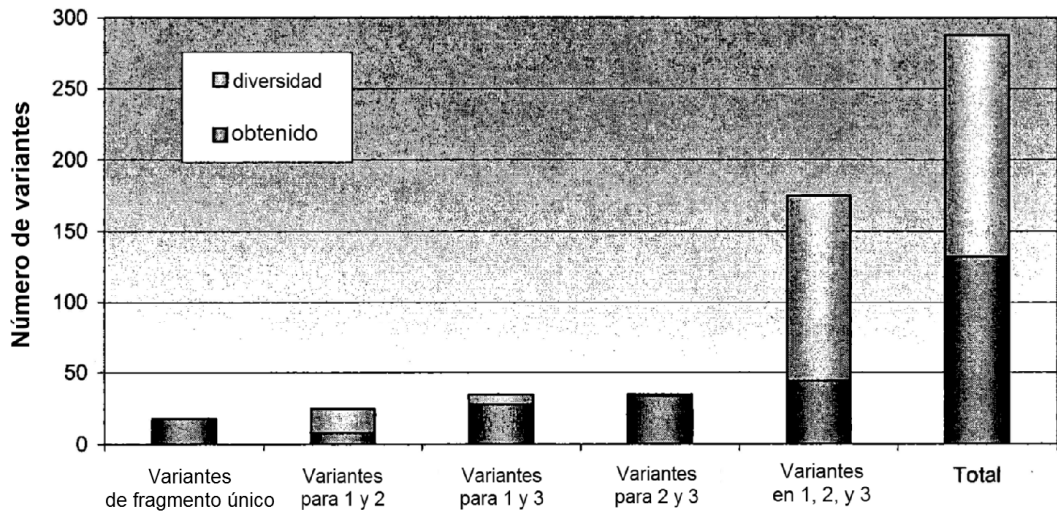


Figura 14A.

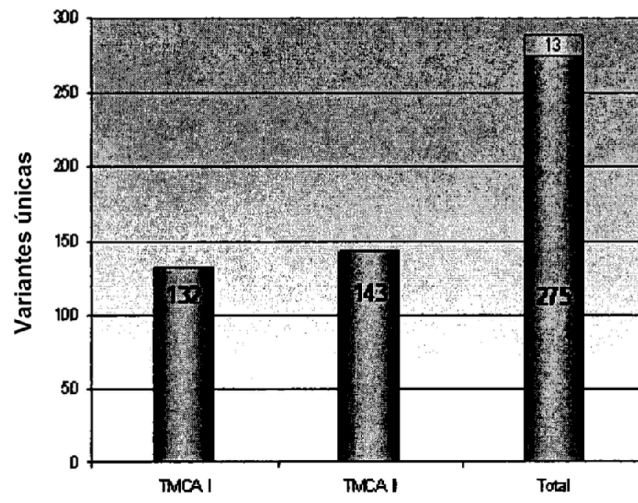


Figura 14B.

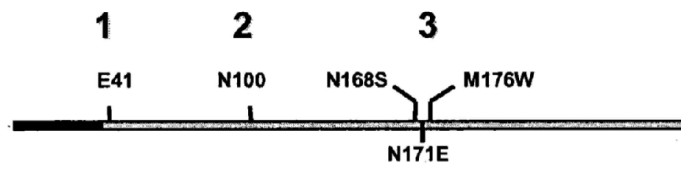


Figura 14C.