

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 582 385**

51 Int. Cl.:

**C12N 15/10** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.12.2007 E 07872389 (7)**

97 Fecha y número de publicación de la concesión europea: **13.04.2016 EP 2097524**

54 Título: **Método de diversificación aleatoria de una secuencia genética que permite preservar la identidad de determinados segmentos internos de dicha secuencia genética**

30 Prioridad:

**20.12.2006 FR 0655723**

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

**12.09.2016**

73 Titular/es:

**PHERECYDES PHARMA (100.0%)  
102 avenue Gaston Roussel  
93230 Romainville, FR**

72 Inventor/es:

**IRIS, FRANÇOIS**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

**ES 2 582 385 T3**

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

## DESCRIPCIÓN

Método de diversificación aleatoria de una secuencia genética que permite preservar la identidad de determinados segmentos internos de dicha secuencia genética

5 La presente invención se dirige a un método muy general para diversificar aleatoriamente una secuencia nucleotídica S por PCR, preservando al mismo tiempo la identidad de determinados dominios o segmentos de esta secuencia, y describe un banco de secuencias nucleotídicas así diversificadas, y más particularmente un banco de secuencias codificantes de las proteínas de orientación selectiva de bacteriófagos modificados.

10 Los bacteriófagos se presentan como virus capaces de infectar específicamente las bacterias y de replicarse en ellas. Su existencia se puso de manifiesto a principios del siglo XX por el británico Frederick Twort y el quebequense Félix d'Hérelle.

15 Los bacteriófagos ocupan todos los nichos ecológicos en los que se encuentran las bacterias. Se manifiestan principalmente en dos formas: forma lisogénica, por la cual pueden permanecer de manera quiescente en el interior de su huésped, o como forma lítica, cuando se replican de manera activa al lisis la célula bacteriana. La forma lítica provoca que los bacteriófagos se liberen en grandes cantidades en el medio en forma infecciosa.

20 A fin de conservar su carácter infeccioso con respecto a sus huéspedes que experimentan a veces mutaciones rápidas, los bacteriófagos evolucionan constantemente. Por ello, presentan naturalmente un grado elevado de especialización para las especies bacterianas que parasitan y son muy diversificados.

25 A partir de su descubrimiento, los bacteriófagos se consideraron como un medio para luchar contra las infecciones bacterianas, mucho antes de la era de los antibióticos.

30 El proceso que consiste en identificar en la naturaleza bacteriófagos específicos de una bacteria patógena a fin de tratar los pacientes infectados por esta bacteria se desarrolló así en Rusia y en los países del antiguo bloque soviético durante la primera mitad del siglo XX.

No obstante, el empleo de antibióticos, cuyo espectro es generalmente mucho más amplio, se generaliza en la segunda mitad del siglo XX de forma masiva, sin explotar todas las posibilidades ofrecidas por los bacteriófagos.

35 Hoy en día, con la aparición de cepas bacterianas multirresistentes a los antibióticos, y habida cuenta de las dificultades halladas por la comunidad científica en desarrollar nuevos antibióticos, los bacteriófagos contribuyen a reavivar un nuevo interés en el tratamiento de infecciones bacterianas difíciles de erradicar, especialmente en casos de infecciones nosocomiales [Thiel, K., *Nature Biotechnology*, 2004, 22:31-36].

40 No obstante, persisten algunas dificultades en la utilización de bacteriófagos, en particular en el hecho en el que las bacterias pueden evadir los bacteriófagos enmascarando o alterando los elementos constitutivos de su pared externa.

45 El ciclo de replicación de los bacteriófagos requiere, en efecto, una etapa de identificación y adhesión del bacteriófago a la pared de la bacteria huésped, que afecta a la capacidad del bacteriófago para infectar la bacteria, es decir, inyectar el material genético contenido en su cápside en el interior del citoplasma de la bacteria.

50 El bacteriófago T4, por ejemplo, es un bacteriófago que infecta las bacterias de tipo *Escherichia coli* cuyo ciclo de replicación dura aproximadamente 30 minutos a 37 °C. Este ciclo de replicación comienza inmediatamente tras la identificación de la bacteria huésped por el bacteriófago, mediante la fase de absorción y penetración. Esto se traduce en el cese inmediato de la expresión génica de la bacteria huésped, la síntesis de las enzimas necesarias en la replicación del fago, 5 minutos tras la infección, la replicación del ADN (iniciada tras 10 minutos) y la formación del virus (iniciada tras 12 minutos). El ciclo de replicación produce la ruptura de la bacteria (tras 30 minutos) y la liberación en el medio de aproximadamente cincuenta bacteriófagos por bacteria lisada.

55 La adhesión a la bacteria se asegura principalmente por las proteínas de la plataforma basal (placa basal) que sirve de anclaje al bacteriófago y la identificación se asegura más particularmente por las proteínas que forman los filamentos periféricos, denominadas "fibras de la cola". Sin embargo, las proteínas de las fibras de la cola y de la placa basal pueden implicarse a la vez en la identificación y adhesión del bacteriófago a la pared bacteriana. El conjunto de estas proteínas llamadas "orientación selectiva" se representa en las figuras 1 y 3 de la presente solicitud.

60 Entre las proteínas implicadas en esta identificación o esta adhesión en el fago T4, pueden citarse más particularmente las glicoproteínas GP12 de la plataforma basal, y las glicoproteínas GP36, GP37 y GP38 de las fibras de la cola.

65 A fin de limitar la emergencia de bacterias resistentes al dispositivo de identificación de bacteriófagos, se propone en

general utilizar simultáneamente diferentes formas de bacteriófagos capaces de orientar selectivamente una misma bacteria.

5 Estos bacteriófagos se extraen a partir de la naturaleza o procedentes de colecciones para formar el conjunto denominado "cóctel bacteriófago".

10 No obstante, conviene para el desarrollo de estos cócteles de bacteriófagos seleccionar individualmente y de forma rigurosa los bacteriófagos que los componen, asegurando, en particular, que estos bacteriófagos son líticos y no lisogénicos o parcialmente lisogénicos, como ocurre con frecuencia en el caso de bacteriófagos extraídos de la naturaleza.

15 La necesidad de ensayar individualmente los bacteriófagos para asegurar su eficacia real hace que el desarrollo de los cócteles de bacteriófagos sea largo y tedioso, y por tanto es preciso prever un cóctel diferente para cada bacteria considerada.

20 La solicitud WO 01/51066 describe dicho preparado de bacteriófagos que comprende seis bacteriófagos diferentes utilizado como conservante de productos alimenticios frescos para destruir la bacteria *Listeria monocytogenes*, responsable de la listeriosis. Este preparado natural se envasa en un pulverizador para pulverizarse en la carne o en productos lácteos. Es seguro para los seres humanos, animales o plantas, ya que los bacteriófagos pueden infectar las bacterias del género *Listeria* y no las células de los organismos pluricelulares.

25 Para afrontar los problemas planteados por la selección de bacteriófagos naturales, en la solicitud WO 06/066224 se propone un método para obtener bacteriófagos cuyas proteínas de orientación selectiva se modifican para orientar específicamente un factor de virulencia dado. Los factores de virulencia son moléculas descritas necesarias para que la bacteria desarrolle una infección. Estas moléculas se consideran elementos estables, menos susceptibles de variar durante la infección que los elementos estructurales externos como, por ejemplo, los lipopolisacáridos. Más particularmente, este método propone seleccionar una proteína procedente de un bacteriófago natural (por ejemplo GP37 del fago T4) capaz de reconocer un factor de virulencia descrito en la bibliografía (por ejemplo OmpC de *E. coli*), de transferir el gen que codifica esta proteína en un bacteriófago lambda y de utilizar el bacteriófago lambda para modificar esta proteína. La modificación de la proteína consiste en llevar a cabo intercambios entre los diferentes dominios involucrados en la identificación del factor de virulencia (por ejemplo, los dominios His de GP37). Los fagos lambda se ensayaron a continuación por su capacidad de adhesión al factor de virulencia diana. Este método, que es similar a la técnica de expresión en fago, permite aislar diferentes variantes del bacteriófago lambda capaces de orientar selectivamente el factor de virulencia, y de este modo disponer de diferentes proteínas de orientación selectiva. Los genes correspondientes a estas proteínas diferentes de orientación selectiva, pueden, posteriormente, transferirse en los bacteriófagos infecciosos. Estos bacteriófagos pueden entonces constituir cócteles de fagos activos con respecto a la bacteria portadora del factor de virulencia diana originalmente.

40 Este método constituye un avance en la obtención de bacteriófagos diversificados para el desarrollo de cócteles de fagos. Sin embargo, lo cierto es que este tipo de cócteles se refiere inicialmente a las especies bacterianas que expresan el factor de virulencia diana.

45 Este método puede aplicarse por tanto a bacterias patógenas cultivables en las que pudo efectuarse una identificación previa de los factores de virulencia.

50 Para conseguir que el empleo de bacteriófagos sea más universal, resultaría útil disponer de bacteriófagos infecciosos con respecto a un mayor número de especies, por ejemplo, creando bacteriófagos cuyo espectro de infectividad puede modificarse o ampliarse. Dichos bacteriófagos podrían utilizarse en detrimento de nuevas especies bacterianas, en particular contra bacterias patógenas emergentes o al inicio de las infecciones nosocomiales.

55 Sin embargo, la producción de fagos modificados en su espectro de infectividad se enfrenta a las limitaciones técnicas cuya obtención de bacteriófagos depende del huésped bacteriano en el cual el bacteriófago se transforma y multiplica. Las etapas de modificación genética de los fagos comprenden generalmente varias etapas de replicación en un único huésped bacteriano.

60 En los experimentos descritos en los documentos de la técnica anterior susodichos, se requieren numerosos ciclos de replicación para modificar los bacteriófagos utilizados para orientar selectivamente los factores de virulencia. La recombinación homóloga que es la técnica observada con más frecuencia para transformar el genoma de los bacteriófagos implica numerosos ciclos de replicación y selección sucesivos en la bacteria. Ahora bien, los bacteriófagos modificados, si consiguen adquirir la capacidad de infectar huéspedes diferentes de sus huéspedes habituales, también pueden perder la capacidad de infectar al huésped utilizado para su replicación. Por consiguiente, se eliminan de la selección y evaden al experimentador. Esto da lugar a una pérdida significativa de la diversidad de fagos modificados que pueden obtenerse.

65 Para superar las dificultades antes mencionadas, la presente invención propone un nuevo método que consiste en

diversificar aleatoriamente secuencias nucleotídicas, particularmente las que codifican las proteínas de orientación selectiva de los bacteriófagos, por inserción en sus genes de secuencias de ADN producidas aleatoriamente.

5 Este método es particularmente útil para producir copias de genes que comprenden segmentos mutados aleatoriamente, en particular para su clonación en vectores de expresión o de recombinación homóloga.

10 El presente método según la invención se desarrolló originalmente para diversificar las proteínas de orientación selectiva de los bacteriófagos y obtener así por recombinación homóloga bacteriófagos recombinantes cuya especificidad de huésped varía.

No obstante, no se limita al campo de bacteriófagos, puesto que puede aplicarse a cualquier secuencia de ADN o ARN.

15 Los dominios de aplicación son numerosos especialmente en el ámbito de la medicina, dado que se pretende modificar una secuencia genética aleatoriamente en sus dominios variables, preservando al mismo tiempo la identidad de los dominios constantes, que generalmente son esenciales para la funcionalidad de la proteína.

20 El método según la invención se detalla en lo sucesivo, así como su aplicación más específica en el campo de los bacteriófagos.

**Figura 1:** representación del bacteriófago T4 que muestra los diferentes elementos constitutivos del bacteriófago. Las proteínas consideradas en el contexto de la presente invención se encuadran.

**Figura 2:** representación del genoma completo del bacteriófago T4. Los genes citados en la presente solicitud se indican con una flecha perpendicular a los marcos abiertos de lectura.

25 **Figura 3:** representación tridimensional de la plataforma basal del bacteriófago y las fibras de la cola implicadas en la identificación y la adhesión del bacteriófago a la bacteria huésped, permitiendo visualizar las proteínas de orientación selectiva modificadas según el método de la invención.

30 **Figura 4:** diagrama que resume el principio de obtención de un banco de bacteriófagos recombinantes según la invención. El marco superior izquierdo representa una bacteria huésped que comprende 3 vectores de recombinación homólogos para introducir oligonucleótidos cuya secuencia se produce aleatoriamente en tres genes codificantes de las proteínas de orientación selectiva del bacteriófago. Estos vectores representan construcciones de ADN en la presente invención, portadoras de una gran diversidad genética. Tras la infección por un bacteriófago y la recombinación homóloga, se obtiene una gran cantidad de bacteriófagos (banco de bacteriófagos) con proteínas de orientación selectiva, que forman "una fuente de diversidad de orientación selectiva". Los bacteriófagos obtenidos se exploran con respecto a nuevos huéspedes potenciales (selección positiva) a fin de seleccionar bacteriófagos capaces de infectar estos huéspedes. Pueden igualmente ensayarse en los huéspedes no bacterianos (células eucariotas) para asegurar que no son perjudiciales para los seres humanos o animales.

35 **Figura 5:** comparación de la secuencia polipeptídica de GP12 del bacteriófago T4 (línea superior) y su homólogo presente en el bacteriófago RB 69 (línea inferior) según el protocolo BLAST. Los aminoácidos comunes a las dos proteínas se indican en la línea intermedia. El símbolo "+" significa que los aminoácidos son similares. La parte N-terminal encuadrada corresponde al dominio de estas dos proteínas que permite el anclaje del bacteriófago en la pared de la bacteria. Este es el dominio de anclaje que se muta aleatoriamente y se inserta por recombinación homóloga en el genoma del fago según la invención. Los segmentos internos D1 a D4 corresponden a las secuencias de la proteína que se conservan durante el método de diversificación aleatoria implementado por PCR según la invención.

40 **Figura 6:** figura que recapitula las etapas del método por PCR implementado según la invención para obtener, en particular, una copia del dominio de anclaje de gp12 en la que se insertan oligonucleótidos producidos aleatoriamente. Los segmentos internos D1 a D4 deseados para conservar la identidad de la secuencia se representan por rectángulos. Estos dominios corresponden a los mencionados en la figura 5 precedente. Los dos segmentos externos delimitan la secuencia del dominio de anclaje de gp12. **A:** el dominio de anclaje de gp12 se amplifica por PCR de alta fidelidad. **B:** se realizan 4 PCR propensas a error de forma independiente utilizando los oligonucleótidos mencionados. **C:** Los productos de las amplificaciones obtenidos en B se purifican y se juntan en una misma reacción de PCR de alta fidelidad. Dichos productos de amplificación se superponen, es posible ensamblar los diferentes fragmentos, pero a condición de que los dominios D1 a D4 se conserven lo suficiente para poder hibridar los cebadores utilizados. **D:** de la PCR realizada en la etapa C, dos fragmentos PA-1 y PB-1 que se ensamblan utilizando los cebadores corresponden a los segmentos externos del dominio de anclaje de GP12. Al final, se obtiene una copia del dominio de anclaje de GP12 cuya secuencia se modificó aleatoriamente excepto en los dominios D1 a D4 cuya identidad se ha conservado. Los cebadores utilizados en el ejemplo de realización se describieron en cada una de las etapas mencionadas en esta figura.

50 **Figura 7:** perfil de secuenciación obtenido directamente en los productos de PCR (gp12-Mut) obtenidos según el método de la invención. Los perfiles se establecieron utilizando el programa de análisis libre ApE a partir de los datos procedentes de un secuenciador automático. **A:** secuenciación de la porción de gp12 situado entre los nucleótidos 1089 y 1110 (SEQ ID NO.1). La parte izquierda conservada corresponde a la parte 3' del dominio D1 (nucleótidos 1089 a 1099). La parte derecha muestra una superposición de picos que reflejan las mutaciones aleatorias producidas durante el método de PCR en la región situada inmediatamente corriente abajo de D1

(nucleótidos 1100 a 1110). **B:** comparación de la secuenciación realizada en una región no mutada de gp12 situada entre los nucleótidos 1195 y 1212 de SEQ ID NO. 1 (*arriba*) y realizada para la misma región en los productos de PCR obtenidos según el método (*abajo*). La región secuenciada se encuentra entre los dominios conservados D1 y D2. Se aprecia la presencia de picos superpuestos (*abajo*) que reflejan la inserción aleatoria de nucleótidos en la secuencia de gp12 inicial.

Descripción del método de diversificación de secuencias genéticas según la invención:

Por consiguiente, la presente invención tiene por objeto más particularmente un método de PCR que permite mutar aleatoriamente una secuencia nucleotídica S, delimitada en sus extremos 5' y 3' por dos segmentos F1 y F2, conservando al mismo tiempo la identidad de al menos un segmento interno D de dicha secuencia nucleotídica, caracterizado por que comprende las siguientes etapas:

i) se efectúa una PCR propensa a error en toda la secuencia S utilizando al menos dos cebadores, uno de los cuales es sentido y el otro antisentido, respectivamente de los segmentos F1 y F2, de modo que se amplifica la secuencia S al introducir mutaciones aleatoriamente;  
 ii) se purifican los productos de amplificación obtenidos, que corresponden a las copias de S mutadas aleatoriamente;  
 iii) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación purificados en la etapa ii), utilizando pares de cebadores sentido y antisentido, respectivamente, al menos:

- un cebador sentido de F1 y un cebador antisentido se hibridan en su totalidad con el segmento D, a fin de amplificar la región F1-D de las copias de S mutadas,
- un cebador sentido se hibrida en su totalidad con el segmento D y un cebador antisentido de F2, a fin de amplificar la región D-F2 de las copias de S mutadas;

iv) se purifican los productos de amplificación obtenidos en la etapa iii), que consisten en copias de los segmentos F1-D y D-F2 de la secuencia S mutada en la etapa i) en las que el segmento D ha conservado su identidad;

v) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación purificados en la etapa iv) utilizando los cebadores sentido y antisentido, respectivamente F1 y F2;  
 vi) se purifican los productos de PCR obtenidos, que corresponden a las secuencias nucleotídicas S mutadas en la etapa i) en las que el segmento D ha conservado su identidad.

Un cebador según la invención es un ácido nucleico monocatenario, apto para hibridarse con una porción, o la totalidad, de una de las cadenas de ADN que forman la totalidad o parte de la secuencia S. Se dice cebador de "sentido" cuando su concatenación de nucleótidos reproduce, en algunos nucleótidos próximos, una parte de la cadena codificante de S. Se dice cebador "antisentido" de S, cuando la concatenación de nucleótidos reproduce, en algunos nucleótidos próximos, una parte de la cadena no codificante de S.

Cuando un cebador se proporciona como "correspondiente a" un segmento dado de la secuencia S, esto significa que puede ser sentido o antisentido en relación con una parte de S tomada en su forma de molécula de ADN bicatenario.

Este método resulta particularmente ventajoso cuando se pretende conservar la identidad de numerosos dominios internos  $D_N$  de la secuencia nucleotídica S (como es el caso de la proteína GP12), siendo considerada N como un número entero igual o superior a 1.

Para  $N=2$ , el método comprende las siguientes etapas:

i) se efectúa una PCR propensa a error en toda la secuencia S utilizando al menos dos cebadores, uno de los cuales es sentido y el otro antisentido, respectivamente de los segmentos F1 y F2, de modo que amplifica la secuencia S, introduciendo mutaciones aleatoriamente;  
 ii) se purifican los productos de amplificación obtenidos;  
 iii) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación purificados en la etapa ii), utilizando pares de cebadores sentido y antisentido, al menos:

- un cebador sentido de F1 y un cebador antisentido se hibridan en su totalidad con  $D_2$ , a fin de amplificar la región F1- $D_2$  de las copias de S mutadas, en las cuales el segmento  $D_2$  ha conservado su identidad;
- un cebador antisentido de F2 y un cebador sentido se hibridan en su totalidad con  $D_1$ , a fin de amplificar la región  $D_1$ -F2 de S mutada en las que el segmento  $D_1$  ha conservado su identidad;

iv) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación obtenidos en la etapa iii) utilizando al menos un par de cebadores sentido y antisentido de F1 y F2;

v) se purifican los productos de PCR obtenidos en la etapa iv), que corresponden a las secuencias nucleotídicas S mutadas aleatoriamente, en las que las secuencias de los segmentos  $D_2$  y  $D_1$  han conservado su identidad.

- Según un aspecto preferente de la invención, este método es útil para introducir oligonucleótidos producidos aleatoriamente en ciertos dominios variables de las secuencias codificantes de las proteínas de orientación selectiva de los bacteriófagos. Las proteínas de orientación selectiva de los bacteriófagos se definen como proteínas implicadas en la identificación y adhesión del bacteriófago a la bacteria huésped. Estas proteínas se seleccionan preferentemente entre las que constituyen las fibras de la cola o de la plataforma basal del fago T4. Una proteína particularmente adaptada al método según la invención es la proteína GP12 de la plataforma basal cuya secuencia nucleotídica corresponde a SEQ ID NO: 1. Otras proteínas de orientación selectiva preferentes son GP36, GP37 o GP38 constitutivas de la parte distal de la fibras de la cola. Evidentemente, también resultan preferentes las proteínas homólogas de las citadas previamente presentes en otros bacteriófagos.
- Por secuencia homóloga, se entiende secuencias con al menos 50 % de identidad con estas últimas, preferentemente al menos 70 %, más preferentemente al menos 90 % presente en dos organismos de especies diferentes.
- El ejemplo se proporciona en la presente solicitud de la modificación de la proteína GP12 del bacteriófago T4, particularmente preferente según la invención. La proteína GP12 presenta un dominio de anclaje en el que se desea introducir la variabilidad mientras se mantiene la integridad de ciertos segmentos señalados D1 a D4 (figura 5), en los que los segmentos corresponden a dominios conservados en las proteínas homólogas de los bacteriófagos de tipo T.
- Preferentemente, estos son los segmentos génicos de las proteínas correspondientes a los dominios de la proteína implicados en las funciones de adhesión o identificación que se modifican, y más preferentemente los segmentos correspondientes a los dominios variables.
- Por dominio variable, se entiende los segmentos de una secuencia que se conservan menos cuando se comparan dos secuencias homólogas entre sí.
- A cambio, resulta preferente conservar la identidad de los dominios más conservados, que generalmente resultan indispensables para la funcionalidad de las proteínas codificadas, en su caso, por estas secuencias.
- Los protocolos de PCR utilizados según la invención son los protocolos convencionales conocidos adecuadamente por el experto en la materia. La forma en que se implementan distingue la invención.
- Por PCR propensa a error, se entiende una reacción de polimerización en cadena efectuada en condiciones que no permiten una replicación fiel de las secuencias de ADN. Este tipo de reacción puede obtenerse mediante la implementación de una polimerasa clásica de tipo *Taq* en condiciones pocas rigurosas y en presencia de sales de manganeso, como se describe en la bibliografía [Cadwell, R. C. *et al.* 1992, *Randomization of genes by PCR mutagenesis, PCR Methods Appl.*, 2:28-33].
- Una PCR de alta fidelidad es, por el contrario, una reacción de polimerización en cadena que permite amplificar una plantilla de ADN con una tasa de error de replicación muy bajo. Este tipo de reacción puede obtenerse mediante la implementación, por ejemplo, de una polimerasa de tipo *Pfu* en condiciones rigurosas, como se describe en la bibliografía [Innis, M. A. *et al.* Eds., *PCR Protocols: a guide to methods and applications*, 1989, Academic Press].
- Las reacciones de PCR de alta fidelidad mencionadas en la etapa iii), que permiten amplificar las subregiones de la secuencia S situadas entre los segmentos conservados, se efectúan preferentemente por separado.
- En los ejemplos de la presente solicitud, la letra F indica un cebador "sentido" (en inglés *forward*) y la letra R indica un cebador antisentido (en inglés *reverse*).
- El presente método puede aplicarse a cualquier secuencia nucleotídica que pretende variar selectiva y aleatoriamente algunos segmentos.
- El método según la invención puede encontrar aplicaciones en numerosos ámbitos terapéuticos.
- Su implementación, de hecho, es particularmente útil cuando se pretende modificar el espectro de actividad, o de interacción con un ligando, de una proteína al modificar aleatoriamente dominios particulares de su secuencia genética.
- Por consiguiente, la invención prevé, por ejemplo, que la secuencia S sea una secuencia génica codificante de una proteína de interés.
- Según un aspecto preferente de la invención, la proteína de interés es una proteína de orientación selectiva de un bacteriófago, tal como la proteína GP12 de un bacteriófago de tipo T.
- No obstante, puede tratarse igualmente de cualquier proteína de interés terapéutico o diagnóstico que se pretende

que varíe los sitios activos o dominios de interacción, como por ejemplo una proteína ligando que se pretende que modifique la especificidad con respecto a un receptor o incluso una inmunoglobulina que se pretende que modifique los dominios variables o hipervariables con el fin de detectar anticuerpos.

5 La presente solicitud no puede dar lugar a una limitación del punto de vista de las numerosas aplicaciones que dejan entrever la modificación de una secuencia genética según la invención.

10 Cuando la secuencia S está presente en forma de ARN, el presente método puede estar precedido por una etapa de retrotranscripción que permite obtener una copia de dicha secuencia en forma de polinucleótido según los protocolos habituales.

Los productos de amplificación obtenidos según el método particular de la invención constituyen un banco de secuencias nucleotídicas S diversificadas aleatoriamente cuya identidad de dominios internos  $D_N$  se ha conservado.

15 La invención describe por consiguiente productos de PCR obtenidos según la invención, caracterizados por que consisten en variantes polinucleotídicas resultantes de la mutación aleatoria de la secuencia S y caracterizados por que uno o varios dominios internos D de dicha secuencia S se han conservado.

20 En otras palabras, la invención permite obtener un banco de polinucleótidos constituido por variantes de una secuencia nucleotídica S mutada por PCR aleatoriamente, caracterizado por que dichas variantes comprenden uno o más segmento(s) interno(s) D de dicha secuencia S intacta o cuya identidad se ha conservado.

25 En general, los dominios internos  $D_N$  conservan, según el método de la invención, más del 50 % de identidad en relación con su secuencia inicial en la proteína de tipo natural, preferentemente más del 70 %, y más preferentemente más del 90 %, incluso más del 99 %, según las condiciones de PCR utilizadas, en particular las condiciones de PCR de alta fidelidad efectuadas en las etapas iii) y v) del método según la invención.

30 Las variantes polinucleotídicas que constituyen el banco de secuencias nucleotídicas pueden clonarse directamente en vectores de expresión, o más preferentemente en vectores de recombinación homóloga, de los cuales el experto tiene a su disposición, y formar un conjunto de construcciones según la invención.

35 El método según la invención, permite en particular, superar la dificultad de mutar por partes un gen aleatoriamente. De hecho, en la técnica anterior se requería producir oligonucleótidos aleatoriamente para ensamblar después estos oligonucleótidos mediante clonación a las partes de la secuencia génica cuya identidad se pretende conservar. Ahora bien, cuando varios oligonucleótidos deben insertarse en diferentes sitios del gen, este trabajo es muy tedioso y el resultado en términos de diversidad de las secuencias obtenidas es decepcionante.

40 Las construcciones de ADN obtenidas según la invención, en su conjunto, permiten por consecuencia más particularmente, integrar en el genoma un número considerable de diferentes secuencias nucleotídicas codificantes de las proteínas modificadas.

45 Por consiguiente, la invención se refiere de igual modo a un método de diversificación de una proteína codificada por una secuencia nucleotídica S, caracterizada por que en un huésped de expresión apropiado se expresan las variantes polinucleotídicas de dicha secuencia nucleotídica S contenida en un producto de PCR o un banco de polinucleótidos como se ha definido previamente.

Este método puede consistir en particular por que:

- 50
- en un vector de expresión se clonan los productos de PCR obtenidos en la etapa vi) del método según la invención, que corresponden a secuencias nucleotídicas S mutadas en las que al menos un dominio D permanece igual, a continuación
  - se transforma una célula huésped que permite la expresión del polipéptido codificado por la secuencia S,
  - se expresa en dicha célula huésped dichas secuencias nucleotídicas S mutadas en las que al menos un dominio D permanece igual, para obtener proteínas diversificadas, y
  - 55 - se purifican las diferentes proteínas obtenidas.

Estas etapas se efectúan por medio de técnicas conocidas adecuadamente por el experto en la materia [Sambrook J., Russell D. W. (2001) *Molecular Cloning, a Laboratory Manual*, CHSL Press].

60 Una vez traducidas en la proteína, las secuencias nucleotídicas contenidas en el banco antes mencionado pueden dar lugar a la expresión de proteínas con una variedad de secuencias polipeptídicas producidas aleatoriamente incluyendo dominios internos  $D_N$  cuya identidad de la secuencia polipeptídica se conserva.

65 A este respecto, la invención describe un banco de proteínas resultantes de la expresión de un banco de secuencias nucleotídicas como se ha indicado previamente, dichas proteínas presentan una diversidad de secuencias polipeptídicas producidas aleatoriamente incluyendo segmentos internos  $D_N$  que conservan su identidad.

Por consiguiente, el presente método permite la modificación diana de cualquier proteína que se desea que cambie determinadas especificidades funcionales, variando aleatoriamente las secuencias de estas proteínas implicadas, por ejemplo, en las interacciones con ligandos, actividades catalíticas, toxicidad o transporte.

5 Aplicación de secuencias diversificadas obtenidas según el método de la invención para la transformación de bacteriófagos mutados en sus proteínas de orientación selectiva:

Según un aspecto preferente de la invención, los productos de amplificación obtenidos se insertan en los genes del bacteriófago codificante para sus proteínas de orientación selectiva mediante recombinación homóloga.

- 10 Las construcciones de ADN preferentes según la invención comprenden preferentemente:
- una región que permite la duplicación de dicha construcción en una bacteria huésped;
  - una región que permite la recombinación homóloga en el genoma del bacteriófago en un gen codificante de una proteína de orientación selectiva, comprendiendo dicha región dos secuencias de ADN homólogas a las secuencias de dicho gen codificante de una proteína de orientación selectiva, las cuales delimitan un segmento de inserción que incluye un oligonucleótido cuya secuencia se produce aleatoriamente, preferentemente según el método de PCR descrito previamente.

20 Según un aspecto preferente de la invención, la región que permite la recombinación homóloga comprende todo o parte del gen codificante de la proteína de orientación selectiva, preferentemente la totalidad de la secuencia del gen.

25 Preferentemente, esta segunda región consiste en un producto de amplificación que puede obtenerse según el método de mutación aleatoria expuesto previamente.

30 Según un aspecto preferente de la invención, numerosos genes codificantes de las proteínas de orientación selectiva del fago se mutan simultáneamente por recombinación homóloga según el método de la invención. Para lograr este resultado, la invención prevé transformar la bacteria huésped sucesivamente utilizando diferentes vectores, cada uno orientándose selectivamente a un gen diferente.

35 Los vectores preferentes que permiten modificar de forma simultánea los genes GP12, GP37 o GP38 del bacteriófago T4 en la bacteria huésped *E. coli*, son, por ejemplo, los vectores pACYC184 (CACT 37033), pBAD18-K (CACT 87397) y RR1 (CACT 87076). Dichos vectores ofrecen la ventaja de poseer marcadores que confieren resistencia a antibióticos diferentes y no comparten secuencias nucleotídicas comunes capaces de provocar recombinaciones entre los diferentes vectores una vez integrados en la bacteria huésped.

Un bacteriófago es preferentemente un bacteriófago lítico, natural o modificado.

40 Preferentemente, el bacteriófago utilizado es un fago de tipo T, tal como los bacteriófagos T4, T5, T6 y T7, se conoce adecuadamente por el experto en la materia y más particularmente, el fago T4, cuyo genoma se ha secuenciado [Miller, E. S. *et al.*, *Bacteriophage T4 genome*, *Mol Microbiol. Biol. Rev.*, 2003, 67(1):86-156]. La secuencia completa del genoma del bacteriófago se dispone en Genbank (AF 158101).

45 Una bacteria huésped es una bacteria utilizada habitualmente para replicar el fago que se pretende modificar. Preferentemente, la bacteria huésped es una cepa que puede transformarse con una construcción de ADN que permite modificar el fago mediante recombinación homóloga.

50 Para la transformación de esta bacteria huésped con construcciones de ADN, se disponen uno o más bancos de bacterias huésped transformadas. Cada una de las bacterias de este banco contiene potencialmente una construcción capaz de transformar por recombinación homóloga una o más proteínas de orientación selectiva del fago de manera diferente.

55 Dicho banco de bacterias huésped ofrece la ventaja de poder multiplicarse y conservarse. Constituye un producto intermedio renovable útil para la producción de bacteriófagos recombinantes cuyas proteínas de orientación selectiva se modifican aleatoriamente.

Gracias al método según la invención, es posible obtener un conjunto de bacteriófagos recombinantes muy diversificado.

60 Estos bacteriófagos forman un banco de bacteriófagos en la presente invención.

65 Para que un banco de bacteriófagos abarque el mayor número posible de bacteriófagos diferentes, es necesario aplicar una cantidad suficiente de bacterias huésped transformadas, ya que es esta cantidad la que determina el número y la diversidad de fagos recogidos.

Si esta cantidad es suficiente, un banco de fagos comprende, como mínimo, al menos  $10^6$ , preferentemente  $10^8$ , más preferentemente  $10^{10}$  variantes diferentes de un mismo bacteriófago, dichas variantes se distinguen por la secuencia de al menos una de sus proteínas de orientación selectiva.

5 La diversidad de los bacteriófagos en el banco puede justificarse por un simple cálculo de recuento.

Por consiguiente, si se considera que preferentemente:

- 10
- 3 genes codificantes de las proteínas de orientación selectiva se modifican; por inserción de al menos 3 oligonucleótidos compuestos de secuencias aleatorias de al menos 12 nucleótidos; y que
  - 1/3 de las secuencias nucleotídicas imponen una modificación polipeptídica en las proteínas de orientación selectiva; y que
  - solo 3 de las 4 bases (A, T, C, G) son propensas a producir una mutación en relación con la proteína original;

15 Se obtiene pues un mínimo de  $3^{24}$  posibilidades de mutaciones a nivel polipeptídico, lo que significa enumerar algunos  $2,8 \cdot 10^{11}$  bacteriófagos potencialmente diferentes.

Los protocolos experimentales siguientes tienen por objeto ilustrar la invención a modo de ejemplo sin limitar el alcance de la invención reivindicada.

20 Preparación de las secuencias de variantes de gp12, gp37 y gp38 de fagos T4

A continuación, se describe el procedimiento utilizado para gp12, pero puede aplicarse sin dificultad por el experto en la materia en la modificación de otros genes codificantes en particular de las proteínas de orientación selectiva.

25 Etapas 1: Preparación del gen gp12

El gen gp12 se amplifica por PCR a partir de ADN genómico del T4 de tipo natural obtenido a partir de un cultivo concentrado de lisado de fagos (véase previamente) utilizando como cebadores

30 gp12F 5'-TGAGTAATAATACATATCAACACG (SEQ ID NO. 2) y  
gp12R 5'-TGATTCTTTACCTTAATTATGTAC (SEQ ID NO. 3).

35 Tras la purificación en un gel de agarosa preparativo, el producto de PCR (gp12A) se utiliza como plantilla en las reacciones de PCR sensibles a error con el fin de introducir mutaciones e inserciones puntuales en la región codificante correspondiente al dominio de unión al receptor de gp12 (véase la figura 5).

Etapas 2: Introducción de mutaciones aleatorias en el dominio de unión al receptor.

40 Una serie de 4 reacciones de PCR encajadas (40 ciclos cada una) se implementa en presencia de Mn<sup>2+</sup> de manera que se induzcan errores de polimerasa aleatorios.

45 Las PCR propensas a error se llevan a cabo en un volumen de reacción de 100 µl, utilizando plantillas y cebadores en concentraciones finales respectivamente iguales a 400 ng y 30 pmol, con dATP 0,2 mM y dGTP (cada uno), 1 mM de dCTP y dTTP (cada uno), 2,5 mM de MgCl<sub>2</sub>, 0,7 mM de MnCl<sub>2</sub> y 5 U de ADN polimerasa *Taq* (New England Biolabs, Inc.) en tampón de reacción 1X. La PCR se efectúa a 96 °C durante 2 min, con 30 ciclos a 95 °C durante 1 min, 56 °C durante 1 min y 72 °C durante 2 min, y una extensión final a 72 °C durante 7 min.

La primera reacción (P1-1) utiliza como cebadores:

50 p12NF1F 5'-TCAAGGTAACCGCATCGTAAC (SEQ ID NO. 4)  
p12NF2R 5'-AAAGACCACGCATGTCAG (SEQ ID NO. 5)

La segunda reacción (P2-1) utiliza como cebadores:

55 p12NF2F 5'-TGCCATGGTGGAAGTGTCA (SEQ ID NO. 6)  
p12NF3R 5'-CACCTAATCTAGGTTTAC (SEQ ID NO. 7)

La tercera reacción (P3-1) utiliza como cebadores:

60 p12NF3F 5'-CTGACATGCGTGGTCTTT (SEQ ID NO. 8)  
p12NF4R 5'-ATGTTTATGATAAGACAT (SEQ ID NO. 9)

La cuarta reacción (P4-1) utiliza como cebadores:

65 p12NF4F 5'-GTAAACCTAGATTAGGTG (SEQ ID NO. 10)

p12NF5R 5'-TCATTCTTTTACCTTAATTAT (SEQ ID NO.11)

Cada uno de estos productos de reacción presenta una superposición parcial con otros dos productos de reacción y los cebadores utilizados corresponden a los dominios invariantes conservados en la proteína gp12. Estos dominios deben preservarse fielmente en las estructuras producidas de genes mutados finales. No obstante, algunos de los fragmentos producidos por las reacciones de PCR sensibles a error pueden mutagenizarse perfectamente en estas regiones. A fin de preservar los dominios intactos, cada uno de los productos mencionados previamente ha de someterse a reacciones de PCR de alta fidelidad que tienen por objeto amplificar selectivamente solo los fragmentos en los que se ha retenido el dominio conservado.

Etapa 3: Amplificación selectiva de los fragmentos deseados.

Esto se efectúa mediante dos series de reacciones de PCR de alta fidelidad que comprenden cada una 25 ciclos.

Las PCR de alta fidelidad se llevan a cabo en un volumen de reacción de 50 µl, utilizando plantillas y cebadores en concentraciones finales respectivamente iguales a 250 ng y 40 pM, en tampón para *pfu* 1X (20 mM de Tris-HCl en pH 9,0, 10 mM de KCl, 1 mM de MgSO<sub>4</sub>, 6 mM de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, Tritón X-100 al 0,1 %, 0,1 mg/ml de SAB) con 200 µM de dNTP y 5 U de polimerasa *pfu* (Promega). Los perfiles de PCR son los siguientes: 94 °C durante 20 s, 45 °C durante 15 s y 72 °C durante 30 s, con repetición durante 20 ciclos.

En la primera serie de reacciones (P1-2 a P4-2), una porción alícuota (aproximadamente 250 ng) de cada uno de los productos de PCR previos se amplifica utilizando los cebadores "F" correspondientes (por ejemplo: p12NF1F) biotinilados al extremo 5'. Esto es necesario para separar los productos deseados de las plantillas no amplificables (dominios conservados mutados) con longitudes prácticamente idénticas y que no pueden separarse por electroforesis en gel.

Los productos de cada reacción se pasan entonces en minicolumnas de exclusión por tamaños (para eliminar el exceso de cebadores), se purifican individualmente en microesferas de estreptavidina, se lavan en tampón de unión, se eluyen, se precipitan y se vuelven a suspender en ddH<sub>2</sub>O.

En la segunda serie de reacciones, una porción alícuota de los productos de reacción P1-2 purificados se mezcla con una cantidad igual a los productos de reacción P2-2. Estos tienen en común el sitio de cebado F2-R1 y F3-R2. En consecuencia, la cadena "-" de los productos de reacción P2-2 sirve como cebador para la cadena "+" de los productos de reacción P1-2 y la extensión con una polimerasa comienza en el sitio de cebado F2-R1 interno de P1-2. La extensión falla para los productos de reacción P1-2 en los que este sitio conservado se ha mutado de manera significativa.

Del mismo modo, la cadena "+" de los productos de reacción P1-2 sirve como cebador para la cadena "-" de los productos de reacción P2-2 y la extensión con una polimerasa comienza en el sitio de cebado F3-R2 interno de P2-2. La extensión falla para los productos de reacción P2-2 en los que este sitio conservado se ha mutado de manera significativa.

Los productos de PCR amplificados con éxito (PA-1) corresponden a una fusión de fragmentos P1-2 y P2-2, en los que en cada uno de los cuatro dominios conservados (F1 + F2-R1+ F3-R2 + F4-R3) se retuvieron en estado no mutado.

Una mezcla similar se prepara a partir de los productos de reacción P3-2 y P4-2. En estas reacciones, una extensión exitosa se basa en los sitios conservados internos F4-R3 (P3-2) y F5-R4 (P4-2) intactos y los productos de PCR resultantes (AP-1 y PB-1) corresponden a una fusión de los fragmentos P3-2 y P4-2, en los que cada uno de los cuatro dominios conservados (F3-R2 + F4-R3 + F5-R4 + R5) se obtuvieron en estado no mutado.

Los productos obtenidos se purifican por electroforesis en gel preparativo dado que las dimensiones de los productos de PCR fusionados son muy diferentes de las de las plantillas individuales.

A fin de aumentar el rendimiento en esta segunda serie de reacciones PCR, el protocolo de extensión puede modificarse de la siguiente manera:

10 ciclos con solo los productos P1-2 y P2-2, y P3-2 y P4-2 en las mezclas de reacción.

Las reacciones se interrumpen, 50 ng de cada uno de los cebadores p12NF1F y p12NF3R se añaden a la mezcla PA-1 mientras que 50 ng de cada uno de los cebadores p12NF3F y p12NF5R se añaden a la mezcla PB-1, permitiendo continuar las reacciones PCR durante 15 ciclos suplementarios.

Etapa 4: Reconstitución del dominio de anclaje mutado de gp12.

Tras la purificación, las porciones alícuotas iguales de los productos PA-1 y PB-1 se utilizan entre sí en una reacción

PCR de alta fidelidad que comprende 30 ciclos con el objetivo de reconstituir el dominio total de unión al receptor de gp12 en fragmentos únicos que contienen las diversas mutaciones introducidas previamente.

5 En este caso, como en la serie precedente de reacciones, los productos PA-1 se utilizan como cebadores de extensión para los productos PB-1 y viceversa.

El producto de PCR fusionado final (gp12BD-Fu) se purifica por electroforesis en gel dado que sus dimensiones son muy diferentes a las de las plantillas individuales.

10 A fin de aumentar el rendimiento, el protocolo de extensión puede modificarse mediante la realización de una reacción que comprende 15 ciclos con solo los productos P1-2 y P2-2, y P3-2 y P4-2 en las mezclas de reacción. Las reacciones se interrumpen, 50 ng de cada uno de los cebadores p12NF1F y p12NF5R se añaden a la mezcla, a continuación, se permite continuar las reacciones PCR durante 15 ciclos suplementarios.

15 Tras la purificación, una porción alícuota (aproximadamente 250 ng) del dominio de unión al receptor de gp12 reconstituido se utiliza como plantilla en una nueva serie de cuatro reacciones PCR sensibles a error, seguido por amplificaciones selectivas y reconstituciones de los dominios como se ha descrito previamente.

20 No obstante, dado que en cada etapa de fusión (etapas 3 y 4 previas) todas las posibles mutaciones introducidas individualmente en cada subdominio se fabrican aleatoriamente en el dominio reconstituido final de unión al receptor, existen limitaciones en cuanto al número de ciclos de mutagénesis que pueden efectuarse sin efecto negativo en los dominios conservados. Debido a las etapas necesarias de amplificación selectiva, se estima que, más allá de 4 ciclos sucesivos, todas las mutaciones introducidas recientemente presentan una importante probabilidad de efectos negativos en un dominio conservado o de introducción de una reversión que restituye la secuencia T4 original, no mutada.

Etapa 5: Reconstitución de la copia modificada del gen gp12 en su conjunto

1) Amplificación del segmento de gp12 corriente arriba del dominio de unión al receptor.

30 El gen gp12, generado por PCR, producido previamente (gp12A) se utiliza como plantilla junto con los cebadores

gp12F 5'-TGAGTAATAATACATATCAACACG (SEQ ID NO.12), y  
gp12AR 5'-GTTACGATGCGGTTACCTTGT (SEQ ID NO.13)

35 De ello resulta una purificación de los productos de amplificación obtenidos en un gel de agarosa preparativo, una precipitación y a continuación una resuspensión en ddH<sub>2</sub>O.

2) Amplificación del dominio de unión mutado de gp12

40 Una porción alícuota (aproximadamente 500 ng) del producto de PCR (gp12B) se utiliza junto con una cantidad igual del dominio de unión al receptor de gp12 reconstituido (gp12BD-Fu) en una reacción PCR de alta fidelidad. Cabe observar que estos dos productos presentan solamente una superposición en común de 20 pb. En consecuencia, el perfil de la reacción PCR ha de modificarse para tener esto en cuenta.

45 Perfil de PCR modificado: A) 30 min a 96 °C seguido de 5 ciclos de 1 min a 94 °C, 10 s a 45 °C, y a continuación B) 5 ciclos de 1 min a 94 °C y 20 s a 50 °C, después C) 5 ciclos de 1 min a 94 °C, 30 s a 50 °C y, finalmente, D) 15 ciclos de 1 min a 94 °C, 1 min a 55 °C, 5 segundos a 72 °C.

50 A fin de aumentar el rendimiento, la reacción puede interrumpirse en la subetapa D previa y los cebadores gp12F y p12NF5R pueden introducirse en la mezcla. A continuación, se deja que la reacción se recupere y se realiza hasta que esta finalice.

55 Los productos finales (gp12-Mut) se purifican en un gel de agarosa preparativo, se precipitan y se vuelven a suspender en ddH<sub>2</sub>O hasta su utilización.

Etapa 6: Verificación por secuenciación de los productos de PCR gp12-Mut obtenidos

60 Los productos de PCR purificados en gel de agarosa se secuencian utilizando un secuenciador automático con el fin de verificar que las mutaciones se han introducido correctamente de forma aleatoria en gp12 mientras se conservan intactos los dominios D1 y D2. La secuenciación se efectúa en dos porciones correspondientes, respectivamente a los nucleótidos 1089 a 1110 (Fig. 6A) y 1195 y 1212 (Fig. 6B) de gp12 (SEQ ID NO.1). Los perfiles de secuenciación obtenidos se muestran en la Fig. 6. Comparando los perfiles de los productos de PCR con respecto a los de la secuencia inicial de gp12, puede notarse que numerosos picos suplementarios se agregan al perfil esperado de gp12. Estos picos adicionales traducen una frecuencia elevada de mutaciones introducidas en la secuencia por el método de PCR. Estos picos no se encuentran en el dominio D1, lo que confirma que este dominio ha conservado

su identidad de secuencia con el de gp12.

Clonación de gp12-Mut mutados en un vector alternativo para recombinación homóloga

5 Los procedimientos previos han permitido generar, de forma similar a gp12, los genes mutantes gp37 y gp38 de T4.

El objetivo de los siguientes procedimientos es introducir estos genes de T4 variantes de manera que favorezca su introducción, por recombinación homóloga, en el genoma del bacteriófago T4 de modo que produzca progenies de T4 con especificidades de orientación selectiva de los huéspedes diferentes al bacteriófago parental.

10 A tal efecto, los genes mutados generados previamente han de introducirse en vectores alternativos (un vector diferente para cada gen). En el caso presente, los vectores pACYC184, pBAD18-K y RR1 se utilizan para clonar cada uno genes mutados.

15 El vector seleccionado (por ejemplo pACYC184 para la clonación de gp12Mut) se corta en el sitio Sma I, se purifica y se vuelve a suspender en 20 µl de ddH<sub>2</sub>O y se mezcla con el producto de PCR gp12Mut en una relación de 1:3. Tras el ligamento, se purifica el ADN, se vuelve a suspender en ddH<sub>2</sub>O y se transforma en células DK8 (CACT 47038) "electrocompetentes". Este vector lleva un gen de resistencia al cloranfenicol (Chl). Por consiguiente, tras la electroporación y un periodo de recuperación de 1 h, las células se transfieren a 10 ml de medio LB que contiene  
20 170 µg/ml de cloranfenicol y se cultivan a 30 °C durante 4 h con ventilación constante. Las células se siembran en placas de medio LB + Chl y se cultivan durante una noche a 30 °C. Algunas colonias se recogen para la verificación por PCR de la presencia del segmento de inserción gp12Mut y algunas colonias positivas se cultivan en medio LB + Chl para la preparación de un cultivo concentrado de células (DK8-p12C).

25 Introducción de genes mutados gp38 y gp37 en la bacteria huésped

Un cultivo fresco de una noche de DK8-p12C se utiliza para la preparación de células electrocompetentes.

30 El vector pBAD18-K se corta en el sitio Sma I, se purifica, se vuelve a suspender en 20 µl de ddH<sub>2</sub>O y se mezcla con el producto de PCR gp37Mut en una relación de 1:3. Tras el ligamento, se purifica el ADN, se vuelve a suspender en ddH<sub>2</sub>O y se transforma en células DK8-p12C "electrocompetentes".

35 El vector utilizado en este caso lleva un gen de resistencia a kanamicina (Kan). Por consiguiente, tras la electroporación y un periodo de recuperación de 1 h, las células se transfieren a 10 ml de medio LB que contiene 170 µg/ml de cloranfenicol y 50 µg/ml de kanamicina y se cultivan a 30 °C durante 4 h con ventilación constante. Las células se siembran en placas de medio LB + Chl + Kan y se cultivan durante una noche a 30 °C. Algunas colonias se recogen para la verificación por PCR de la presencia del segmento de inserción gp12Mut y gp37Mut y algunas colonias positivas se cultivan en medio LB + Chl + Kan para la preparación de un cultivo concentrado de células (DK8-p12C-p37K).

40 Un cultivo fresco de una noche de DK8-p12C-p37K se utiliza para la preparación de células electrocompetentes. El vector RR1 se corta en el sitio Sma I, se purifica, se vuelve a suspender en 20 µl de ddH<sub>2</sub>O y se mezcla con el producto de PCR gp37Mut en una relación de 1:3. Tras el ligamento, se purifica el ADN, se vuelve a suspender en ddH<sub>2</sub>O y se transforma en células DK8-p12C-p37K "electrocompetentes". El vector utilizado en este caso lleva un  
45 gen de resistencia a ampicilina (Amp). Por consiguiente, tras la electroporación y un periodo de recuperación de 1 h, las células se transfieren a 10 ml de medio LB que contiene 170 µg/ml de cloranfenicol + 50 µg/ml de kanamicina + 60 µg/ml de ampicilina y se cultivan a 30 °C durante 4 h con ventilación constante. Las células se siembran en placas de medio LB + Chl + Kan + Amp y se cultivan durante una noche a 30 °C. Algunas colonias se recogen para la verificación por PCR de la presencia del segmento de inserción gp12Mut, gp37Mut y gp38Mut y algunas colonias  
50 positivas se cultivan en medio LB + Chl + Kan + Amp para la preparación de un cultivo concentrado de células (DK8-p12C-p37K-P38A).

Queda ahora construir un huésped capaz de presentar una potencia de recombinación sumamente eficaz.

55 Construcción de bacterias huésped *E. coli* "mini-λ"

Para obtener la recombinación eficaz del ADN donante en los fondos *recA*<sup>+</sup> o *recA*<sup>-</sup>, se prepara el huésped *E. coli* que contiene un profago λ que lleva los genes de recombinación *exo*, *bet* y *gam* bajo el control de un represor cl de λ sensible a la temperatura. Los genes *exo*, *bet* y *gam* pueden activarse fácilmente a 42 °C y se inhiben a 32 °C.  
60 Cuando las funciones de λ se activan durante un breve periodo de 5 min, las células se vuelven más recombinógenas y absorben el ADN lineal sin su destrucción, λ Gam inhibe el ataque del ADN lineal por la nucleasa RecBCD de *E. coli* y *Exo* y *Beta* generan una actividad de recombinación para este ADN lineal. Más importante aún, esta recombinación es eficaz con homología de ADN limitadas a 30 a 50 pb en los extremos de los sustratos que consisten en ADN lineal.

65 Los oligonucleótidos 5' GTATGCATGCTGGGTGTGG (MARf) y 5' CGCACTCTCGATTCGTAGAGCCTCG (MARr) se

## ES 2 582 385 T3

utilizan como cebadores para la amplificación por PCR del ADN de la región *attP-cro* utilizando ADN de  $\lambda$  *cl857* como plantilla.

5 Una vez generado el profago  $\lambda$  por PCR, una posibilidad consiste en clonar el profago  $\lambda$  en el plásmido pFN476 (CACT 86962) en un número reducido de copias con selección por LacZ, por ejemplo.

10 El vector PFN476 se corta en el sitio Sma I (extremo romo), se purifica, se vuelve a suspender en 20  $\mu$ l de ddH<sub>2</sub>O y se mezcla con el producto de PCR del profago  $\lambda$  en una relación de 1:3. Tras el ligamento, el DNA se purifica, se vuelve a suspender en ddH<sub>2</sub>O y se transforma en células DK8-p12C-p37K-P38A "electrocompetentes". Tras la recuperación, las células se siembran en placas de medio LB X-gal + Chl + Kan + Amp (véase previamente) y se incuban a 30 °C.

15 Algunas colonias blancas se seleccionan para la verificación por PCR de la presencia del profago  $\lambda$ . Una colonia positiva de profagos  $\lambda$  se cultiva a continuación durante una noche a 30 °C en medio LB + Chl + Kan + Amp para la preparación de un cultivo concentrado de células (DK8-T4Mut- $\lambda$ ) para su uso en posteriores manipulaciones.

En esta etapa, los huéspedes transformados se inducen por  $\lambda$  a temperatura elevada, lacZ positivas y contienen copias de genes gp12, gp37 y gp38 de T4 mutados preparados para la recombinación homóloga.

### 20 Producción de la progenie del bacteriófago T4 con intervalos de huésped amplios

Un cultivo fresco de una noche con células DK8-T4Mut- $\lambda$  se prepara en medio LB + Chl + Kan + Amp a 30 °C.

25 Los cultivos para la infección por T4 se inician con un volumen igual o inferior a 0,05 ml de células de un cultivo de una noche para 10 ml de medio LB + Chl + Kan + Amp a fin de garantizar que las células pasen a la fase de crecimiento exponencial antes de la adición del bacteriófago.

30 Para mejorar la ventilación, estos cultivos se multiplican en matraces de Erlenmeyer de 250 ml con brazo lateral, se taponan de forma no hermética en un baño de agua con agitación a 30 °C.

250 ml de células en medio LB + Chl + Kan + Amp se cultivan a una densidad de  $3 \times 10^8$  células por ml a 30 °C con agitación.

35 Las porciones alícuotas de 10 ml de células en crecimiento exponencial se transfieren a 40 ml de medio LB + Chl + Kan + Amp precalentado a 42 °C y se incuban durante exactamente 15 min a 42 °C con ventilación constante. Se añade triptófano a una concentración de 0,02 mg/ml y seguido por el bacteriófago T4 con una multiplicidad de aproximadamente 10 partículas por célula. Los cultivos se transfieren a un baño de agua a 30 °C, permitiendo el crecimiento durante exactamente 25 min.

40 El objetivo en este caso consiste en aislar la progenie de la primera ráfaga y detener la propagación antes de que esta progenie de bacteriófagos de primera generación pueda transformarse en reproductores.

### Recuperación de la progenie de los bacteriófagos

45 Las células se recogen por centrifugación a 5.000 r/min durante 5 minutos y se recupera el sobrenadante, se añaden unas gotas de cloroformo y la mezcla se centrifuga de nuevo durante 10 minutos a 6.000 r/min. El sobrenadante, excluyendo el cloroformo, se ajusta a una concentración en tampón SM 1X (10 mM de MgSO<sub>4</sub>, 100 mM de NaCl, gelatina al 0,01 % y 50 mM de Tris-HCl [pH 7,5]) utilizando un tampón concentrado 5X, y se almacena a 4 °C antes del análisis.

50 Las células sedimentadas se vuelven a suspender en 8 ml de solución de Tris-clorhidrato 0,05 M en pH 8,0 con sacarosa al 25 %. Se añadió una porción alícuota de 1,6 ml de lisozima (5 mg/ml) y la mezcla se incubó durante 5 min a 0 °C. Se añade una porción alícuota de 3,2 ml de solución EDTA 0,2 M y la mezcla se incubó durante un periodo suplementario de 15 minutos a 0 °C. El lisado celular se ajusta a 500 mM de Tris-HCl en pH 7,4, 100 mM de MnCl<sub>2</sub> con tampón concentrado 10X, se equilibra a 15 °C y se incubó con 10 U de Dnasa 1 (Sigma) durante 2 h. La mezcla se centrifuga a 6.000 r/min durante 10 min. El sobrenadante límpido se retira cuidadosamente, se ajusta a una concentración en tampón SM 1X y se almacena como anteriormente.

### Verificación de las extensiones de los intervalos de huésped

60 Se preparan cultivos de cepas bacterianas patógenas (*Yersinia sp.*, *Salmonella sp.*, *E. coli* O157 H7, *Enterobacter sakazakii*, etc.).

65 Las porciones alícuotas de 3 ml se infectan a continuación con 1 ml de cultivo concentrado de la progenie y se incuban a 30 °C con agitación. Se determinó la turbidez del cultivo por colorimetría inmediatamente después de la infección y luego cada 60 min. Una caída significativa de la turbidez del cultivo en un periodo de 5 horas indica que

las partículas capaces de infectar el huésped ensayado estaban presentes en la progenie de T4 recombinante. Se añaden algunas gotas de cloroformo al cultivo o a los cultivos, lo que provoca una caída repentina de la turbidez, y los cultivos se centrifugan a 5.000 r/min durante 5 min.

- 5 El sobrenadante se recupera y se utiliza como sustancia infecciosa para la producción de partículas de bacteriófagos asignados a la bacteria ensayada, que en la naturaleza, no puede ser atacada por el bacteriófago natural de tipo T4.

LISTADO DE SECUENCIAS

- 10 <110> PHERECYDES PHARMA
- <120> Método de diversificación aleatoria de una secuencia genética que permite preservar la identidad de determinados segmentos internos de dicha secuencia genética
- 15 <130> BIF 117147 WO
- <160> 13
- <170> PatentIn version 3.1
- 20 <210> 1  
<211> 1582  
<212> ADN  
<213> bacteriófago T4
- 25 <220>  
<221> misc\_feature  
<222> (1) .. (1582)  
<223> Región codificante del gen gp12 del bacteriófago T4
- 30 <400> 1

ES 2 582 385 T3

gagtaataat acatatcaac acgtttctaa tgaatctcgt tatgtaaaat ttgatcctac 60  
 cgatacgaat tttccaccgg agattactga tgttcacgct gctatagcag ccatttctcc 120  
 tgctggagta aatggagttc ctgatgcacg gtcaacaaca aagggaaatc tatttattcc 180  
 cactgaacag gaagttatag atggaactaa taataccaaa gcagttacac cagcaacggt 240  
 ggcaacaaga ttatcttctc caaatgcaac tgaaactggt tacggattaa caagatattc 300  
 aaccaatgat gaagccattg ccggagttaa taatgaatct tctataactc cagctaaatt 360  
 tactgtcgcc ctttaataatg cgtttgaaac gcgagtttca actgaatcct caaatgggtg 420  
 tattaanaat tcatctctac cgcaagcatt agctgggtga gatgatacta ctgcaatgac 480  
 tccattaanaa acacagcagt tagctattaa attaattgag caaattgctc cttctgaaac 540  
 cacagctacc gaatcggacc aaggtgttgt tcaattagca acagtagcgc aggttcgtca 600  
 gggaaacttta agagaaggct atgcaatttc tcttatacag tttatgaatt catcttctac 660  
 tgaagaatat aaaggcgtaa ttaaattagg aacacaatca gaagttaact cgaataatgc 720  
 ttctgttgcg gttactggcg caactcttaa tggctcgtgt tctacgacgt caatgagagg 780  
 cgtagttaa ttaactacaa ccgccggtc acagagtgga ggcgatgctt catcagcctt 840  
 agcttggaat gctgacgta tccagcaaag aggtgggtcaa attatctatg gaacactccg 900  
 cattgaagac acatttacia tagctaattg tggagcaaat attacgggta ccgtcagaat 960  
 gactggcgggt tatattcaag gtaaccgcat cgtaacacia aatgaaattg atagaactat 1020  
  
 tctgtcggg gctattatga tgtgggccc tgatagtctt cctagtgatg cttggcgctt 1080  
 ctgccatggt ggaactgttt cagcgtcaga ttgtccatta tatgcttcta gaattggaac 1140  
 aagatatggc ggaaccat caaatcctgg attgcctgac atgcgtggctc tttttgttcg 1200  
 tggttctggt cgtggttctc acttaacaaa tccaaatggt aatggtaatg accaatttgg 1260  
 taaacctaga ttaggtgtag gttgtaccgg tggatatggt ggtgaagtac agatacaca 1320  
 gatgtcttat cataaacatg ctgggtgatt tgggtgagcat gatgatctgg gggcattcgg 1380  
 taatacccgt agatcaaat ttgttggtac acgtaaagga cttgactggg ataaccgttc 1440  
 atacttcacc aatgacggat atgaaattga ccagaaatca caacgaaatt ccaaatatac 1500  
 attaaatcgt cctgaattaa ttggaatga aacacgtcca tggaaacatt ctttaacta 1560  
 cataattaag gtaaaagaat ga 1582

<210> 2  
 <211> 24  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 <221> Cebador  
 10 <222> (1) .. (24)

## ES 2 582 385 T3

<223> Cebador gp12F para la amplificación del gen gp12 del bacteriófago T4

5 <400> 2  
 tgagtaataa tacatatcaa cacg 24

<210> 3  
 <211> 25  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <221> Cebador  
 <222> (1)..(25)  
 <223> Cebador gp12R para la amplificación del gen gp12 del bacteriófago T4

15 <400> 3  
 tgattctttt accttaatta tgtac 25

<210> 4  
 <211>21  
 <212> ADN  
 <213> Secuencia artificial

25 <220>  
 <221> Cebador  
 <222> (1)..(21 )  
 <223> Cebador P12NF1F interno el dominio del anclaje de gp12 del bacteriófago T4

30 <400> 4  
 tcaaggtaac cgcacgtgaa c 21

<210> 5  
 <211> 18  
 <212> ADN  
 <213> Secuencia artificial

35 <220>  
 <221> Cebador  
 <222> (1)..(18)  
 <223> Cebador P12NF2R interno del dominio de anclaje de gp12 del bacteriófago T4

40 <400> 5  
 aaagaccacg catgtcag 18

45 <210> 6  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial

50 <220>  
 <221> Cebador  
 <222> (1)..(20)  
 <223> Cebador P12NF2F interno del dominio de anclaje de gp12 del bacteriófago T4

55 <400> 6  
 tgccatggtg gaactgttca 20

<210> 7  
 <211> 18  
 <212> ADN  
 <213> Secuencia artificial

60 <220>  
 <221> Cebador  
 <222> (1)..(18)  
 <223> Cebador P12NF3R interno del dominio de anclaje de gp12 del bacteriófago T4

65

ES 2 582 385 T3

	<400> 7 cacctaactc acgtttac	18
5	<210> 8 <211> 18 <212> ADN <213> Secuencia artificial	
10	<220> <221> Cebador <222> (1)..(18) <223> Cebador P12NF3F interno del dominio de anclaje de gp12 del bacteriófago T4	
15	<400> 8 ctgacatgcg tggcttt	18
20	<210> 9 <211> 18 <212> ADN <213> Secuencia artificial	
25	<220> <221> Cebador <222> (1)..(18) <223> Cebador P12NF4R interno del dominio de anclaje de gp12 del bacteriófago T4	
30	<400> 9 atgttatga taagacat	18
35	<210> 10 <211> 18 <212> ADN <213> Secuencia artificial	
40	<220> <221> Cebador <222> (1)..(18) <223> Cebador P12NF4F interno del dominio de anclaje de gp12 del bacteriófago T4	
45	<400> 10 gtaaaccctag attaggtg	18
50	<210> 11 <211> 21 <212> ADN <213> Secuencia artificial	
55	<220> <221> Cebador <222> (1) .. (21) <223> Cebador P12NF5R interno del dominio de anclaje de gp12 del bacteriófago T4	
60	<400> 11 tcattctttt accttaatta t	21
65	<210> 12 <211> 24 <212> ADN <213> Secuencia artificial	
	<220> <221> Cebador <222> (1) .. (24) <223> Cebador gp12F para la amplificación del gen gp12 del bacteriófago T4	

# ES 2 582 385 T3

<400> 12  
tgagtaataa tacatatcaa cacg 24

5 <210> 13  
<211> 21  
<212> ADN  
<213> Secuencia artificial

10 <220>  
<221> Cebador  
<222> (1)..(21)  
<223> Cebador gp12AR para la amplificación del gen gp12 del bacteriófago T4

15 <400> 13  
gttacgatgc ggttaccttg t 21

## REIVINDICACIONES

1. Método de PCR que permite mutar aleatoriamente una secuencia nucleotídica S, delimitada en sus extremos 5' y 3' por dos segmentos F1 y F2, mientras conserva al mismo tiempo la identidad de un segmento interno D de dicha secuencia nucleotídica S, **caracterizado por que** comprende las siguientes etapas:
- 5
- i) se efectúa una PCR propensa a error en toda la secuencia S utilizando al menos dos cebadores, uno de los cuales es sentido y el otro antisentido, respectivamente, de los segmentos F1 y F2, de modo que se amplifica la secuencia S al introducir mutaciones aleatoriamente;
  - 10 ii) se purifican los productos de amplificación obtenidos, que corresponden a las copias de S mutadas aleatoriamente;
  - iii) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación purificados en la etapa ii), utilizando pares de cebadores sentido y antisentido, respectivamente, al menos:
  - 15 - un cebador sentido de F1 y un cebador antisentido se hibridan en su totalidad con el segmento D, a fin de amplificar la región F1-D de las copias de S mutadas,
  - un cebador sentido se hibrida en su totalidad con el segmento D y un cebador antisentido de F2, a fin de amplificar la región D-F2 de las copias de S mutadas;
  - 20 iv) se purifican los productos de amplificación obtenidos en la etapa iii), que consisten en copias de los segmentos F1-D y D-F2 de la secuencia S mutada en la etapa i) en las que el segmento D ha conservado su identidad;
  - v) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación purificados en la etapa iv) utilizando los cebadores sentido y antisentido, F1 y F2 respectivamente;
  - 25 vi) se purifican los productos de PCR obtenidos, que corresponden a las secuencias nucleotídicas S mutadas en la etapa i) en las que el segmento D ha conservado su identidad.
2. Método para diversificar por PCR una secuencia nucleotídica S, delimitada en sus extremos 5' y 3' por dos segmentos F1 y F2, mientras conserva al mismo tiempo la identidad de dos dominios internos D<sub>1</sub> y D<sub>2</sub> de dicha secuencia nucleotídica S, **caracterizado por que** comprende las siguientes etapas:
- 30
- i) se efectúa una PCR propensa a error en toda la secuencia S utilizando al menos dos cebadores, uno de los cuales es sentido y el otro antisentido, respectivamente de los segmentos F1 y F2, de modo que amplifica la secuencia S introduciendo mutaciones aleatoriamente;
  - 35 ii) se purifican los productos de amplificación obtenidos;
  - iii) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación purificados en la etapa ii), utilizando pares de cebadores sentido y antisentido, al menos:
  - 40 - un cebador sentido de F1 y un cebador antisentido se hibridan en su totalidad con D<sub>2</sub>, a fin de amplificar la región F1-D<sub>2</sub> de las copias de S mutadas, en las que el segmento D<sub>2</sub> ha conservado su identidad;
  - un cebador antisentido de F2 y un cebador sentido se hibridan en su totalidad con D<sub>1</sub>, a fin de amplificar la región D<sub>1</sub>-F2 de S mutada, en la que el segmento D<sub>1</sub> ha conservado su identidad;
  - 45 iv) se efectúa una PCR de alta fidelidad a partir de los productos de amplificación obtenidos en la etapa iii) utilizando al menos un par de cebadores sentido y antisentido de F1 y F2;
  - v) se purifican los productos de PCR obtenidos en la etapa iv), que corresponden a las secuencias nucleotídicas S mutadas aleatoriamente, en las que las secuencias de los segmentos D<sub>2</sub> y D<sub>1</sub> han conservado su identidad.
3. Método según las reivindicaciones 1 o 2, **caracterizado por que** la secuencia S es una secuencia genética que codifica una proteína de interés.
- 50
4. Método según la reivindicación 3, **caracterizado por que** la proteína de interés es una proteína de orientación selectiva de un bacteriófago, tal como la proteína GP12 de un bacteriófago de tipo T.
- 55
5. Método según la reivindicación 3, **caracterizado por que** la proteína de interés es una proteína ligando que se pretende que modifique la especificidad con respecto a un receptor.
6. Método según la reivindicación 3, **caracterizado por que** la proteína de interés es una inmunoglobulina que se pretende que modifique los dominios variables o hipervariables.
- 60

# El fago T4

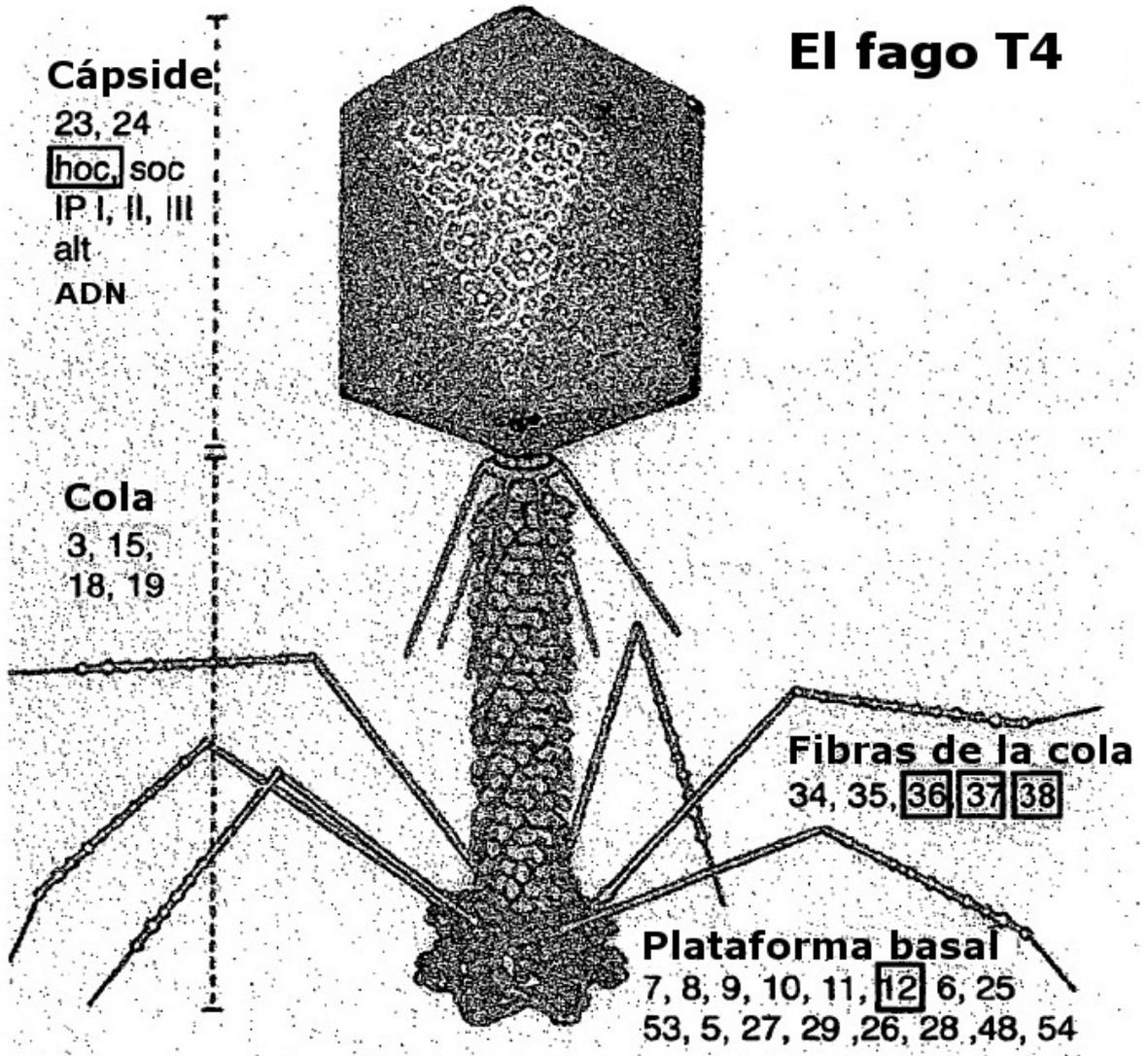


Fig. 1

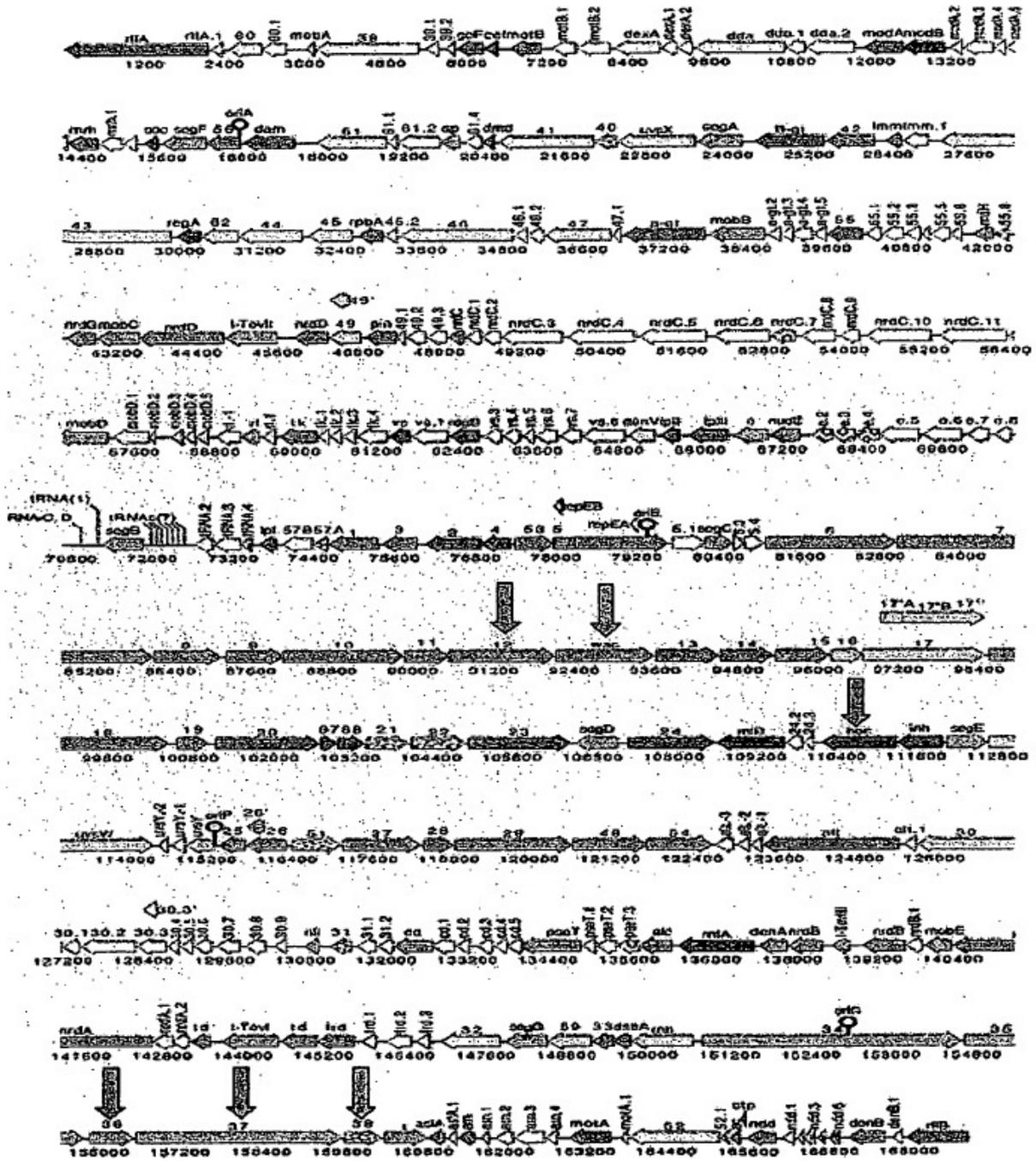


Fig. 2

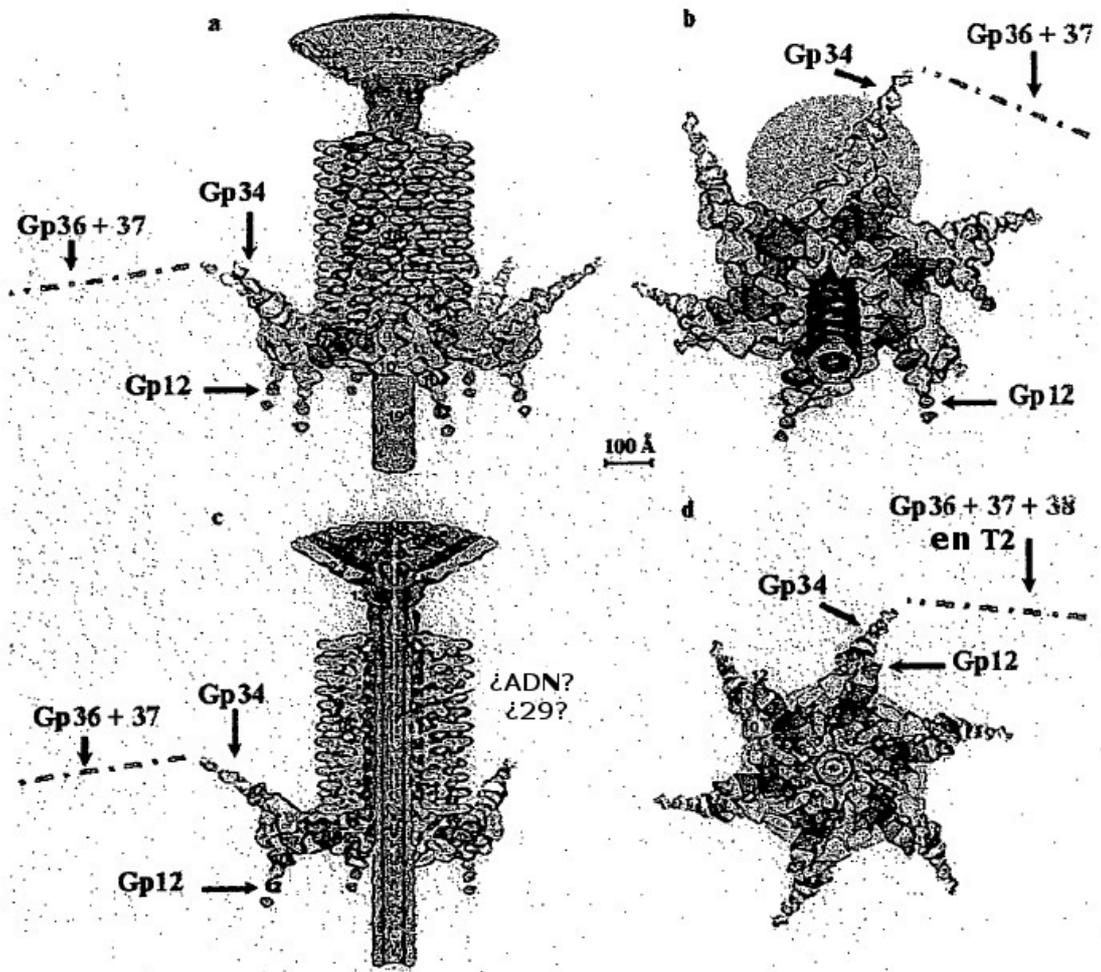


Fig. 3

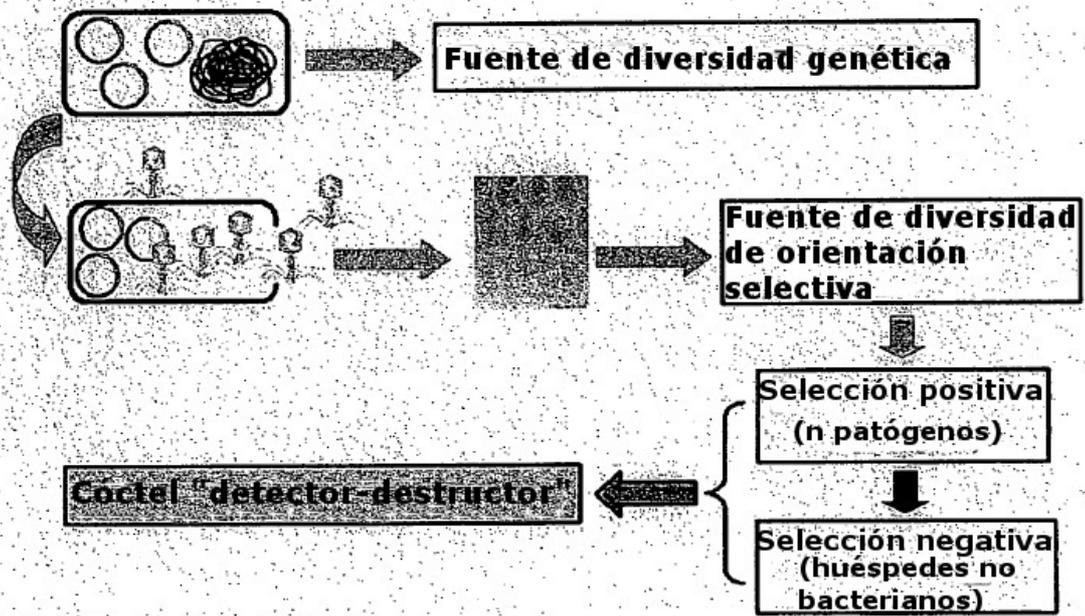


Fig. 4

1 MSNNTYQHVSNESRYVKFDPTDTNFPPEITDVHAAIAAISPAQVNGVVPDASSTTKGILFI 60  
 MSNNTYQHVSNES YV+FDPT +NF IT+V AA+A+IS GV GVPDAS KG++ +  
 1 MSNNTYQHVSNESVYVEFDPTGSNFDSSITNVQAALASISAYGVKGVDPDASEAEKGVIOQL 60

61 PTEQEVIDGTNNTKAVTPATLRLSYPNATETVYGLTRYSTNDEAIAGVNNESITPAK 120  
 TEQEV+DG N+TKAVTPATL RL YPNA+ET YG+T+Y+T +EAIAG + SITP K  
 61 ATEQEVLDGFNSTKAVTPATLNARLQYPNASETQYGVTKYATQEEAIAGTLDTVSITPLK 120

121 FTVALNNAFETRVSTESSNGVIKISSLPQALAGADDTTAMTPLKTOQLAIKLIQIAPSE 180  
 ++N F TR STE++NGVIKI++ ALAG+DDTTAMTPLKTOQLAIKLI+QIAP+  
 121 LNQTIDNTFSTRYSTETTINGVIKIATQTAALAGSDDTTAMTPLKTOQLAIKLISQIAPNN 180

181 TTATESDQGVVQLATVAQVRQGTLRREGYAISPYTFMNSSSTEYKGVIKLGTQSEVNSNN 240  
 A+ES GVV+LATVAQ RQGTLRREGYAISPYTFMNS +T+EYKGV+LGTQ+E+NSN  
 181 DPASESITGVVRLATVAQTRQGTLRREGYAISPYTFMNSVATQYKGVIRLGTQAEINSNL 240

241 ASVAVTGATLNGRGSTTSMRGVVKLTFTTAGSQSGDASSALAWNADVIQRRGGQIITYGTL 300  
 VAVTG TLNGRG+T SMRGVVKLT AG GD+S ALAWNADVI RGGQ I G+L  
 241 GDVAVTGETLNGRGATGSMRGVVKLTQAGVAPEGDSSGALAWNADVINTRGGQTINGSL 300

301 RIEDTFTIANGGANITGTVRMTGG-YIQGN [E1] DRTIPVGAIMMWAADSLPSDA 358  
 ++ ANG + GG + G+ [E1] +PVG I M+A DS P  
 301 NLD--HLTANG-----IWSRGGMWKNGE [E1] SERVPVGTIQMFAGDSAP-PG 350

359 WRF [D1] BASDCPLYASRIGTRYGGNPSNPGI [D2] FVRGSGRGSHLTNPVNGND 418  
 W [D1] P Y + +GTR+GG+ +NPG+ [D2] FVRG+G GSH+ N G D  
 351 WVL [D1] EGDOFPDYRNVVGTREGGDWNPNPGI [D2] FVRGAGTGSHILNN--RGOD 408

419 QEF [D3] VGCTGGYVGEVQIQMSY [D4] FEGFGEHDDLGA-FGNTRRSNEVGTREKGLDW 477  
 +SK [D3] VGC G +VG VQ QQMSY [D4] +GE A FG + ++GTRK DW  
 409 GY [D3] VGCDGMHVGGVQAQMSY [D4] GGEFORHEAPFGASVYQGYLGTREKYS DW 468

478 DNRSYFTNDGYEIDPESQRNSKYTLNRPELIGNETRPWNISLN [E2] 525  
 DN SYFTNDG+E+ R++ TLNR LIG ETRPWNISLN [E2]  
 469 DNASYFTNDGFELG--GHRDATGTLNREGLIGYETRPWNISLN [E2] 514

Fig. 5

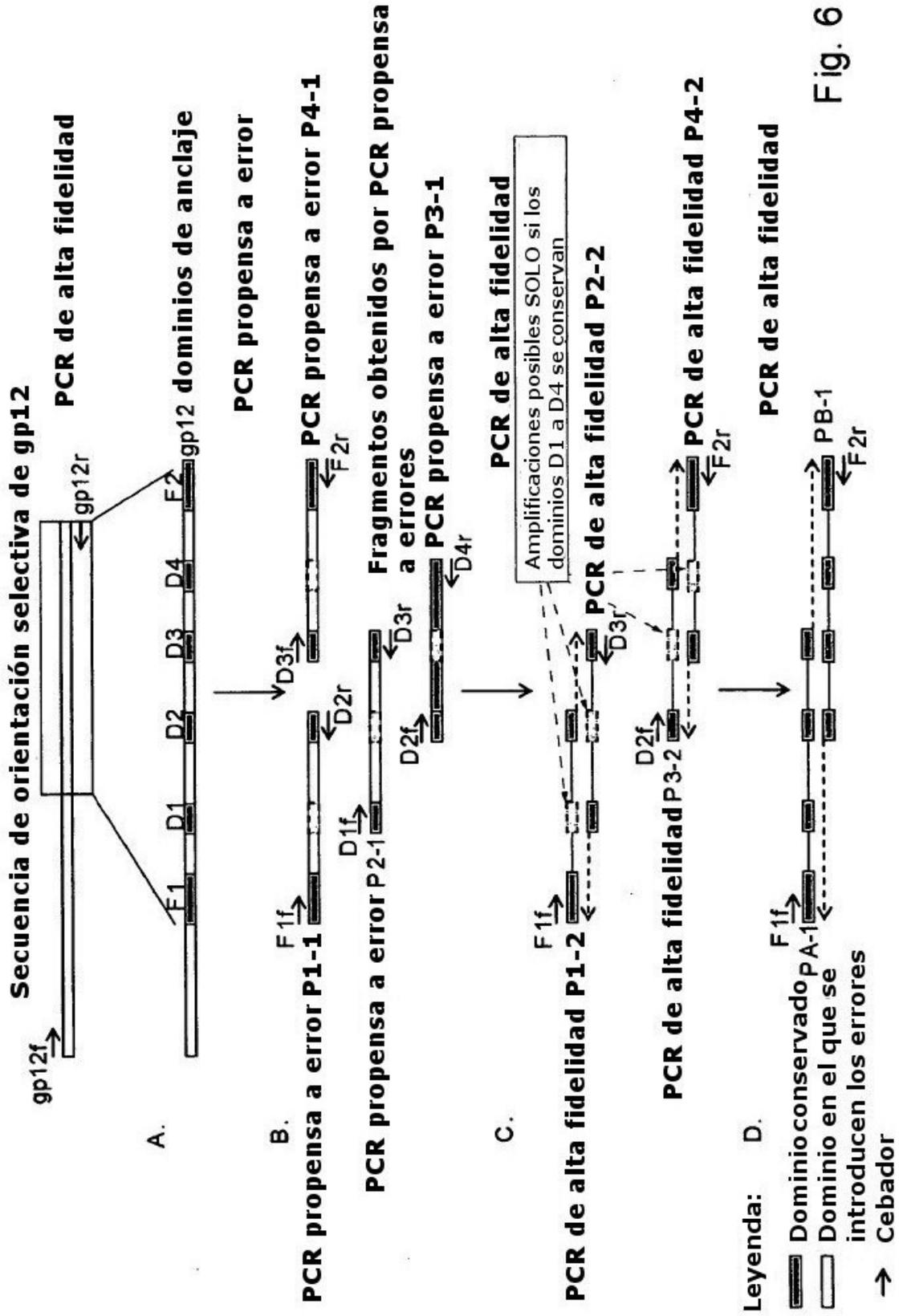


Fig. 6

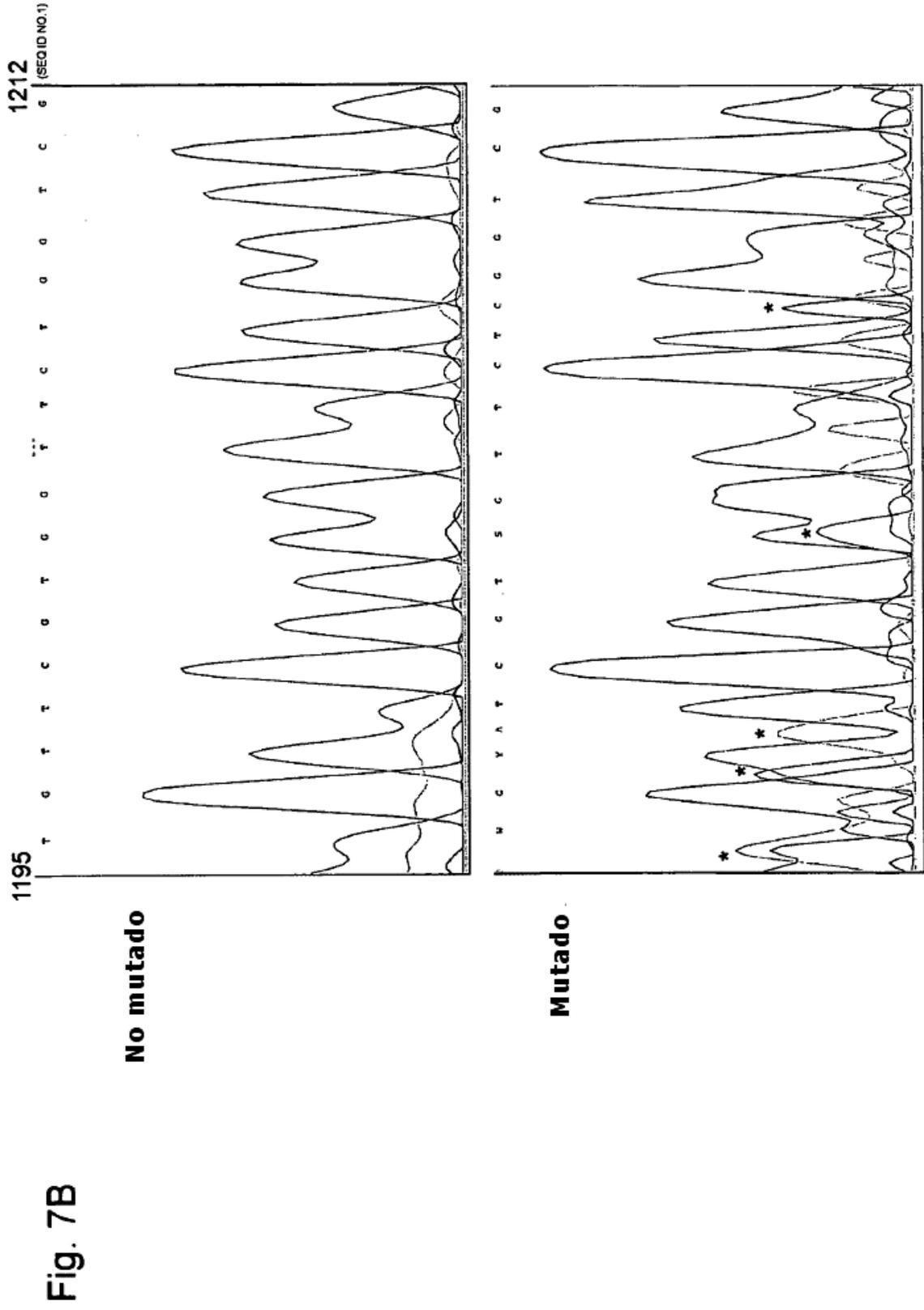


Fig. 7A

