

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 498 765**

51 Int. Cl.:

**C12N 5/00** (2006.01)

**C12N 9/78** (2006.01)

**C12N 15/00** (2006.01)

**A01K 67/027** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.04.2010 E 10723293 (6)**

97 Fecha y número de publicación de la concesión europea: **02.07.2014 EP 2414507**

54 Título: **Mutantes de la citidina desaminasa inducida por activación (AID) y procedimientos de uso**

30 Prioridad:

**03.04.2009 US 166349 P**

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

**25.09.2014**

73 Titular/es:

**MEDICAL RESEARCH COUNCIL (100.0%)  
20 Park Crescent  
London W1B 1AL, GB**

72 Inventor/es:

**WANG, MENG;  
YANG, ZIZHEN;  
RADA, CRISTINA y  
NEUBERGER, MICHAEL**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

**ES 2 498 765 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

Mutantes de la citidina desaminasa inducida por activación (AID) y procedimientos de uso

**Antecedentes de la invención**

5 Los mecanismos naturales de generación de diversificación de anticuerpos explotan el proceso de la hipermutación somática (HMS) para desencadenar la evolución de las regiones variables de inmunoglobulina, de modo que se generan rápidamente el repertorio de anticuerpos secundarios asociados con la respuesta humoral. La HMS *in vivo* representa un proceso altamente eficiente que es capaz de explorar rápidamente estructuras plegadas productivas y desarrollar anticuerpos de alta afinidad de un modo que representa el proceso natural para la optimización de anticuerpos. Por tanto, ha existido un interés significativo en intentar replicar la HMS *in vitro* para crear un simple  
10 proceso sólido que podría imitar los procesos naturales de la maduración de la afinidad directamente en un contexto de células de mamífero para seleccionar y desarrollar anticuerpos inmunogénicamente tolerados y altamente expresados en células de mamífero (Cumbers et al., Nat Biotechnol., 20(11): 1129 - 1134 (2002); Wang et al., Prot. Eng. Des. Sel., 17(9): 569 - 664 (2004); Wang et al., Proc. Natl. Acad. Sci. USA., 101 (48): 16745 - 16749 (2004); Ruckerl et al., Mol. Immunol., 43 (10): 1645 - 1652 (2006); Todo et al., J. Biosci. Bioeng., 102(5): 478 - 81 (2006);  
15 Arakawa et al., Nucleic Acids Res., 36(1): e1 (2008)).

No obstante, los anticuerpos nativos que se han aislado de un ser humano o animal individual a menudo fracasan a la hora de demostrar propiedades de afinidad óptimas porque un techo de afinidad intrínseca inherente al sistema inmunológico previene la discriminación *in vivo*, y por tanto la selección, de anticuerpos con afinidades más potentes que aproximadamente 100 pM (Batista and Neuberger, Immunity, 8(6): 751 - 91998, (1998) and EMBO J., 19(4): 513  
20 - 20 (2000).

El uso de bibliotecas de expresión en fagos puede abordar algunos de estos problemas y se ha demostrado que los abordajes basados en la expresión en fagos son capaces de producir de forma rutinaria anticuerpos de alta afinidad. No obstante, desde una perspectiva teórica, dichas bibliotecas estáticas están limitadas de forma inherente en su tamaño y alcance porque las bibliotecas más grandes ( $10^{12}$ ) únicamente pueden explorar una pequeña fracción del potencial repertorio inmunológico innato. Adicionalmente, no es posible codesarrollar de forma simultánea anticuerpos a través de abordajes de expresión en fagos en base a una buena expresión en mamíferos y a una afinidad elevada, lo que conduce a potenciales problemas de fabricación cadena abajo que sean el resultado de la por otro lado mala expresión en células huésped de mamífero. Adicionalmente, el uso de mutagénesis aleatoria en combinación con expresión en fagos carece de un perfil de selectividad inherente hallado en los procesos naturales de la maduración de la afinidad de los anticuerpos, lo que a menudo da lugar a problemas de inmunidad humana antihumana o perfiles de reactividad cruzada indeseables.

El uso de una línea celular cultivada para desarrollar un anticuerpo frente a un antígeno diana específico usando hipermutación somática *in Vitro* se demostró primero usando la línea celular de linfoma de Burkitt humano Cumbers et al., Nat. Biotechnol., 20(11): 1129 - 1134 (2002)). Ramos, y otras líneas de linfocitos B, también se han usado con éxito para desarrollar genes que no son de anticuerpos que se han integrado aleatoriamente en el ADN cromosómica de la célula huésped (Wang et al., Prot. Eng. Des. Sel., 17(9): 569 - 664 (2004) y Proc. Natl. Acad. Sci. USA., 101(48): 16745 - 16749 (2004)). Adicionalmente, se ha demostrado una hipermutación somática eficiente en genes que no son de anticuerpos en líneas de linfocitos B usando vectores episomales, con o sin elementos reguladores en cis específicos de Ig (Ruckerl et al., Mol. Immunol., 43 (10): 1645 - 1652 (2006)). Aunque algunas líneas celulares de Ramos muestran tasas relativamente altas de hipermutación constitutiva, las líneas de linfocitos B en general muestran tasas relativamente lentas de división celular y son difíciles de transfectar con una eficiencia elevada, lo que limita su utilidad práctica para una evolución dirigida.

La línea celular de linfocitos B de la bolsa de Fabricio de pollo, DT40, diversifica su gen ligero de Ig reordenado mediante conversión génica del molde del seudogen V. No obstante, si la conversión génica se bloquea mediante la delección del parólogo Rad51, XRCC2 (Sale et al., Nature, 412: 921 - 6 (2001) o la delección de los donantes de conversión del seudogen (Arakawa et al., Nucleic Acids Res., 36(1): e1 (2008)), la línea celular expresa hipermutación constitutiva en cultivo. Por comparación con las células de Ramos, las células DT40 tienen un tiempo de generación significativamente más corto (12 horas), son susceptibles a ser dianas génicas dirigidas y se han usado con éxito para la evolución dirigida tanto de anticuerpos endógenos (Seo et al., Nat. Biotechnol., 23(6): 731 - 5 (2005); Nat. Protoc., 1(3): 1502 - 6 (2006); Biotechnol. Genet. Eng. Rev., 24: 179 - 93 (2007); Todo et al., J. Biosci. Bioeng., 102(5): 478 - 81 (2006)) como de proteínas que no son anticuerpos (Arakawa et al., Nucleic Acids Res., 36(1): e1 (2008)).

Aunque los derivados de linfocitos B, tales como RAMOS y DT40, se han usado con éxito para la evolución dirigida, el uso fiable de estas células es un proceso sólido para la evolución dirigida se complica con una serie de factores, incluidos: (i) la necesidad de insertar el gen de interés en un sitio definido en el locus Ig de la célula huésped con el fin de alcanzar un nivel elevado de mutagénesis (Parsa et al., Mol Immunol., 44(4): 567 - 75 (2007), y (ii) la biología natural compleja de la hipermutación somática que actúa en los loci de inmunoglobulinas endógenas en estas células. Adicionalmente, dichas líneas celulares modificadas genéticamente exhiben una inestabilidad clonal significativa en los índices de HMS (Zhang et al., Int. Immunol., 13: 1175 - 1184 (2001), Martin et al., Proc. Natl. Acad. Sci. USA., 99(19): 12304 - 12308 (2002) and Nature, 415(6873): 802 - 806 (2002); Ruckerl et al., Mol.  
60

Immunol., 41: 1135 - 1143 (2004)), y no proporcionan ningún medio simple para regular o controlar la hipermutación, es decir pasar a mutagénesis después de la selección de un fenotipo deseado.

El uso de linfocitos no B para iniciar la hipermutación somática dirigida en un gen de interés se ha descrito con éxito en una serie de grupos (Martin et al., Proc. Natl. Acad. Sci. USA., 99(19): 12304 - 12308 (2002) and Nature, 415(6873): 802 - 806 (2002); McBride et al., Proc. Natl. Acad. Sci. USA, 103(23): 8798 - 803 (2006); Jovanic et al., PLoS ONE, 23;3(1): e1480 (2008); publicación de la solicitud de patente de EE.UU. 09/0075378; publicaciones solicitud de patente internacional WO 08/103474A1 y WO 08/103475A1), y estas líneas celulares también proporcionan una transferencia génica eficiente, nivel elevado de expresión proteica, características de crecimiento óptimas y son fácilmente sensibles al cultivo en suspensión y citometría de flujo.

La citidina desaminasa inducida por activación (AID) pertenece a la familia APOBEC de las enzimas citidina desaminasas. La AID se expresa en los linfocitos B activados y es necesaria para iniciar la hipermutación somática (Muramatsu et al., Cell, 102(5): 553 - 63 (2000); Revy et al., Cell, 102(5): 565 - 75 (2000); Yoshikawa et al., Science, 296(5575): 2033 - 6 (2002)) creando mutaciones puntuales en el ADN subyacente que codifica los genes de anticuerpos (Martin et al., Proc. Natl. Acad. Sci. USA., 99(19): 12304 - 12308 (2002) y Nature, 415(6873): 802 - 806 (2002); Petersen-Mart et al., Nature, 418(6893): 99 - 103 (2002)). La AID también es un factor proteico esencial para la recombinación por cambio de clase y la conversión génica (Muramatsu et al., Cell, 102(5): 553 - 63 (2000); Revy et al., Cell, 102(5): 565 - 75 (2000)).

El descubrimiento de que la AID es responsable de iniciar la hipermutación somática ha abierto la posibilidad de usar líneas celulares no linfocitos B para crear sistemas más definidos, estables y controlables para usar la hipermutación somática.

A pesar de estos avances, todavía existen los principales retos en lo que respecta al desarrollo de un sistema práctico para la hipermutación somática, incluyendo (1) la capacidad para dirigir la hipermutación somática a un gen de interés y lejos de los genes estructurales, (2) las tasas relativamente bajas y la naturaleza de las mutaciones alcanzadas usando AID exógena en comparación con la hipermutación somática *in vivo* y (3) los tiempos de duplicación celular relativamente largos necesarios para cultivar una población celular a partir de un único clon celular entre ciclos de mutagénesis.

Por tanto, existe una necesidad específica de mejores composiciones y procedimientos para mejorar la eficiencia de los sistemas de hipermutación somática. La presente invención proporciona dichas composiciones y procedimientos.

### **Breve resumen de la invención**

La invención proporciona una molécula de ácido nucleico aislada y purificada que comprende una secuencia de nucleótidos que codifica una citidina desaminasa inducida por activación (AID) mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos.

En una realización, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en el residuo 34, el residuo 82 y el residuo 156.

En otra realización, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en el residuo 10 y al menos una sustitución de aminoácidos en el residuo 156.

En otra realización más, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en el residuo 35 y al menos una sustitución de aminoácidos en el residuo 145.

En una realización adicional, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en el residuo 34 y al menos una sustitución de aminoácidos en el residuo 160.

En otra realización, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en el residuo 43 y al menos una sustitución de aminoácidos en el residuo 120.

La invención también proporciona una molécula de ácido nucleico aislada y purificada que comprende una secuencia de nucleótidos que codifica una proteína (AID) mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos dos sustituciones de aminoácidos, en la que al menos una sustitución está en el residuo 57 y al menos una sustitución está en el residuo 145 o 81 y en la que la proteína AID mutante funcional tiene al menos una mejora de 10 veces la actividad en comparación con la proteína AID humana en un ensayo de papilación bacteriana.

En todavía otra realización, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la

secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 82.

5 En una realización adicional, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 34.

En una realización adicional, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 157.

10 En otra realización más, la secuencia de aminoácidos de la proteína AID mutante funcional difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en los residuos 10, 82 y 156.

15 La invención también proporciona una molécula de ácido nucleico aislada y purificada que comprende una secuencia de nucleótidos que codifica una citidina desaminasa inducida por activación (AID) mutante funcional cuya secuencia de aminoácidos difiere de una secuencia de aminoácidos seleccionada del grupo de una proteína AID canina (SEC ID N° 3), proteína AID murina (SEC ID N° 4), proteína AID de rata (SEC ID N° 5), proteína AID bovina (SEC ID N° 6), proteína AID de pollo (SEC ID N° 7) en al menos una sustitución de aminoácidos, en la que un residuo seleccionado del grupo que consiste en el residuo 34, el residuo 82 y el residuo 156, en la que la proteína AID mutante funcional tiene al menos una mejora de 10 veces la actividad en comparación con la proteína AID humana en un ensayo de papilación bacteriana.

20 También se proporciona un vector de expresión que comprende una molécula de ácido nucleico que comprende una secuencia de nucleótidos que codifica una citidina desaminasa inducida por activación (AID) mutante funcional.

La invención también proporciona una célula aislada que comprende la molécula de ácido nucleico que codifica una proteína AID mutante funcional.

25 También se proporciona un animal transgénico que comprende la molécula de ácido nucleico que codifica una proteína AID mutante funcional.

30 La invención también proporciona un procedimiento de preparación de un producto génico que tiene una propiedad deseada, en el que el procedimiento comprende expresar un ácido nucleico que codifica el producto génico en una población de células, en la que la población de células expresa, o se puede inducir que exprese, una proteína AID mutante funcional, con lo cual la expresión de la proteína AID mutante funcional induce una mutación en el ácido nucleico que codifica el producto génico.

También se proporciona un procedimiento para mutar un organismo que tienen un fenotipo deseado que comprende expresar o inducir la expresión de una proteína AID mutante funcional en el organismo, con lo cual la expresión de la proteína AID mutante funcional induce una mutación dentro del ADN cromosómico del organismo.

### **Breve descripción de las figuras**

35 La FIG. 1a incluye imágenes de papilas en colonias bacterianas que expresan o no expresan AID humana. La FIG. 1b es un gráfico de barras que cuantifica las papilas en las colonias bacterianas que expresan AID humana, APOBEC1 (A1) o APOBEC3G (A3G). La FIG. 1c es un diagrama que representa los dos ADNc de APOBEC3G obtenidos mediante detección selectiva en una biblioteca de ADNc de bazo humano de papilas. La FIG. 1d incluye imágenes de papilas en colonias bacterianas que expresan la proteína AID indicada y enumera las frecuencias de mutación respecto del vector. La FIG. 1e incluye imágenes de papilación mediante AID Mut 1, 1 como función de la concentración de arabinosa. La FIG. 1f es un gráfico que representa la eficiencia del cultivo en placas de mutantes de AID humana.

45 La FIG. 2 es un gráfico que ilustra determinados mutantes de AID humana funcionales identificados en las detecciones de papilación. Los números representan la frecuencia media de mutación en Rif<sup>r</sup> de cada mutante AID humana respecto del vector.

50 La FIG. 3a es un diagrama de secuencia de AID humana que ilustra las posiciones e identidades de las mutaciones funcionales que se identificaron en la AID humana y de pez globo. La FIGS. 3b compara el nivel de expresión de las proteínas de fusión GST-AID mutantes mediante transferencia de tipo Western. Las FIGS. 3c,d son gráficos que cuantifican la actividad desaminasa y dirigen la especificidad de las proteínas de fusión GST-AID mutantes.

La FIG. 4a es un gráfico que ilustra determinadas mutantes de AID de pez globo identificadas en las detecciones de papilación. La FIG. 4b es un gráfico de barras que compara la frecuencia de mutación relativa de las mutantes de AID de pez globo con respecto a Rif<sup>r</sup> a 18° C y 37° C.

La FIG. 5a incluye gráficos de citometría de flujo de la expresión de IgM y GFP en clones de DT40 individuales

que expresan la proteína AID indicada. La FIG. 5a también contiene un gráfico de la pérdida de IgM en 12 transfectantes clonales independientes que expresan la proteína indicada. La FIG. 5b contiene un diagrama de la distribución de las mutaciones IgVλ observadas en células DT40 transfectadas. La FIG. 5b también contiene gráficos circulares que representan el número de mutaciones IgVλ después de clasificar según la pérdida de IgM. La FIG. 5b también muestra la expresión de AID mediante transferencia de tipo Western. La FIG. 5c incluye gráficos de citometría de flujo del cambio a IgG1 en linfocitos B deficientes en AID transducidos con los retrovirus indicados. El gráfico de barras en la FIG. 5c cuantifica el cambio de IgG1 respecto a la AID de tipo silvestre y la transferencia de tipo Western en la FIG. 5c muestra la expresión de AID mediante transferencia de tipo Western.

La FIG. 6a es un diagrama que ilustra una translocación recíproca entre c-myc y el locus de IgH y representa los cebadores (flechas) y sondas (P) usados para la detección de la translocación. La FIG. 6b es una transferencia de tipo Southern de translocaciones c-myc-IgH derivadas de los cromosomas 15 y 12 tras la amplificación mediante PCR del ADN genómico de linfocitos C deficientes en AID transducidos con el retrovirus indicado.

La FIG. 7 es una alineación LOGO que ilustra que las mutaciones funcionales identificadas en la criba de papilación bacteriana acercan la secuencia de AID a la de APOBEC3s.

La FIG. 8 enumera los números de acceso en GenBank/Ensembl de las secuencias de AID y APOBEC3 de mamífero usadas para generar la FIG. 7.

La FIG. 9 es un diagrama de secuencia de AID humana y AID de pez globo (fugu) que ilustra las posiciones e identidades de las mutaciones funcionales identificadas.

La FIG. 10a es una alineación de la secuencia de ácido nucleico de secuencias AID usadas en los experimentos celulares 293-c18 descritos en el Ejemplo 14. Los residuos en recuadros indican cambios entre las secuencias mutantes 7,3 y silvestre. La FIG. 10b es una alineación de la secuencia de aminoácidos de las secuencias AID usadas en los experimentos celulares 293-c18 descritos en el Ejemplo 14. Los residuos en recuadros indican cambios entre las secuencias mutantes 7,3 y silvestre. La mutación de L en A en MutE and Mut 7.3 desactiva la función de la señal de exportación nuclear. Un punto indica un codón de terminación y un guión especifica las posiciones en las que no hay aminoácido correspondiente.

### **Descripción detallada de la invención**

La invención proporciona una molécula de ácido nucleico aislada y purificada que comprende una secuencia de nucleótidos que codifica una proteína (AID) mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos, en la que la proteína AID mutante funcional tiene al menos una mejora de 10 veces la actividad en comparación con la proteína AID humana en un ensayo de papilación bacteriana.

Con “molécula de ácido nucleico” se pretende abarcar un polímero de ADN o ARN, es decir un polinucleótido que puede ser monocatenario o bicatenario y que puede contener nucleótidos no naturales o alterados. Los términos “ácido nucleico” y “polinucleótido”, como se usan en el presente documento, se refieren a una forma de nucleótidos de cualquier longitud, ribonucleótidos (ARN) o desoxirribonucleótidos (ADN). Estos términos hacen referencia a la estructura primaria de la molécula y, por tanto, incluyen ADN monocatenario y bicatenario y ARN monocatenario y bicatenario. Los términos incluyen, como equivalentes, análogos de ARN o ADN formados por análogos nucleotídicos y polinucleótidos modificados tales como, entre otros, polinucleótidos metilados y/o protegidos.

El término “nucleótido”, como se usa en el presente documento, se refiere a una unidad monomérica de un polinucleótido que consiste en una base heterocíclica, un azúcar y uno o más grupos fosfato. Las bases de origen natural (guanina, (G), adenina, (A), citosina, (C), timina, (T) y uracilo (U)) normalmente son derivados de purina o pirimidina, aunque debe entenderse que también se incluyen análogos de bases de origen natural o no naturales. El azúcar de origen natural es la pentosa (azúcar de cinco carbonos) desoxirribosa (que forma ADN) o ribosa (que forma ARN), aunque se debe entender que también están incluidos los análogos de azúcar de origen natural y no natural. Normalmente, los ácidos nucleicos están unidos mediante enlaces fosfatos para formar ácidos nucleicos o polinucleótidos, aunque en la materia se conocen muchos otros enlaces (p. ej., fosforotioatos, boranofosfatos y similares).

Las expresiones “polinucleótido sintético”, “gen sintético” o “polipéptido sintético”, como se usan en el presente documento, significan que la correspondiente secuencia de polinucleótidos o porción de la misma, o secuencia de aminoácidos o porción de la misma, deriva de una secuencia que se ha diseñado o sintetizado *de novo* o modificado en comparación con la secuencia de origen natural equivalente. Los polinucleótidos sintéticos o genes sintéticos se pueden preparar mediante procedimientos conocidos en la materia, incluyendo, entre otros, síntesis química de secuencias de ácido nucleico o aminoácidos o amplificar mediante PCR (o sistemas de amplificación enzimática similares). Los genes sintéticos normalmente son diferentes de los genes no modificados o genes de origen natural, bien a nivel de aminoácidos o a nivel de polinucleótidos (o ambos) y normalmente se localizan dentro del contexto de las secuencias control de la expresión sintética. Por ejemplo, las secuencias génicas sintéticas pueden incluir secuencias de aminoácidos o polinucleótidos que se han modificado, por ejemplo mediante sustitución, delección o adición de uno o más aminoácidos o nucleótidos, de modo que se proporciona una secuencia de aminoácidos o una

secuencia de codificación polinucleotídica que es diferente de la secuencia fuente. Los genes sintéticos o secuencias de polinucleótidos no necesariamente codifican proteínas con diferentes aminoácidos frente al gen natural. Por ejemplo, también pueden abarcar secuencias de polinucleótidos sintéticas que incorporan diferentes codones pero que codifican el mismo aminoácido, es decir los cambios nucleotídicos representan mutaciones silentes a nivel de aminoácidos. En una realización, los genes sintéticos exhiben una susceptibilidad alterada a la HMS en comparación con el gen de origen natural o no modificado. Los genes sintéticos se pueden modificar repetidamente usando los procedimientos descritos en el presente documento, en cada repetición sucesiva, una correspondiente secuencia de polinucleótidos o secuencia de aminoácidos, deriva, completa o parcialmente, a partir de una secuencia que se ha diseñado o sintetizado *de novo* o modificado en comparación con la secuencia no modificada equivalente.

Como se usa en el presente documento, un "codón" se refiere a los tres nucleótidos que, cuando se transcriben y traducen, codifican un único residuo de aminoácidos o en el caso de UUA, UGA o UAG codifica una señal de terminación. Los codones que codifican aminoácidos son bien conocidos en la técnica.

El uso óptimo de codones está indicado por las frecuencias de uso de codones para genes expresados, por ejemplo, como se muestra en el gráfico de uso de codones del programa "Human - High.cod" del Wisconsin Sequence Analysis Package, Versión 8,1, Genetics Computer Group, Madison, Wisc. El uso de codones también se describe en, por ejemplo, R. Nussinov, "Eukaryotic Dinucleotide Preference Rules and Their Implications for Degenerate Codon Usage," J. Mol. Biol., 149: 125 - 131 (1981). Los codones que se usan con más frecuencia en los genes humanos altamente expresados son, posiblemente, los codones óptimos para la expresión en células huésped humanas y, por tanto, forman la base para construir una secuencia de codificación sintética.

Por "aislado" se quiere decir la eliminación de un ácido nucleico de su ambiente natural. Por "purificado" se quiere decir que en un ácido nucleico dado, se haya eliminado uno de la naturaleza (incluyendo ADN y ARNm genómico) o sintetizado (incluyendo ADNc) y/o amplificado en condiciones de laboratorio, se ha aumentado la pureza, en la que "pureza" es un término relativo, no "pureza absoluta". No obstante, debe entenderse que los ácidos nucleicos y las proteínas se pueden formular con diluyentes o adyuvantes y aislarse con fines prácticos. Por ejemplo, los ácidos nucleicos se mezclarán con un vehículo o diluyente aceptable cuando se usa para introducir en las células.

La expresión "citidina desaminasa inducida por activación" o ("AID") se refiere a miembros de la familia AID/APOBEC de ARN/ADN que editan citidina desaminasas capaces de mediar en la desaminación de citosina en uracilo dentro de una secuencia de ADN. (Véase, Conticello et al., Mol. Biol. Evol., 22: 367 - 377 (2005) y la patente de EE.UU. 6.815.194).

La expresión "AID de tipo silvestre" se refiere a una secuencia de aminoácidos de origen natural de una proteína AID. Las proteínas AID de tipo silvestre adecuadas incluyen todas las formas vertebradas de AID, incluyendo, por ejemplo, primates, roedores, aves y osteictios. Ejemplos representativos de secuencias de aminoácidos de AID de tipo silvestre incluyen, sin limitaciones, AID humana (SEC ID N° 1 o SEC ID N° 2), AID canina (SEC ID N° 3), AID murina (SEC ID N° 4), AID de rata (SEC ID N° 5), AID bovina (SEC ID N° 6), AID de pollo (SEC ID N° 7), AID porcina (SEC ID N° 8), AID de chimpancé (SEC ID N° 9), AID de macaco (SEC ID N° 10), AID de caballo (SEC ID N° 11), AID de Xenopus (SEC ID N° 12), AID de pez globo (fugu) (SEC ID N° 13) y de pez cebra (SEC ID N° 14).

El término "homólogo de AID" se refiere a las enzimas de la familia Apobec e incluyen, por ejemplo, Apobec-1, Apobec3C o Apobec3G (descrito en, por ejemplo, Jarmuz et al., Genomics, 79: 285 - 296 (2002)). La expresión "actividad AID" incluye la actividad mediada por AID y homólogos de AID.

Un "mutante de AID" como se usa en el presente documento se refiere a una secuencia de aminoácidos de AID que difiere de una secuencia de aminoácidos de AID de tipo silvestre en al menos un aminoácido. Una secuencia de aminoácidos de tipo silvestre se puede mutar para producir un mutante de AID mediante cualquier procedimiento conocido en la técnica, tal como, por ejemplo, mediante inserción, delección y/o sustitución. Por ejemplo, se pueden introducir mutaciones en una secuencia de ácido nucleico que codifica AID de tipo silvestre aleatoriamente o de un modo específico de un sitio. Las mutaciones aleatorias se pueden generar mediante, por ejemplo, PCR propensa a error de una secuencia molde de AID. Un medio preferido para introducir mutaciones aleatorias es el kit de mutagénesis aleatoria Genemorph II (Stratagene, LaJolla, CA). Las mutaciones específicas de sitio se pueden introducir, por ejemplo, ligando un vector de expresión un oligonucleótido sintetizado que comprende el sitio modificado. Como alternativa, se pueden usar procedimientos de mutagénesis específica de sitio dirigida a oligonucleótido, tales como los divulgados en Walder et al., Gene, 42: 133 (1986); Bauer et al., Gene, 37: 73 (1985); Craik, Biotechniques, 12 - 19 (January 1995); y las patentes de EE.UU. N° 4.518.584 y 4.737.462. Un medio preferido para introducir mutaciones específicas de sitio es el kit de mutagénesis dirigida a sitios QuikChange (Stratagene, LaJolla, CA).

Las expresiones "mutante de AID funcional", "mutante AID funcional" o "proteína AID mutante funcional" se refieren cada una a una proteína AID mutante que conserva todo o parte de la actividad biológica de una AID de tipo silvestre o que exhibe una mayor actividad biológica en comparación con una proteína AID silvestre". La actividad biológica de una AID de tipo silvestre incluye, entre otros, la desaminación de citosina en uracilo dentro de una secuencia de ADN, papilación en un ensayo de mutagénesis bacteriana, hipermutación somática de un gen diana y cambio de clase de inmunoglobulina. Una proteína AID mutante puede conservar cualquier parte de la actividad

biológica de una proteína AID de tipo silvestre. Deseablemente, la proteína AID mutante conserva al menos un 75% (p. ej., 75 %, 80 %, 90 % o más) de la actividad biológica de la AID de tipo silvestre. Preferentemente, la proteína AID mutante conserva al menos un 90 % (p. ej., 90 %, 95 %, 100 % o más) de la actividad biológica de la AID de tipo silvestre.

5 En una realización preferida, la proteína AID mutante exhibe un incremento de la actividad biológica en comparación con una proteína AID de tipo silvestre. A este respecto, la AID mutante funcional tiene al menos una mejora de 10 veces la actividad en comparación con la proteína AID de tipo silvestre medida mediante un ensayo de papilación bacteriana. En la materia se sabe que los ensayos de papilación bacteriana son útiles para detectar mutantes de *E. Coli* que son defectuosos en algún aspecto de la reparación del ADN (Nghiem et al., Proc. Natl. Acad. Sci. USA, 85: 2709 - 2713 (1988) y Ruiz et al., J. Bacteriol., 175: 4985 - 4989 (1993)). El ensayo de papilación bacteriana puede usar células CC102 de *Escherichia coli* que alojan una mutación de sentido erróneo dentro del gen *lacZ*, las células CC102 de *E. Coli* dan lugar a colonias blancas en placas con lactosa-MacConkey. En estas colonias blancas a menudo se puede distinguir un número pequeño de microcolonias rojas o "papilas" (normalmente 0 - 2 por colonia), que refleja inversores Lac<sup>+</sup> que aparecen de forma espontánea. Los clones bacterianos que exhiben una frecuencia elevada de mutación espontánea (es decir, "clones mutadores") se pueden identificar en virtud de un mayor número de papilas. Los ensayos de papilación bacteriana se pueden usar para buscar mutantes de AID funcionales que tienen una mayor actividad en comparación con la AID de tipo silvestre, Los ensayos de papilación bacteriana se describen con detalle en los Ejemplos.

20 En una realización, la AID mutante funcional tiene al menos una mejora de 10 (p. ej., 10, 30, 50 o más) veces la actividad en comparación con la proteína AID de tipo silvestre en un ensayo de papilación bacteriana. Preferentemente, la AID mutante funcional tiene al menos una mejora de 100 (p. ej., 100, 200, 300 o más) veces la actividad en comparación con la AID de tipo silvestre. Más preferentemente, la AID mutante funcional tiene al menos una mejora de 400 (p. ej., 400, 500, 1000 o más) veces la actividad en comparación con la AID de tipo silvestre.

25 La proteína AID mutante funcional comprende una secuencia de aminoácidos que difiere de la secuencia de aminoácidos de una proteína AID silvestre en al menos una sustitución de aminoácido. La proteína AID de tipo silvestre puede ser cualquier proteína AID de vertebrados, incluidos los descritos en el presente documento. Deseablemente, la proteína AID de tipo silvestre es una proteína AID humana, de la cual existen al menos dos variantes conocidas (es decir, SEC ID N° 1 y la SEC ID N° 2). Proteínas AID adicionales de vertebrados incluyen sin limitaciones, AID canina (SEC ID N° 3), AID murina (SEC ID N° 4), AID de rata (SEC ID N° 5), AID bovina (SEC ID N° 6), AID de pollo (SEC ID N° 7), AID porcina (SEC ID N° 8), AID de chimpancé (SEC ID N° 9), AID de macaco (SEC ID N° 10), AID de caballo (SEC ID N° 11), AID de *Xenopus* (SEC ID N° 12), AID de pez globo (fugu) (SEC ID N° 13) (y de pez cebra (SEC ID N° 14).

35 Un experto en la técnica apreciará que aunque existe un grado elevado de homología entre las proteínas AID de vertebrados, existe un número variable de sustituciones, deleciones e inserciones de aminoácidos en cada una de las proteínas AID de vertebrados respecto a la AID humana (SEC ID N° 1 o SEC ID N° 2). Como tal, la presente invención abarca las mutaciones descritas en el presente documento cuando se incorporan en la posición análoga de cualquier proteína AID de vertebrados. Un experto en la técnica puede determinar la posición análoga en cualquier proteína AID de vertebrados realizando una alineación de secuencia de la proteína AID de vertebrados homóloga con la de la AID humana (SEC ID N° 1 o SEC ID N° 2) usando cualquier programa de alineación basado en ordenador conocido en la técnica (p. ej., BLAST o ClustalW2).

45 Normalmente, una proteína AID de tipo silvestre contiene una secuencia de exportación nuclear cerca del extremo C de la proteína. En una realización de la invención se puede mutar un residuo o una pluralidad de residuos que median en la exportación nuclear de la AID de tipo silvestre y se puede generar una proteína AID mutante funcional que comprende una secuencia de aminoácidos que difiere de la secuencia de aminoácidos de una proteína AID que tiene una secuencia de exportación nuclear mutada en al menos una sustitución de aminoácido adicional. Ejemplos de proteínas AID caninas que tiene una secuencia de exportación nuclear mutada que puede servir como secuencia de referencia en la que se puede insertar una o más mutaciones identificadas en el presente documento como productoras de una mutante de AID funcional incluyen un mutante L198A (SEC ID N° 70) y un mutante D187E, D188E, D191E, T195I y L198A (SEC ID N° 71).

50 Una "sustitución" de aminoácidos se refiere a la sustitución de un aminoácido en una posición o residuo dado por otro aminoácido en la misma posición o residuo dentro de una secuencia de polipéptidos.

Los aminoácidos se agrupan ampliamente en "aromáticos" o "alifáticos". Un aminoácido aromático incluye un anillo aromático. Ejemplos de aminoácidos "aromáticos" incluyen histidina (H o His), fenilalanina (F o Phe), tirosina (Y o Tyr) y triptófano (W o Trp). Los aminoácidos no aromáticos se agrupan ampliamente en "alifáticos". Ejemplos de aminoácidos "alifáticos" incluyen glicina (G o Gly), alanina (A o Ala), valina (V o Val), leucina (L o Leu), isoleucina (I o Ile), metionina (M o Met), serina (S o Ser), treonina (T o Thr), cisteína (C o Cys), prolina (P o Pro), ácido glutámico (E o Glu), ácido aspártico (A o Asp), asparagina (N o Asn), glutamina (Q o Gln), lisina (K o Lys) y arginina (R o Arg).

60 Los aminoácidos alifáticos se pueden subdividir en cuatro subgrupos. El "subgrupo no polar alifático grande" consiste en valina, leucina e isoleucina, el "subgrupo alifático ligeramente polar" consiste en metionina, serina, treonina y cisteína, el "subgrupo alifático polar/cargado" consiste en ácido glutámico, ácido aspártico, asparagina,

glutamina, lisina y arginina, y el “subgrupo de residuos pequeños” consiste en glicina y alanina. El grupo de aminoácidos cargados/polares se puede subdividir en tres subgrupos: el “subgrupo cargado positivamente” que consiste en lisina y arginina, el “subgrupo cargado negativamente” que consiste en ácido glutámico y ácido aspártico y el “subgrupo polar” que consiste en asparagina y glutamina.

- 5 Los aminoácidos aromáticos se pueden subdividir en dos subgrupos: el “subgrupo de anillo de nitrógeno” que consiste en histidina y triptófano, y el “subgrupo de fenilo” que consiste en fenilalanina y tirosina.

La expresión “sustitución de aminoácidos conservadora” o “mutación conservadora” hace referencia a la sustitución de un aminoácido por otro aminoácido con una propiedad común. Un modo funcional de definir propiedades comunes entre aminoácidos individuales es analizar las frecuencias normalizadas de los cambios de aminoácidos entre las correspondientes proteínas de organismos homólogos (Schulz, G. E. and R. H. Schirmer, Principles of Protein Structure, Springer-Verlag, New York (1979)). De acuerdo con estos análisis, los grupos de aminoácidos se pueden definir cuando los aminoácidos dentro de un grupo se intercambian preferentemente entre sí y, por tanto, se asemejan entre sí principalmente en su impacto sobre la estructura global de la proteína (Schulz, G. E. y R. H. Schirmer, citado anteriormente).

- 15 Ejemplos de mutaciones conservadoras incluyen sustituciones de aminoácidos dentro de los subgrupos anteriores, por ejemplo lisina por arginina y viceversa, de modo que se pueda mantener una carga positiva: ácido glutámico por ácido aspártico y viceversa de modo que se pueda mantener una carga negativa; serina por treonina de modo que se pueda mantener un –OH libre; y glutamina por asparagina de modo que se pueda mantener un -NH<sub>2</sub> libre.

Las “mutaciones semiconservadoras” incluyen sustituciones de aminoácidos por los mismos grupos indicados anteriormente, que no comparten el mismo subgrupo. Por ejemplo, la mutación de ácido aspártico por asparagina o asparagina por lisina implica cada una aminoácidos dentro del mismo grupo pero en diferentes subgrupos.

“Mutaciones no conservadoras” implican sustituciones de aminoácidos entre grupos diferentes, por ejemplo lisina por triptófano o fenilalanina por serina, etc.

En una realización preferida, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de AID de tipo silvestre en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en el residuo 34, residuo 82 y residuo 156. Estos residuos se pueden sustituir solos o en cualquier combinación. En las realizaciones en las que el residuo 34 lisina (K) está sustituido, preferentemente está sustituido por un ácido glutámico (E) o un residuo de ácido aspártico (D). En las realizaciones en las que el residuo 82 treonina (T) está sustituido, preferentemente está sustituido por una isoleucina (I) o un residuo de leucina (L). En las realizaciones en las que el residuo 156 ácido glutámico (E) está sustituido, preferentemente está sustituido por un residuo de glicina (G) o un residuo de alanina (L). Además, cuando el residuo de aminoácido 156 está sustituido (Solo o en combinación con una sustitución en el residuo 34 y/o el residuo 82), también puede ser deseable generar una proteína mutante AID funcional como sustituciones de aminoácidos en los residuos 9, 13, 38, 42, 96, 115, 132, 157, 180, 181, 183, 197, 198, o combinaciones de los mismos. En particular, (a) la sustitución de aminoácidos en el residuo 9 puede ser metionina (M) o lisina (K), (b) la sustitución de aminoácidos en el residuo 13 puede ser fenilalanina (F) o triptófano (W), (c) la sustitución de aminoácidos en el residuo 38 puede ser glicina (G) o alanina (A), (d) la sustitución de aminoácidos en el residuo 42 puede ser isoleucina (I) o leucina (L), (e) la sustitución de aminoácidos en el residuo 96 puede ser glicina (G) o alanina (A), (f) la sustitución de aminoácidos en el residuo 115 puede ser tirosina (Y) o triptófano (W), (g) la sustitución de aminoácidos en el residuo 132 puede ser ácido glutámico (E) o ácido aspártico (D), (h) la sustitución de aminoácidos en el residuo 180 puede ser isoleucina (I) o alanina (A), (i) la sustitución de aminoácidos en el residuo 181 puede ser metionina (M) o valina (V), (j) la sustitución de aminoácidos en el residuo 183 puede ser isoleucina (I) o prolina (P), (k) la sustitución de aminoácidos en el residuo 197 puede ser arginina (R) o lisina (K), (l) la sustitución de aminoácidos en el residuo 198 puede ser valina (V) o leucina (L), y (m) la sustitución de aminoácidos en el residuo 157 puede ser treonina (T) o lisina (K).

En otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 10 y al menos una sustitución de aminoácidos en el residuo 156. Estos residuos se pueden sustituir solos o en cualquier combinación. En las realizaciones en las que el residuo de aminoácido 10 (lisina) está sustituido, preferentemente está sustituido por un ácido glutámico (E) o un residuo de ácido aspártico (D). En las realizaciones en las que el residuo 156 (ácido glutámico) está sustituido, preferentemente está sustituido por un residuo de glicina (G) o un residuo de alanina (L). En las realizaciones en las que los aminoácidos en los residuos 10 y 156 están sustituidos, puede ser también deseable incluir sustituciones de aminoácidos en los residuos 13, 34, 82, 95, 115, 120, 134, 145, o combinaciones de los mismos. En particular, (a) la sustitución de aminoácidos en el residuo 13 puede ser fenilalanina (F) o triptófano (W), (b) la sustitución de aminoácidos en el residuo 34 puede ser ácido glutámico (E) o ácido aspártico (D), (c) la sustitución de aminoácidos en el residuo 82 puede ser isoleucina (I) o leucina (L), (d) la sustitución de aminoácidos en el residuo 95 puede ser serina (S) o leucina (L), (e) la sustitución de aminoácidos en el residuo 115 puede ser tirosina (Y) o triptófano (W), (f) la sustitución de aminoácidos en el residuo 120 puede ser arginina (R) o asparagina (N), y (g) la sustitución de aminoácidos en el residuo 145 puede ser leucina (L) o isoleucina (I).

En otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 35 y al menos una sustitución de aminoácidos en el residuo 145. Los residuos de aminoácidos 35 y 145 pueden estar sustituidos por cualquier aminoácido adecuado. El aminoácido en el residuo 35 está sustituido, preferentemente, por glicina (G) o alanina (A). El aminoácido en el residuo 145 está sustituido, preferentemente, por leucina (L) o isoleucina (I).

En otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 34 y al menos una sustitución de aminoácidos en el residuo 160. Los residuos de aminoácidos 34 y 160 pueden estar sustituidos por cualquier aminoácido adecuado. El aminoácido en el residuo 34 está sustituido, preferentemente, por ácido glutámico (E) o ácido aspártico (D). El aminoácido en el residuo 160 está sustituido, preferentemente, por ácido glutámico (E) o ácido aspártico (D).

En otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 43 y al menos una sustitución de aminoácidos en el residuo 120. Los residuos de aminoácidos 43 y 120 pueden estar sustituidos por cualquier aminoácido adecuado. El aminoácido en el residuo 43 está sustituido, preferentemente, por prolina (P). El aminoácido en el residuo 120 está sustituido, preferentemente, por arginina (R).

En otra realización más, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos dos sustituciones de aminoácidos, en las que al menos una sustitución está en el residuo 57 y al menos una sustitución en un residuo 145 u 81. Estos residuos pueden estar sustituidos solos o en cualquier combinación (p. ej., la sustitución de los residuos 57 y 145 o la sustitución de los residuos 57 y 81). Preferentemente, el aminoácido en el residuo 57 está sustituido, preferentemente, por glicina (G) o alanina (A). Cuando el aminoácido en el residuo 145 está sustituido, preferentemente, por leucina (L) o isoleucina (I). Cuando el aminoácido en el residuo 81 está sustituido, preferentemente, por tirosina (Y) o triptófano (W).

En todavía otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 82. Los residuos de aminoácidos 156 y 82 pueden estar sustituidos por cualquier aminoácido adecuado. El aminoácido en el residuo 156 está sustituido, preferentemente, por glicina (G) o alanina (A). El aminoácido en el residuo 82 está sustituido, preferentemente, por leucina (L) o isoleucina (I).

En otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 34. Los residuos de aminoácidos 156 y 34 pueden estar sustituidos por cualquier aminoácido adecuado. El aminoácido en el residuo 156 está sustituido por glicina (G) o alanina (A). El aminoácido en el residuo 34 está sustituido, preferentemente, por ácido glutámico (E) o ácido aspártico (D).

En otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 157. Los residuos de aminoácidos 156 y 157 pueden estar sustituidos por cualquier aminoácido adecuado. El aminoácido en el residuo 156 está sustituido, preferentemente, por glicina (G) o alanina (A). El aminoácido en el residuo 120 está sustituido, preferentemente, por arginina (R) o asparagina (N).

En todavía otra realización, la molécula de ácido nucleico codifica un mutante de AID funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de la AID de tipo silvestre en al menos una sustitución de aminoácidos en el residuo 10 y al menos una sustitución de aminoácidos en los residuos 10, 82 y 156. Estos residuos se pueden sustituir solos o en cualquier combinación. En una realización preferida, la molécula de ácido nucleico codifica una AID mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de AID de tipo silvestre en las sustituciones de aminoácidos en los residuos 10, 82 y 156. En las realizaciones en las que los aminoácidos en los residuos 10, 82 y 156 están sustituidos, también puede ser deseable incluir sustituciones de aminoácidos en los residuos 9, 15, 18, 30, 34, 35, 36, 44, 53, 59, 66, 74, 77, 88, 93, 100, 104, 115, 118, 120, 142, 145, 157, 160, 184, 185, 188, 192 o combinaciones de las mismas. En concreto, (a) la sustitución de aminoácidos en el residuo 9 puede ser serina (S), metionina (M), o triptófano (W), (b) la sustitución de aminoácidos en el residuo 10 puede ser ácido glutámico (E) o ácido aspártico (D), (c) la sustitución de aminoácidos en el residuo 15 puede ser tirosina (Y) o leucina (L), (d) la sustitución de aminoácidos en el residuo 18 puede ser alanina (A) o leucina (L), (e) la sustitución de aminoácidos en el residuo 30 puede ser tirosina (Y) o serina (S), (f) la sustitución de aminoácidos en el residuo 34 puede ser ácido glutámico (E) o ácido aspártico (D), (g) la sustitución de aminoácidos en el residuo 35 puede ser serina (S) o lisina (K), (h) la sustitución de aminoácidos en el residuo 36 puede ser cisteína (C), (i) la sustitución de aminoácidos en el residuo 44 puede ser arginina (R) o lisina (K), (j) la sustitución de aminoácidos en el

residuo 53 puede ser tirosina (I) o glutamina (Q), (k) la sustitución de aminoácidos en el residuo 57 puede ser alanina (A) o leucina (L), (l) la sustitución de aminoácidos en el residuo 59 puede ser metionina (M) o alanina (A), (m) la sustitución de aminoácidos en el residuo 66 puede ser treonina (T) o alanina (A), (n) la sustitución de aminoácidos en el residuo 74 puede ser histidina (H) o lisina (K), (o) la sustitución de aminoácidos en el residuo 77 puede ser serina (S) o lisina (K), (p) la sustitución de aminoácidos en el residuo 82 puede ser isoleucina (I) o leucina (L), (q) la sustitución de aminoácidos en el residuo 88 puede ser serina (S) o treonina (T), (r) la sustitución de aminoácidos en el residuo 93 puede ser leucina (L), arginina (R), o lisina (K), (s) la sustitución de aminoácidos en el residuo 100 puede ser ácido glutámico (E), triptófano (W), o fenilalanina (F), (t) la sustitución de aminoácidos en el residuo 104 puede ser isoleucina (I) o alanina (A), (u) la sustitución de aminoácidos en el residuo 115 puede ser tirosina (I) o leucina (L), (v) la sustitución de aminoácidos en el residuo 118 puede ser ácido glutámico (E) o valina (V), (x) la sustitución de aminoácidos en el residuo 120 puede ser arginina (R) o leucina (L), (i) la sustitución de aminoácidos en el residuo 142 puede ser ácido glutámico (E) o ácido aspártico (D), (z) la sustitución de aminoácidos en el residuo 145 puede ser leucina (L) o tirosina (I), (aa) la sustitución de aminoácidos en el residuo 156 puede ser glicina (G) o alanina (A), (bb) la sustitución de aminoácidos en el residuo 157 puede ser glicina (G) o lisina (K), (cc) la sustitución de aminoácidos en el residuo 160 puede ser ácido glutámico (E) o ácido aspártico (D), (dd) la sustitución de aminoácidos en el residuo 184 puede ser asparagina (N) o glutamina (Q), (ee) la sustitución de aminoácidos en el residuo 185 puede ser glicina (G) o ácido aspártico (D), (ff) la sustitución de aminoácidos en el residuo 188 puede ser glicina (G) o ácido glutámico (E), and (gg) la sustitución de aminoácidos en el residuo 192 puede ser treonina (T) o serina (S).

La proteína mutante de AID funcional puede diferir de la proteína AID de tipo silvestre puede ser cualquiera de las sustituciones de aminoácidos divulgada en el presente documento, solas o en cualquier combinación. Como alternativa, la proteína mutante de AID funcional puede tener sustituciones de aminoácidos adicionales en comparación con una secuencia de aminoácidos de AID de tipo silvestre (p. ej., una secuencia de aminoácidos de AID humana de SEC ID N° 1 o SEC ID N° 2). Por ejemplo, una proteína mutante de AID funcional puede tener una cualquiera, o una combinación de ellas, de las siguientes sustituciones de aminoácidos con respecto a la SEC ID N° 1 o SEC ID N° 2: N7K, R8Q, Q14H, R25H, Y48H, N52S, H156R, R158K, L198A, R9K, G100W, A138G, S173T, T195I, F42C, A138G, H156R, L198F M6K, K10Q, A39P, N52A, E118D, K10L, Q14N, N52M, D67A, G100A, V135A, Y145F, R171H, Q175K, R194K, inserción de K tras el residuo 118, y D119E.

La invención también proporciona moléculas de ácido nucleico que codifican mutantes de AID funcionales que comprenden una mutación por truncamiento en el extremo C. La generación de una mutación por truncamiento en el extremo C está dentro de la experiencia en la materia y se puede realizar, por ejemplo, de acuerdo con los procedimientos descritos anteriormente para generar mutantes de AID. Por ejemplo, la mutación por truncamiento en el extremo C se puede generar mediante la inserción de un codón de terminación en o distal al residuo 181 de la secuencia de aminoácidos de AID.

Ejemplos de sustituciones de aminoácidos preferidas que producen proteínas mutantes funcionales de AID en el contexto de la invención se ilustran en la FIG. 2.

En el contexto de la invención, una mutantes funcionales de AID también incluye una secuencia de ácido nucleico que codifica una proteína AID de tipo silvestre en la que una porción de la secuencia de ácido nucleico se deletorea o reemplaza con una secuencia de ácido nucleico de un homólogo de AID (p. ej., Apobec-1, Apobec3C o Apobec3G). A este respecto, las proteínas APOBEC3 humanas, como la AID humana, son capaces de desaminar la citosina (C) en el ADN, pero mientras que la AID prefiere apuntar a los residuos de C flanqueados por una purina flanqueante en 5', las APOBEC3 prefiere una pirimidina en 5' flanqueante con APOBEC3 individuales que difieren con respecto a la preferencia del nucleótido flanqueante en 5' específico. La comparación de las secuencias génicas de APOBEC3 humanas sugiere que una tira de aproximadamente ocho aminoácidos localizada a aproximadamente 60 residuos del extremo carboxi del dominio de la proteína desempeña un papel importante en la determinación de esta preferencia de nucleótidos flanqueantes. En vista de la estructura cristalina de APOBEC2 y la estructura cristalina de la TadA ARNt-adenosina desaminasa en complejo con un sustrato oligonucleotídico, esta secuencia de 60 aminoácidos en AID y APOBEC3 probablemente forme un contacto con el sustrato de ADN. Por tanto, en una realización de la invención, un mutante funcional de AID puede comprender una secuencia de ácido nucleico que codifica una proteína AID de tipo silvestre en la que los residuos de aminoácidos 115 – 223 de la AID humana se eliminan y reemplazan con la correspondiente secuencia de las proteínas APOBEC3 (p. ej., APOBEC3C, APOBec3F y APOBEC3G).

La invención proporciona adicionalmente moléculas de ácido nucleico que codifican una proteína de fusión que comprende un mutante de AID funcional y un segundo polipéptido condensados en el mismo marco. Por ejemplo, la generación de las proteínas de fusión está dentro de la experiencia en la técnica y puede implicar el uso de enzimas de restricción o técnicas de clonación recombinantes.

En una realización, el segundo polipéptido de la proteína de fusión puede comprender una "señal de localización nuclear" o "SLN". Las expresiones "señal de localización nuclear" y "SLN" hacen referencia a un dominio o dominios capaces de mediar en la importación nuclear de una proteína o polinucleótido, o retención de los mismos, dentro del núcleo de una célula. Una "señal de importación nuclear fuerte" representa un dominio o dominios capaces de mediar en más del 90% de la localización subcelular en el núcleo cuando están operativamente unidos a una

proteína de interés. Ejemplos representativos de SLN incluyen, entre otros, señales de localización nuclear monopartitas, señales de localización nuclear bipartitas y motivos en los extremos N y C. Los dominios básicos en el extremo N normalmente forman la secuencia consenso K-K/R-X-K/R, que se descubrió primero en el antígeno T grande de SV40 y que representa una SLN monopartita. Un ejemplo no limitante de una SLN en el dominio básico en el extremo N es PKKKRKV (SEC ID N° 76). También se conocen señales de localización nuclear bipartitas que contienen dos grupos de aminoácidos básicos separados por un espaciador de aproximadamente 10 aminoácidos, como se ilustra mediante la SLN de la nucleoplasmina: KR[PAATKKAGQA]K KKK (SEC ID N° 77). Los motivos en los extremos N y C incluyen, por ejemplo, el dominio M9 ácido de hnRNP A1, la secuencia KIPK (SEC ID N° 78) en el represor de la transcripción de levaduras Mata2 y las señales complejas de U snRNP. La mayoría de estas SLN parecen estar reconocidas directamente por receptores específicos de la familia  $\beta$  de las importinas.

En otra realización, el segundo polipéptido puede ser una pareja de fusión conocida en la técnica porque facilita la purificación y mejora la solubilidad del polipéptido al que está condensado, por ejemplo, marcador de polihistidina, NusA, bacterioferritina (BFR), GrpE, tioredoxina (TRX) o glutatión-S-transferasa (GST). La purificación de las proteínas de fusión está dentro de la experiencia en la técnica.

En otra realización más, el segundo polipéptido puede ser un polipéptido indicador, tal como una proteína autofluorescente (p. ej., GFP, EGFP). Las proteínas autofluorescentes proporcionan un ensayo fácil para la identificación de la expresión de un polinucleótido (y el producto polipeptídico) de interés. Dado que la actividad del polipéptido indicador (y, por inferencia, su nivel de expresión) se puede monitorizar cuantitativamente usando un clasificador de flujo, se pueden analizar muchos transfectantes independientes bien secuencialmente o en la población en volumen. Las células con la mejor expresión se pueden someter a detección selectiva o seleccionar de la población. Esto es útil al seleccionar una célula recombinante que comprende un mutante de AID funcional de acuerdo con la presente invención.

En una realización adicional de la invención, las moléculas de ácido nucleico que codifican las mutantes de AID funcionales de la invención pueden optimizarse por codones para reducir o aumentar el número de motivos de hipermutación somática (HMS). Como se usa en el presente documento, "hipermutación somática" o "HMS" se refiere a la mutación de una secuencia de polinucleótidos iniciada o asociada con la acción de la AID, mutante de la AID funcional, uracilo glicosilasa y/o polimerasas propensas a errores sobre dicha secuencia de polinucleótidos. Con el término se pretende incluir la mutagénesis que se produce como consecuencia de la reparación propensa a errores de la lesión inicial, incluyendo la mutagénesis mediada por la maquinaria de reparación de apareamientos erróneos y enzimas relacionadas.

La expresión "sustrato para HMS" hace referencia a una secuencia de polinucleótidos sintética o semisintética sobre la que actúa la AID y/o las ADN polimerasas propensas a errores para efectuar un cambio en la secuencia de ácidos nucleicos de la secuencia de polinucleótidos sintética o semisintética.

Como se usa en el presente documento, la expresión "punto caliente de HMS" o "punto caliente" se refiere a una secuencia de polinucleótidos, o motivo, de 3-6 nucleótidos que exhibe una mayor tendencia a sufrir hipermutación somática, como se determina mediante análisis estadístico de las mutaciones de HSM en los genes de anticuerpos. Asimismo, como se usa en el presente documento, la expresión "punto frío de HMS" o "punto frío" se refiere a un polinucleótido, o motivo, de 3-6 nucleótidos que exhibe una menor tendencia a sufrir hipermutación somática, como se determina mediante análisis estadístico de las mutaciones de HSM en los genes de anticuerpos. Una clasificación relativa de varios motivos para la HMS, así como puntos calientes y puntos fríos canónicos en los genes de anticuerpos, se describen en la publicación de solicitud de patente de EE.UU. 09/0075378 y en la publicación de solicitud de patente internacional WO 08/103475 y el análisis estadístico se puede extrapolar al análisis de las mutaciones de HMS en los genes que no son de anticuerpos (p. ej., genes de AID) se describen en el presente documento.

La expresión "motivo de hipermutación somática" o "motivo de HMS" se refiere a una secuencia de polinucleótidos que incluye, o se puede alterar para que incluya, uno o más puntos calientes o puntos fríos, y que codifica un conjunto definido de aminoácidos. Los motivos de HMS pueden ser de cualquier tamaño pero convenientemente se basan en polinucleótidos de aproximadamente 2 a aproximadamente 20 nucleótidos de tamaño o de aproximadamente 3 a aproximadamente 9 nucleótidos de tamaño. Los motivos de HMS pueden incluir cualquier combinación de puntos calientes y puntos fríos o pueden carecer de puntos calientes y puntos fríos.

Las expresiones "codón de HMS de punto caliente preferido", "motivo de HMS de punto caliente preferido", "codón de punto caliente de HMS preferido" y "motivo de punto caliente de HMS preferido" se refieren todas a un codón que incluye, entre otros, los codones AAC, TAC, TAT, AGT o AGC. Dichas secuencias pueden estar incluidas potencialmente dentro del contexto de un motivo de HMS más grande, recluta mutagénesis mediada por HMS y genera una diversidad de aminoácidos dirigida en dicho codón.

Como se usa en el presente documento, una secuencia de ácido nucleico se ha "optimizado para HMS" si la secuencia de ácido nucleico, o una porción de la misma, se ha alterado para aumentar o disminuir la frecuencia y/o localización de puntos calientes y/o puntos fríos dentro de la secuencia de ácido nucleico. Una secuencia de ácido nucleico se ha hecho "susceptible a HMS" si la secuencia de ácido nucleico, o una porción de la misma, se ha alterado para aumentar la frecuencia y/o localización de puntos calientes dentro de la secuencia de ácido nucleico o

disminuir la frecuencia (densidad) y/o localización de los puntos fríos dentro de la secuencia de ácido nucleico. Por el contrario, una secuencia de ácido nucleico se ha hecho “resistente a HMS” si la secuencia de ácido nucleico, o una porción de la misma, se ha alterado para disminuir la frecuencia (densidad) y/o localización de puntos calientes dentro del marco de lectura abierto de la secuencia de ácido nucleico. En general, se puede preparar una secuencia que tiene una propensión mayor o menor a sufrir mutagénesis mediada por HMS alterando el uso de codones y/o los aminoácidos codificados por la secuencia de ácido nucleico.

La optimización de una secuencia de ácido nucleico se refiere a modificar aproximadamente 1 %, aproximadamente 2 %, aproximadamente 3 %, aproximadamente 4 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 50 %, aproximadamente 75 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 96 %, aproximadamente 97 %, aproximadamente 98 %, aproximadamente 99 %, aproximadamente 100 %, o cualquier intervalo entre ellos, de los nucleótidos en la secuencia de ácido nucleico. La optimización de una secuencia de polinucleótidos también se refiere a modificar aproximadamente 1, aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 10, aproximadamente 20, aproximadamente 25, aproximadamente 50, aproximadamente 75, aproximadamente 90, aproximadamente 95, aproximadamente 96, aproximadamente 97, aproximadamente 98, aproximadamente 99, aproximadamente 100, aproximadamente 200, aproximadamente 300, aproximadamente 400, aproximadamente 500, aproximadamente 750, aproximadamente 1000, aproximadamente 1500, aproximadamente 2000, aproximadamente 2500, aproximadamente 3000 o más o cualquier intervalo entre ellos, de los nucleótidos en la secuencia de ácido nucleico, de forma que algunos de estos o todos los nucleótidos se optimizan para la mutagénesis mediada por HMS. La reducción de la frecuencia (densidad) de puntos calientes y/o puntos fríos se refiere a reducir aproximadamente un 1 %, aproximadamente 2 %, aproximadamente 3 %, aproximadamente 4 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 50 %, aproximadamente 75 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 96 %, aproximadamente 97 %, aproximadamente 98 %, aproximadamente 99 %, aproximadamente 100 %, o cualquier intervalo entre ellos, de los puntos calientes y/o puntos fríos. El incremento de la frecuencia (densidad) de puntos calientes y/o puntos fríos se refiere a aumentar aproximadamente un 1 %, aproximadamente 2 %, aproximadamente 3 %, aproximadamente 4 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 20 %, aproximadamente 25 %, aproximadamente 50 %, aproximadamente 75 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 96 %, aproximadamente 97 %, aproximadamente 98 %, aproximadamente 99 %, aproximadamente 100 %, o cualquier intervalo entre ellos, de los puntos calientes y/o puntos fríos.

La posición o el marco de lectura de un punto caliente o puntos frío también es un factor que gobierna si la mutagénesis mediada por HMS que puede dar lugar a una mutación silente con respecto a la secuencia de aminoácidos resultante o causa cambios conservadores, semiconservadores o no conservadores a nivel de aminoácidos. Los parámetros del diseño se pueden manipular para potenciar adicionalmente la susceptibilidad o resistencia relativa de una secuencia de nucleótidos a la HMS. Por tanto, tanto el grado de reclutamiento de la HMS como el marco de lectura del motivo se consideran en el diseño de las secuencias de ácido nucleico susceptibles a la HMS y resistentes a la HMS.

La invención también proporciona un vector que comprende la molécula de ácido nucleico que codifica una AID mutante funcional. Un “vector” o “vector de clonación” es un replicón, tal como un plásmido, fago o cósmico, en el que se puede introducir otro segmento polinucleotídico para producir la replicación del segmento insertado. Los vectores existen normalmente como ADN circular bicatenario y en intervalo de tamaño de unas pocas kilobases (kb) a cientos de kb. Vectores de clonación preferidos se han modificado con respecto a plásmidos de origen natural para facilitar la clonación y la manipulación recombinante de las secuencias de polinucleótidos. Muchos de estos vectores son bien conocidos en la técnica; véase, por ejemplo, Sambrook et al, "Molecular Cloning: A Laboratory Manual," segunda edición, Cold Spring Harbor Laboratory, (1989), y Maniatis et al., Cell Biology: A Comprehensive Treatise, Vol. 3, Gene Sequence Expression, Academic Press, NY, pp. 563 - 608(1980).

La expresión “vector de expresión”, como se usa en el presente documento, se refiere a un vector usado que expresa determinados polinucleótidos dentro de una célula huésped o un sistema de expresión *in vitro*. La expresión incluye plásmidos, epitomas, cósmidos, retrovirus o fagos. El vector de expresión se puede usar para expresar una secuencia de ADN que codifica una proteína deseada y un aspecto incluye una unidad transcripcional que comprende un montaje de las secuencias de control de la expresión. La elección del promotor y otros elementos reguladores generalmente varía de acuerdo con la célula huésped prevista o sistema de expresión *in vitro*.

Como se usa en el presente documento, un “sistema de expresión *in vitro*” se refiere a sistemas sin células que permiten la transcripción o la transcripción y traducción acopladas de moldes de ADN. Dichos sistemas incluyen, por ejemplo, el sistema de reticulocitos de conejo, así como nuevos sistemas de síntesis sin células (J. Biotechnol., 110: 257 - 63 (2004); Biotechnol. Annu. Rev., 10: 1 - 30 (2004)).

“Secuencias de control de la expresión” son secuencias reguladoras de ADN, tales como promotores, potenciadores, señalización de poliadenilación, terminadores de la transcripción, sitios internos entrada al ribosoma (IRES) y similares, que proporcionan la expresión de una secuencia de codificación en una célula huésped. En la técnica se conocen ejemplos de secuencias de control de la expresión y se describen en Goeddel; Gene Expression

Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990).

Un "promotor" es una secuencia de ADN a la que se puede unir la ARN polimerasa en una célula e iniciar la transcripción de una secuencia de codificación cadena abajo (dirección 3'). Como se usa en el presente documento, la secuencia promotora está unida a su extremo 3' mediante el sitio de inicio de transcripción y se extiende cadena arriba (dirección 5') para incluir el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima de los niveles basales. Dentro de la secuencia promotora se encontrará un sitio de inicio de la transcripción (convenientemente definido mediante mapeo con la nucleasa S1), así como dominios de unión a la proteína (secuencias consenso) responsables de la unión de la ARN polimerasa. Los promotores eucariotas a menudo, aunque no siempre, contendrán cajas "TATA" y cajas "CAT". Los promotores procariotas contienen secuencias de Shine-Dalgarno además de las secuencias consenso en -10 y -35.

En la técnica se conoce un gran número de promotores, incluyendo promotores constitutivos, inducibles y represibles, de varias fuentes diferentes. Fuentes representativas incluyen, por ejemplo, los tipos de células virales, de mamífero, de insectos, vegetales, de levaduras y bacterianas) y promotores adecuados de estas fuentes están disponibles fácilmente o se pueden producir de forma sintética en base a las secuencias disponibles públicamente on line o, por ejemplo, de depósitos tales como la ATCC, así como de otras fuentes comerciales o individuales. Los promotores pueden ser unidireccionales (es decir, inician la transcripción en una dirección) o bidireccionales (es decir, inician la transcripción en dirección 3' o 5'). Ejemplos no limitantes de promotores incluyen, por ejemplo, el sistema de expresión bacteriana T7, el sistema de expresión bacteriana (araA), el promotor de citomegalovirus (CMV), el promotor SV40, el promotor RSV. Promotores inducibles incluyen el sistema de Ensayo (patentes de EE.UU. n° 5.464.758 y 5.814.618), el sistema inducible Ecdysone (No et al., Proc. Natl. Acad. Sci., 93: 3346 - 3351

(1996); el sistema T-REX™ (Invitrogen, Carlsbad, CA), LacSwitch® (Stratagene, San Diego, CA) y el sistema de Cre-ERT recombinasa inducible por tamoxifeno (Indra et al., Nuc. Acid. Res., 27: 4324 - 4327 (1999); Nuc. Acid. Res., 28: e99 (2000); patente de EE.UU. 7.112.715). Véase, en general, Kramer & Fussenegger, Methods Mol. Biol., 308: 123 - 144 (2005)) o cualquier promotor conocido en la materia adecuado para la expresión en las células deseadas.

Si se usa un sistema inducible, tal como el sistema controlado por ensayo se puede añadir doxiciclina al medio para inducir la expresión del ácido nucleico que codifica un mutante de AID funcional durante un periodo de tiempo (p. ej., 1 hora (h), 2 horas, 4 horas, 6 horas, 8 horas, 10 horas, 15 horas, 20 horas, 24 horas o cualquier otro tiempo) antes del análisis mediante un ensayo adecuado. Se puede dejar que las células crezcan durante un determinado tiempo para proporcionar una diversificación constante, por ejemplo para 1 - 3 generaciones de células o, en determinados casos, 3-6 generaciones o, en algunos casos, de 6 a 10 generaciones, o más.

Como se usa en el presente documento, un "promotor mínimo" se refiere a una secuencia promotora parcial que define el sitio de inicio de la transcripción pero que, por sí mismo, no puede, en su caso, iniciar la transcripción con eficiencia. La actividad de dichos promotores mínimos depende de la unión de los activadores, tales como el transactivador controlado por tetraciclina, a sitios de unión operablemente unidos.

Los términos "IRES" o "sitio interno de entrada al ribosoma" se refieren a un elemento polinucleotídico que actúa para potenciar la traducción de una secuencia de codificación codificada por un ARN mensajero policistrónico. Los elementos IRE median en la iniciación de la traducción reclutando directamente y uniendo los ribosomas a una molécula de ARN mensajero (ARNm), sorteando el protector de 7-metil-guanosina implicado en el barrido típico del ribosoma. La presencia de una secuencia de IRES puede aumentar el nivel de traducción independiente del protector de una proteína deseada. Las primeras publicaciones se refieren de forma descriptiva a secuencias de IRES como "potenciadores de la traducción". Por ejemplo, se describen "potenciadores de la traducción" de ARN cardioviral en las patentes de EE.UU. N° 4.937.190 y 5.770.428.

El término "potenciador", como se usa en el presente documento, se refiere a una secuencia de ADN que aumenta la transcripción de, por ejemplo, un gen o secuencia de codificación a la que está operablemente unido. Los potenciadores pueden estar a muchas kilobases de la secuencia de codificación y pueden mediar en la unión de factores reguladores, patrones de metilación del ADN o cambios en la estructura del ADN. Un gran número de potenciadores, desde varias fuentes diferentes, son bien conocidos en la técnica y están disponibles como o en polinucleótidos clonados (de, por ejemplo, depósitos tales como la ATCC, así como otras fuentes comerciales o individuales). Una serie de polinucleótidos que comprende promotores (tales como el promotor del CMV de uso habitual) también comprenden secuencias potenciadoras. Los potenciadores unidos operablemente se pueden localizar cadena arriba, dentro o cadena debajo de las secuencias de codificación. La expresión "potenciadores de Ig" se refiere a elementos potenciadores derivados de regiones potenciadoras mapeadas dentro del locus de Ig (dichos potenciadores incluyen, por ejemplo, los potenciadores de la cadena pesada ( $\mu$ ) 5', los potenciadores de la cadena ligera ( $\kappa$ ) 5', potenciadores intrónicos de  $\kappa$  y  $\mu$  y potenciadores 3' (véase, generalmente, Paul WE (ed) Fundamental Immunology, 3ª edición, Raven Press, New York (1993) páginas 353 - 363; patente de EE.UU. 5.885.827).

"Secuencias terminadoras" son las que dan lugar a la terminación de la transcripción. Las secuencias de la terminación se conocen en la técnica e incluyen, entre otros, terminadores de poli A (p. ej., Poli A de Bgh y Poli A de SV40) Normalmente, una señal de terminación de la transcripción incluirá una región de la región no traducida en 3' (o "3' ut"), un intrón opcional (también denominado secuencia intermedia o "IVS") y una o más señales de

poliadenilación ("p(A)" o "pA." Las secuencias de terminación también se pueden preferir como "IVS-pA," "IVS+p(A)," "3' ut+p(A)" or "3' ut/p(A)." Los terminadores naturales o sintéticos se pueden usar como región terminadora.

5 Los términos "poliadenilación", "secuencia de poliadenilación" y "señal de poliadenilación", "Poly A," "p(A)" or "pA" se refieren a una secuencia de ácido nucleico presente en un transcrito de ARN que permite la poliadenilación del transcrito, cuando está en presencia de la enzima poliadenilil transferasa. En la técnica se conocen muchas señales de poliadenilación. Ejemplos no limitantes incluyen la señal de poliadenilación de la hormona de crecimiento variante humana, la señal de poliadenilación tardía de SV40 y la señal de poliadenilación de la hormona de crecimiento bovina.

10 Un "vector de expresión episomal" es capaz de replicarse en una célula huésped y persiste como segmento extracromosómico de ADN dentro de la célula huésped en presencia de presión selectiva adecuada (véase, por ejemplo, Conese et al., *Gene Therapy* 11: 1735 - 1742 (2004)). Vectores de expresión episomal disponible comercialmente representativos incluyen, entre otros, plásmidos episomales que usan el antígeno 1 nuclear de Epstein Barr 1 (EBNA1) y el origen de replicación (oriP) del virus de Epstein Barr Virus (EBV) (oriP). Los vectores  
15 pREP4, pCEP4, pREP7 de Invitrogen, pcDNA3.1 de Invitrogen, y pBK-CMV de Stratagene representan ejemplos no limitantes de un vector episomal que usa el antígeno T y el origen de replicación de SV40 en lugar EBNA1 y oriP.

Un "vector de expresión integrante" puede integrarse aleatoriamente en el ADN de la célula del huésped o puede incluir un sitio de recombinación para permitir la recombinación específica entre el vector de expresión y el cromosoma de las células huésped. Dichos vectores de expresión de integración pueden usar las secuencias de control de la expresión endógenas de los cromosomas de la célula huésped para efectuar la expresión de la proteína deseada. Ejemplos de vectores que integran de un modo específico de sitio incluyen, por ejemplo, los componentes del sistema flip-in de Invitrogen (p. ej., pcDNA™5/FRT), o el sistema cre-lox, de forma que se puede encontrar en los vectores pExchange-6 Core de Stratagene. Ejemplos de vectores que se integran en los cromosomas de la célula huésped de un modo aleatorio incluyen, por ejemplo, pcDNA3.1 (cuando se introducen en ausencia del antígeno T)  
25 de Invitrogen, pCI o pFN10A (ACT) Flexi® de Promega.

Vectores de expresión viral disponibles comercialmente representativos incluyen, entre otros, el sistema Per.C6 basado en adenovirus disponible en Crucell, Inc., el pLP1 basado en lentivirus de Invitrogen y los vectores retrovirales pFB-ERV más pCFB-EGSH de Stratagene.

30 Como alternativa, el vector de expresión se puede usar para introducir e integrar un promotor fuerte o secuencias potenciadoras en un locus en la célula para modular la expresión de un gen endógeno de interés (Capecchi MR. *Nat Rev Genet.*, 6(6): 507 - 12 (2005); Schindehutte et al., *Stem Cells*, 23(1): 10 - 5 (2005)). Este abordaje también se puede usar para insertar un promotor inducible, tal como el promotor Tet-On (patentes de EE.UU. N° 5.464.758 y 5.814.618), en el ADN genómico de la célula para proporcionar la expresión inducible de un gen endógeno de interés. La construcción de activación también puede incluir secuencia(s) diana para permitir la recombinación  
35 homóloga o no homóloga de la secuencia de activación en un locus deseado específico del gen de interés (véase, por ejemplo, Garcia-Otin and Guillou, *Front. Biosci.*, 11: 1108 - 36 (2006)). Como alternativa, un sistema de recombinasa inducible, tal como el sistema Cre-ER, se puede usar para activar un transgén en presencia de 4-hidroxitamoxifeno. (Indra et al., *Nuc. Acid. Res.*, 27(22): 4324 - 4327 (1999); *Nuc. Acid. Res.*, 28(23): e99 (2000); patente de EE.UU. 7.112.715).

40 El vector de la presente invención puede comprender un "gen marcador seleccionable". La expresión "gen marcador seleccionable", como se usa en el presente documento, hace referencia a polinucleótidos que permiten que las células portadoras de polinucleótidos se seleccionan específicamente en presencia de un correspondiente agente selectivo. Los marcadores seleccionables pueden ser positivos, negativos o bifuncionales. Los marcadores seleccionables positivos permiten la selección de células portadoras del marcador, mientras que los marcadores  
45 seleccionables negativos permiten eliminar selectivamente las células portadoras del marcador. El polinucleótido marcador seleccionable puede estar unido directamente a los polinucleótidos que se van a expresar o introducirse en la misma célula mediante cotransfección. Se han descrito varios de estos polinucleótidos marcadores, incluyendo, por ejemplo, marcadores bifuncionales (es decir positivos/negativos) (véase, por ejemplo, las publicaciones de solicitud de patente internacional WO 92/08796 y WO 94/28143), genes de resistencia a fármacos (p. ej., ampicilina) y proteínas que confieren resistencia a fármacos citostáticos o citocidas (p. ej., la proteína DHFR) (véase, por ejemplo, *Proc. Natl. Acad. Sci. USA*, 77: 3567 (1980), O'Hare et al., *Proc. Natl. Acad. Sci. USA*, 78: 1527 (1981), Mulligan & Berg, *Proc. Natl. Acad. Sci. USA*, 78: 2072 (1981), Colberre-Garapin et al., *J. Mol. Biol.*, 150: 1 (1981), Santerre et al., *Gene*, 30: 147 (1984), Kent et al., *Science*, 237: 901 - 903 (1987), Wigler et al., *Cell*, 11: 223 (1977), Szybalska & Szybalski, *Proc. Natl. Acad. Sci. USA*, 48:2026 (1962), Lowy et al., *Cell*, 22:817 (1980), y las  
50 patentes de EE.UU. N° 5.122.464 y 5.770.359).

El vector puede comprender un "gen indicador". Un "gen indicador" se refiere a un polinucleótido que confiere la capacidad para detectarse específicamente (o detectarse y seleccionarse), normalmente cuando se expresa con una célula de interés. En la técnica se conocen numerosos sistemas de genes indicadores e incluye, por ejemplo, la fosfatasa alcalina (Berger, J., et al., *Gene*, 66: 1 - 10 (1988); Kain, SR., *Methods Mol. Biol.*, 63: 49 - 60 (1997)), la agalactosidasa (patente de EE.UU. 5.070.012), cloranfenicol acetiltransferasa (Gorman et al., *Mol. Cell. Biol.*, 2: 1044  
60

- 51 (1982)), beta glucuronidasa, peroxidasa, beta lactamasa (las patentes de EE.UU. N° 5.741.657 y 5.955.604), anticuerpos catalíticos, luciferasas (las patentes de EE.UU. N° 5.221.623; 5.683.888; 5.674.713; 5.650.289; y 5.843.746) y proteínas fluorescentes naturales (Tsien, RY, Annu. Rev. Biochem., 67: 509 - 544 (1998)). La expresión "gen indicador" también incluye cualquier péptido que se pueda detectar específicamente en base al uso de uno o más anticuerpos, epítomos, parejas de unión, sustratos, enzimas de modificación, receptores o ligandos que son capaces, o se desea (o no se desea), de interactuar con el péptido de interés para crear una señal detectable. Los genes indicadores también incluyen genes que pueden modular el fenotipo celular. La proteína indicadora, cuando sirve para tal fin de detección, no se tiene que condensar con la proteína AID mutante. Puede estar codificada por el mismo polinucleótido (p. ej., un vector), que también codifica la proteína AID mutante y se puede cointroducir y coexpresar en una célula diana.

Los vectores de expresión también pueden incluir polinucleótidos antisentido, ribozimas o ARNs para reducir la expresión de las secuencias diana (véase, por ejemplo, Sioud M, & Iversen, Curr. Drug Targets, 6: 647 - 53 (2005); Sandy et al., Biotechniques, 39:215 - 24 (2005)).

La invención también proporciona una célula que comprende una molécula de ácido nucleico que codifica una AID mutante funcional o un vector que comprende la molécula de ácido nucleico que codifica una AID mutante funcional. Las expresiones "células", "cultivos celulares", "línea celular", "células huésped recombinantes", "células receptoras" y "células huésped" a menudo se usan de forma intercambiable e incluyen células sujeto primarias y cualquier progenie de las mismas, sin considerar el número de transferencias. Debe entenderse que no toda la progenie es exactamente idéntica a la célula parental (debido a mutaciones deliberadas o inadvertidas o diferencias en el ambiente). No obstante, dicha progenie alterada se incluye en estos términos, siempre que la progenie conserve la misma funcionalidad que la célula transformada originalmente. Por ejemplo, aunque no limitada a la misma, dicha característica podría ser la capacidad para producir una proteína recombinante concreta. Una "línea celular positiva mutadora" es una línea celular que contiene factores celulares que son suficientes para funcionar en combinación con otros elementos vectores para efectuar hipermutación. La línea celular puede ser cualquiera de las conocidas en la técnica o se describen en el presente documento. Un "clon" es una población de células derivada de una única célula o ancestro común por mitosis.

Los sistemas de expresión basada en células y de hipermutación incluyen cualquier sistema de expresión procarionota o eucariota adecuado. Los sistemas preferidos son aquellos que se pueden cultivar fácil y fiablemente, tienen velocidades de crecimiento razonablemente rápidos, tienen sistemas de expresión bien caracterizados y se pueden transformar o transfectar fácil y eficientemente.

Células microbianas útiles incluyen, entre otras, células de los géneros *Bacillus*, *Escherichia* (tales como *E. coli*), *Pseudomonas*, *Streptomyces*, *Salmonella*, *Erwinia*, *Bacillus subtilis*, *Bacillus brevis*. Células procarionotas particularmente útiles incluyen las diversas cepas de *Escherichia coli* (p. ej., K12, HB101, (ATCC N° 33694) DH5 $\alpha$ , DH10, MC1061 (ATCC N° 53338), y CC102)

Muchas cepas de células de levadura conocidas por los expertos en la técnica también están disponibles como células huésped para la expresión de polipéptidos, incluyendo los de los géneros *Hansenula*, *Kluyveromyces*, *Pichia*, *Rhinosporidium*, *Saccharomyces*, y *Schizosaccharomyces*, y otros hongos. Células de levadura preferidas incluyen, por ejemplo, *Saccharomyces cerevisiae* y *Pichia pastoris*.

Adicionalmente, cuando se desee se pueden utilizar sistemas de células de insectos en los procedimientos de la presente invención. Dichos sistemas se describen en, por ejemplo, Kitts et al., Biotechniques, 14: 810 - 8e17 (1993); Lucklow, Curr. Opin. Biotechnol., 4: 564 - 572 (1993); and Lucklow et al., J. Virol., 67: 4566 - 4579 (1993). Células de insectos preferidas incluyen Sf-9 y HI5 (Invitrogen, Carlsbad, Calif.).

La célula que comprende el ácido nucleico que codifica un mutante de AID funcional es, preferentemente, una célula de mamífero. En la técnica también se conoce una serie de células huésped de mamífero adecuadas y muchas están disponibles en la Colección Americana de Cultivos Tipo (ATCC, Manassas, VA). Ejemplos de células de mamífero adecuadas incluyen, entre otras, células de ovario de hámster chino (CHO) (N° ATCC CCL61) células CHO DHFR (Urlaub et al., Proc. Natl. Acad. Sci. USA, 97: 4216 - 4220 (1980)), células renales embrionarias humanas (HEK) 293 o 293T (ATCC N° CRL1573) y células 3T3 (N° ATCC CCL92). La selección de células huésped de mamífero adecuadas y procedimientos de transformación, cultivo, amplificación, detección selectiva y producción y purificación de productos se conocen en la técnica. Otras líneas de células de mamífero adecuadas son las líneas celulares COS-1 de mono (N° ATCC CRL165) y COS-7 (N° ATCC CRL1651) y la línea celular CV-1 (N° ATCC CCL70). Otras células huésped de mamífero de ejemplo incluyen líneas celulares de primate y líneas celulares de roedores, incluyendo líneas celulares transformadas. También son adecuadas las células diploides normales, cepas de células derivadas de cultivos *in vitro* de tejido primario, así como explantes primarios. Las células candidatas pueden ser genotípicamente deficientes en el gen de selección o pueden contener un gen de selección de acción dominante. Otras líneas de células de mamífero incluyen, entre otros, células N2A de neuroblastoma de ratón, HeLa, células L-929 de ratón, líneas 3T3 derivadas de ratones Swiss, Balb-c o NIH, líneas células de hámster BHK o HaK, que están disponibles en la ATCC.

También están dentro del alcance de la invención las líneas celulares linfoides y derivadas linfoides, tales como una línea celular de origen pre-linfocito B. Ejemplos específicos incluyen, entre otros, RAMOS (CRL-1596), Daudi (CCL-

213), EB-3 (CCL-85), DT40 (CRL-2111), 18 - 81 (Jack et al., Proc. Natl. Acad. Sci. USA, 85: 1581 - 1585 (1988)), células Raji (CCL-86), y derivados de los mismos.

Un mutante de AID funcional de la presente invención se puede introducir en una célula mediante "transfección", "transformación" o "transducción". La "transfección", "transformación" o "transducción", como se usa en el presente documento, hacen referencia a la introducción de uno o más polinucleótidos exógenos en una célula huésped usando uno de procedimientos físicos o químicos. Los expertos en la técnica incluyen muchas técnicas de transfección, incluyendo, entre otros, coprecipitación de ADN en fosfato cálcico (véase *Methods in Molecular Biology*, Vol. 7, Gene Transfer and Expression Protocols, Ed. E. J. Murray, Humana Press (1991)); DEAE-dextrano; electroporación, transfección mediada por liposomas catiónicos, bombardeo con micropartículas de tungsteno (Johnston, S. A., *Nature*, 346: 776 - 777 (1990)); y coprecipitación de ADN en fosfato de estroncio (Brash D. E. et al. *Molec. Cell. Biol.*, 7: 2031 - 2034 (1987)). Los vectores fagos o retrovirales se pueden introducir en las células huésped después del crecimiento de partículas infecciosas en empaquetamiento de células que están disponibles comercialmente.

La invención también proporciona un procedimiento de preparación de un producto génico que tiene una propiedad deseada, en el que el procedimiento comprende expresar un ácido nucleico que codifica el producto génico en una población de células, en la que la población de células expresa, o se puede inducir que exprese, una proteína AID mutante funcional de la presente invención, con lo cual la expresión de la proteína AID mutante funcional induce una mutación en el ácido nucleico que codifica el producto génico. Las descripciones del mutante de AID funcional, células y procedimientos de transfectar y expresar moléculas de ácido nucleico en las células expuestas anteriormente en relación con otras realizaciones de la invención también son aplicables a los mismos aspectos del procedimiento citado anteriormente.

Deseablemente, la proteína de AID mutante funcional induce una mutación en el ácido nucleico que codifica el producto génico a modo de hipermutación somática (HMS). El uso de AID en los sistemas de HMS se describe con detalle en la publicación de solicitud de patente de EE.UU. 09/0075378 y las publicaciones de solicitud de patente internacional WO/08103474 y WO 08/103475. Como se usa en el presente documento, la expresión "producto génico de interés" o "proteína de interés" se refiere a proteínas, o porciones de las mismas, para las cuales se desea que el ácido nucleico que codifica el producto génico esté optimizado para la HMS mediante un mutante de AID funcional con el fin de crear rápidamente, seleccionar e identificar mejores variantes de dicho producto génico. Dichas secuencias de ácido nucleico optimizadas se pueden hacer más susceptibles a la HMS como resultado del uso de codones (como se describe en el presente documento), de modo que se inducen cambios de aminoácidos cuando el polinucleótido se somete a una mutante de AID funcional y se somete a detección selectiva para la función mejorada. Por el contrario, dichas secuencias de ácido nucleico optimizadas se pueden hacer más resistentes a la HMS (como se describe en el presente documento), de modo que se disminuyen los cambios de aminoácidos cuando el polinucleótido se somete a una mutante de AID funcional como resultado del uso de codones, y se somete a detección selectiva para la función mejorada.

Cualquier proteína para la cual el aminoácido, o la correspondiente secuencia de nucleótidos, se conoce o está disponible (p. ej., se puede clonar en un vector como se describe en el presente documento) y un fenotipo o función se pueden mejorar, es un candidato para usar en el procedimiento de la invención. Ejemplos de proteínas adecuadas incluyen, por ejemplo, proteínas de superficie, proteínas intracelulares, proteínas de membrana y proteínas secretadas de cualquier fuente sintética o no modificada. El producto génico es, preferentemente, una cadena pesada de anticuerpo, o porción del mismo, una cadena ligera de anticuerpo, o porción del mismo, una enzima, un receptor, una proteína estructural, un cofactor, un polipéptido, un péptido, un intracuerpo, un marcador seleccionable, una toxina, factor de crecimiento, hormona peptídica, o cualquier otra proteína que se puede optimizar.

El producto génico puede ser cualquier enzima adecuada, incluyendo las enzimas asociadas con fermentación microbiana, modificación genética de la vía metabólica, fabricación de proteínas, bioremedios y crecimiento y desarrollo de plantas (véase, por ejemplo Olsen et al., *Methods Mol. Biol.*, 230: 329 - 349 (2003); Turner, *Trends Biotechnol.*, 21(11): 474 - 478 (2003); Zhao et al., *Curr. Opin. Biotechnol.*, 13(2): 104 - 110 (2002); y Mastrobattista et al., *Chem. Biol.*, 12(12): 1291 - 300 (2005)).

Receptores adecuados para usar en el procedimiento de la invención incluyen, entre otros, receptores unidos a células tales como anticuerpos (receptores de linfocitos B), receptores de linfocitos T, receptores de Fc, receptores proteicos acoplados a proteínas G, receptores de citoquinas, receptores de hidratos de carbono y receptores basados en Avimer™. Dichos receptores se pueden alterar mediante HMS para mejorar uno o más de los rasgos siguientes, afinidad, avidéz, selectividad, termoestabilidad, estabilidad proteolítica, solubilidad, dimerización, plegamiento, inmunotoxicidad, acoplamiento a las cascadas de transducción de señal y expresión.

Productos génicos adecuados para usar en el procedimiento de la invención también incluyen moléculas capaces de modular la farmacocinética y/o la farmacodinámica de otras proteínas biológicamente activas, por ejemplo lípidos y polímeros, tales como poliaminas poliamidas, polietilenglicol y otros poliéteres. Otros ejemplos de productos génicos adecuados para usar en el procedimiento de la invención incluyen polipéptidos tales como VEGF, receptor del VEGF, subunidad A de la toxina diftérica, toxina de *B. pertussis*, quimioquinas CC (p. ej., CCL1-CCL28),

5 quimioquinas CXC (p. ej., CXCL1 - CXCL16), quimioquinas C (p. ej., XCL1 y XCL2) y quimioquinas CX3C (p. ej., CX3CL1), IFN-gamma, IFN-alfa, IFN-beta, TNF-alfa, TNF-beta, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-10, IL-12, IL-13, IL-15, TGF-beta, TGF-alfa, GM-CSF, G-CSF, M-CSF, TPO, EPO, factor de crecimiento humano, factor de crecimiento de fibroblastos, cofactores nucleares, miembros de la familia Jak y Stat, moléculas de señalización de la proteína G tales como receptores de quimioquinas, JNK, Fos-Jun, NF-κB, I-κB, CD40, CD4, CD8, B7, CD28 y CTLA-4. Los procedimientos para seleccionar un producto génico (p. ej., proteína) de interés como un candidato adecuado para la mutación y optimización mediante HMS, así como ensayos de detección selectiva relacionados, se divulgan adicionalmente en la publicación de solicitud de patente de EE.UU. 09/0075378 y las publicaciones de solicitud de patente internacional WO/08103474 y WO 08/103475

10 En una realización preferida de la invención, la secuencia de ácido nucleico que está sujeta a mutación por la proteína mutante de AID funcional codifica un anticuerpo, o una porción del mismo. Las secuencias de ácido nucleico que codifican todos los anticuerpos de origen natural de la línea germinal, madurados por afinidad, sintéticos o semisintéticos, así como fragmentos de los mismos, se puede usar en la presente invención. En general, dichas secuencias que codifican anticuerpos se pueden alterar mediante HMS para mejorar uno o más de los  
15 siguientes rasgos funcionales: afinidad, avidez, selectividad, termoestabilidad, estabilidad proteolítica, solubilidad, plegamiento, inmunotoxicidad y expresión. Dependiendo del formato de anticuerpo, se pueden generar bibliotecas que comprenden bibliotecas separadas de la cadena pesada y la cadena ligera que se pueden coexpresar en una célula huésped. En determinadas realizaciones, se pueden secretar (o liberar) anticuerpos de longitud completa y/o se pueden expresar en la superficie de la membrana plasmática de la célula huésped. En otras realizaciones más,  
20 las bibliotecas de las cadenas pesadas y ligeras se pueden insertar en el mismo vector de expresión o en vectores de expresión diferentes para permitir la coevolución simultánea de ambas cadenas de anticuerpos.

Por tanto, el procedimiento de la invención proporciona la capacidad para sortear la necesidad de inmunización *in vivo* para seleccionar anticuerpos que se unen a epítopos de superficie clave que están alineados con la producción de efectos biológicos más sólidos sobre la función proteica diana. Adicionalmente, los anticuerpos de mamífero procesan intrínsecamente patrones óptimos de uso de codones para la HMS dirigida, simplificando  
25 considerablemente las estrategias de diseño del molde. Para determinados antígenos, la inmunización *in vivo* conduce a la selección de epítopos que no afecta a la función diana, de modo que se dificulta la selección de potentes y eficaces candidatos a anticuerpos. En otras realizaciones más, el procedimiento de la invención puede proporcionar la evolución rápida de los anticuerpos dirigidos a sitios que tienen una actividad potente por naturaleza del papel del dicho epítipo en la determinación de la función proteica diana. Esto proporciona la capacidad para  
30 seleccionar proteínas diana para la posición óptima del epítipo y producir los mejores fármacos anticuerpos para su uso en la clínica.

El procedimiento de la invención se puede usar para aumentar la densidad de los puntos calientes en los subdominios específicos de anticuerpos o fragmentos de los mismos (p. ej., F(ab')<sub>2</sub>, Fab', Fab, Fv, scFv, dsFv, dAb o un polipéptido de unión monocatenario), que puede tener como resultado una mejora en una característica (p. ej., incremento de la afinidad de unión, incremento de la avidez de unión y/o disminución de la unión inespecífica). El procedimiento de la invención también se puede usar para generar anticuerpos sintéticos con mayores puntos calientes en el dominio constante (p. ej., Fc), que puede tener como resultado una mayor afinidad de unión por un receptor Fc ((FcR), de modo que se modula las cascadas de señales. Las cadenas pesadas y las cadenas ligeras, o porciones e las mismas, se pueden modificar simultáneamente usando los procedimientos descritos en el presente documento.

Los intracuerpos se pueden modificar usando el procedimiento de la invención para mejorar o potenciar el plegamiento de la cadena pesada y/o ligera en la reducción del ambiente del citoplasma. Como alternativa, o además de, un intracuerpo sFv se puede modificar para estabilizar los armazones que puedan plegarse  
45 adecuadamente en ausencia de puentes disulfuro intradominios. Los intracuerpos también se pueden modificar para aumentar, por ejemplo, una o más de las características siguientes: Afinidad de unión, avidez de unión, accesibilidad al epítipo, competición con proteínas endógenas por el epítipo diana, semivida, secuestro diana, modificación postraduccional de la proteína diana etc. Dado que los intracuerpos actúan dentro de la célula, su actividad es más análoga a las metodologías de ensayo para los ensayos de la actividad enzimática.

50 Los procedimientos para diseñar y crear bibliotecas de anticuerpos, así como los procedimientos para identificar epítopos óptimos que proporcionan la selección de anticuerpos con una actividad superior, reactividad de especies cruzadas y actividad de bloqueo se conocen en la técnica (véase, por ejemplo, la publicación de solicitud de patente de EE.UU. 09/0075378 y publicaciones de solicitud de patente internacional WO 08/103474 y WO 08/103475). Cribas específicas para detectar y seleccionar anticuerpos expuestos en la superficie o secretados con mejores rasgos son bien conocidas en la técnica. Dichas cribas pueden implicar varias rondas de selección en base a la  
55 selección simultánea de múltiples parámetros, por ejemplo afinidad, avidez, selectividad y termoestabilidad con el fin de desarrollar el mejor anticuerpo global.

Se apreciará que existe una variedad de otras secuencias de nucleótidos componentes, tales como secuencias de codificación y elementos genéticos, que un experto en la técnica preferirá que la proteína de AID mutante funcional no mute para mantener la integridad global del sistema. Estas secuencias de nucleótidos componentes se describen  
60 en el presente documento e incluirán, sin limitaciones, (i) marcadores seleccionables, (ii) genes indicadores, (iii)

señales reguladoras genéticas, (iv) enzimas o factores accesorios usados para HMS potenciada de alto nivel o su regulación o medición (p. ej., AID o un mutante de AID funcional, pol eta, factores de transcripción y MSH2, (v) componentes de transducción de señal (p. ej., quinasas, receptores, factores de transcripción y (vi) dominios o subdominios de proteínas (p. ej., señales de localización nuclear, dominios transmembrana, dominios catalíticos, dominios de interacción proteína-proteína y otros motivos conservados, dominios y subdominios de la familia proteica).

Dependiendo de la naturaleza del producto génico de interés y la cantidad de información disponible sobre el producto génico de interés, el experto en la técnica puede seguir cualquier combinación de las estrategias siguientes antes, o junto con, la práctica del procedimiento de la invención para preparar un producto génico de interés con una propiedad deseada.

1. Sin optimización de HMS: Aunque puede ser deseable potenciar el número de puntos calientes dentro de la secuencia de ácido nucleico que codifica un producto génico de interés, cabe destacar que se prevé que cualquier secuencia de ácido nucleico sin modificar sufre una cantidad determinada de HMS y se puede usar en el procedimiento de la invención sin optimizar o cualquier conocimiento específico de la secuencia real. Además, determinadas proteínas (p. ej., anticuerpos) comprenden de forma natural secuencias de ácido nucleico que han desarrollado un uso adecuado de codones y no requieren modificación de codones. Como alternativa, puede ser deseable potenciar el número de puntos calientes dentro de la secuencia de ácido nucleico que codifica un producto génico de interés (p. ej., regiones almacén de anticuerpos o fragmentos de los mismos).

2. Optimización de puntos calientes para HMS global; En algunos aspectos, el número de puntos calientes en una secuencia de ácido nucleico que codifica un producto génico se puede aumentar, como se describe con detalle en la publicación de solicitud de patente de EE.UU. 09/0075378 y la publicación de solicitud de patente internacional WO 08/103475. Este abordaje se puede aplicar a la totalidad de la región de codificación de la secuencia de ácido nucleico, de modo que hace que la totalidad de la secuencia de ácido nucleico sea más susceptible a la HMS. Este abordaje se puede preferir si se sabe relativamente poco sobre las relaciones de la actividad de la estructura del producto génico o entre isotipos relacionados.

3. Optimización de puntos calientes para HMS selectiva: Como alternativa, una secuencia de ácido nucleico que codifica la proteína de interés se puede modificar de forma selectiva y/o sistemática mediante sustitución dirigida de regiones de interés con regiones variables sintéticas, como se describe en la publicación de solicitud de patente de EE.UU. 09/0075378 y la publicación de solicitud de patente internacional WO 08/103475, que proporcionan una elevada densidad de puntos calientes y máxima diversidad de sembrado mediante HMS en loci específicos.

Un experto en la técnica entenderá, en base a lo anterior, que cualquiera o todos los abordajes anteriores se pueden realizar junto con el procedimiento de la invención. No obstante, es probable que los procedimientos para optimización de puntos calientes por HMS y modificación puntos calientes por HMS conduzcan a una optimización más rápida y más eficiente de la función proteica.

Tras el diseño de una secuencia de ácido nucleico optimizada con HMS que codifica el producto génico de interés, se puede sintetizar usando metodología estándar y secuenciar para confirmar la síntesis correcta. Una vez que la secuencia de la secuencia de ácido nucleico se ha confirmado, la secuencia de ácido nucleico se puede insertar en un vector, tal como se describe en el presente documento, y, después, el vector se puede introducir en una célula huésped, tal como se describe en el presente documento. Los potenciadores (p. ej., potenciadores de Ig) se pueden insertar en un vector para aumentar la expresión y/o dirigir la HDM iniciada por la proteína mutante de AID funcional a la secuencia de ácido nucleico que codifica un producto génico de interés.

De acuerdo con el procedimiento de la invención, cualquiera de los vectores descritos en el presente documento se puede cotransfectar en una célula huésped con un vector separado que contiene la secuencia de ácido nucleico que codifica un mutante de AID funcional como se describe en el presente documento. En un aspecto, los vectores descritos en el presente documento se pueden transfectar en una célula huésped que contiene (y expresa) la proteína AID endógena. En otro aspecto, los vectores descritos en el presente documento pueden cotransfectarse en una célula huésped que contiene una proteína AID endógena con un vector separado que contiene la secuencia de ácido nucleico de un mutante de AID funcional de forma que el mutante de AID funcional se sobreexpresa en la célula. En otro aspecto más, los vectores descritos en el presente documento se pueden modificar para incluir una secuencia de ácido nucleico que codifica un mutante de AID funcional para la transfección en una célula huésped que contiene, o no, una proteína AID endógena. En una realización preferida, la mutante de AID funcional es una AID sintética que está codificada por una secuencia de ácido nucleico que es resistente a la HMS.

Tras la introducción de uno o más ácidos nucleicos en un vector de expresión, el vector se puede amplificar, purificar, introducir en una célula huésped usando técnicas de transfección y caracterizar usando técnicas de biología molecular estándar. El ADN plasmídico purificado se puede introducir en una célula huésped usando técnicas de transfección/transformación estándar y los transformantes/transfectantes resultantes cultivados en medio que contiene antibióticos, agentes seleccionables y/o señales de activación /transactivador (p. ej., agentes inducibles tales como doxiciclina) para inducir la expresión de la secuencia de ácido nucleico que codifica el producto génico de interés.

El procedimiento de la invención puede comprender además introducir en la célula o población de células uno o más de los siguientes (i) al menos una secuencia de ácido nucleico que se ha alterado, completa o parcialmente, a partir de una correspondiente secuencia de ácido nucleico de tipo silvestre para influir positivamente sobre la tasa de HMS experimentada por dicha secuencia de ácido nucleico o una secuencia de ácido nucleico que tiene un porcentaje naturalmente alto de puntos calientes antes de cualquier modificación y/o (ii) una secuencia de ácido nucleico que se ha alterado, completa o parcialmente, para influir negativamente a la tasa de HMS.

En un aspecto, el procedimiento de la invención puede comprender además introducir en la célula o población de células una o más secuencias de ácido nucleico que se ha alterado con respecto a una correspondiente secuencia de ácido nucleico de tipo silvestre para influir positivamente sobre la tasa de HMS. La secuencia de ácido nucleico puede codificar, por ejemplo, uno o más factores para HMS (p. ej., AID, Pol eta, UDG), uno o más genes de marcadores seleccionables o uno o más genes indicadores.

En otro aspecto, el procedimiento de la invención puede comprender además introducir en la célula o población de células una o más secuencias de ácido nucleico que se ha alterado con respecto a una correspondiente secuencia de ácido nucleico de tipo silvestre para influir positivamente sobre la tasa de HMS. La secuencia de ácido nucleico puede codificar, por ejemplo, una enzima, receptor, factor de transcripción, proteína estructural, toxina, cofactor o proteína de unión específica de interés.

En otro aspecto más, el procedimiento de la invención puede comprender además introducir en la célula o población de células una secuencia de ácido nucleico que tiene una tasa intrínsecamente alta de HMS tal como, por ejemplo, una secuencia de ácido nucleico que codifica una cadena pesada de inmunoglobulina o una cadena ligera de inmunoglobulina, o una región hipervariable de un gen de anticuerpo.

La célula o población de células del procedimiento de la invención puede comprender además uno o más de los elementos adicionales siguientes (i) un sistema inducible para regular la expresión de AID, un homólogo de AID o un mutante de AID funcional de la presente invención, (ii) uno o más potenciadores de Ig, (iii) una o más cajas E, (iv) uno o más factores auxiliares para HMS, (v) uno o más factores para la expresión episomal estable, tal como EBNA1, EBP2 o ori-P, (vi) uno o más genes marcadores seleccionables, (vii) uno o más vectores secundarios que contienen el gen para AID, un homólogo de AID o un mutante de AID funcional de la presente invención o (viii) una combinación de los mismos.

En otro aspecto de la invención, el procedimiento comprende expresar dos secuencias de ácido nucleico, cada una de ellas codifica un producto génico de interés, en el que ambas secuencias de ácido nucleico se localizan cerca de un promotor y se expresan y codesarrollan en la misma célula de forma simultánea. El promotor puede ser un promotor bidireccional tal como un promotor de CMV bidireccional. En otra realización, las dos secuencias de ácido nucleico de interés se colocan en frente de dos promotores unidireccionales. Los dos promotores pueden ser el mismo promotor o diferentes promotores. Las dos secuencias de ácido nucleico de interés pueden estar en el mismo vector o en diferentes vectores.

La célula o población de células se expresa de forma constitutiva o se puede inducir para expresar una proteína AID mutante funcional como se describe en el presente documento. La expresión de la proteína de AID mutante funcional induce una mutación en la secuencia de ácido nucleico que codifica el producto génico. La célula o población de células también puede expresar otros factores que potencian la mutación mediada por AID de la secuencia de ácido nucleico. Como resultado del procedimiento de la invención, se consigue una diversificación constante de la secuencia de la secuencia de ácido nucleico que codifica el producto génico de interés. Después de un periodo de tiempo adecuado (p. ej., 2-10 divisiones celulares), las células huésped resultantes, que incluyen variantes del producto génico de interés, se pueden someter a detección selectiva e identificar los mutantes mejorados y separar de la población celular. Las células se pueden cultivar repetidamente, analizar y seleccionar como se ha descrito en el presente documento para enriquecer de forma selectiva las células que expresan una secuencia de ácido nucleico que codifica un producto génico de interés que exhibe una propiedad deseada. En la técnica se conocen los ensayos adecuados y estrategias de enriquecimiento (p. ej., clasificación de células activada por fluorescencia (FACS), separación por afinidad, actividad enzimática, unión al receptor, estimulación del crecimiento etc.) y se describen en, por ejemplo, la publicación de solicitud de patente de EE.UU. 09/0075378 y las publicaciones de solicitud de patente internacional WO 08/103475 y WO 08/103474.

En una realización de la invención, la secuencia de ácido nucleico que codifica el producto génico de interés se puede modificar genéticamente de forma que el producto génico de interés se exprese en la superficie celular. A este respecto, una proteína expresada en la superficie celular se puede crear mediante la creación de una molécula quimérica de una proteína de interés acoplada en el marco con un dominio transmembrana adecuado. En el caso de expresión en células de mamífero, se puede usar, por ejemplo, un dominio transmembrana de MHC de tipo 1, como el de H2kk (incluido el dominio peri-transmembrana, dominio transmembrana y dominio citoplasmático, número de acceso en NCBI Gene AK153419). Asimismo, la expresión en superficie de proteínas en células procariontas (tales como *E. coli* and *Staphylococcus*), células de insecto y levaduras está bien establecida en la técnica (véase, por ejemplo, Winter et al., *Annu. Rev. Immunol.*, 12: 433 - 55 (1994); Pluckthun A. *Bio/Technology*, 9: 545 - 551 (1991); Gunneriusson et al., *J. Bacteriol.*, 78: 1341 - 1346 (1996); Ghiasi et al., *Virology*, 185: 187 - 194 (1991); Boder and Wittrup, *Nat. Biotechnol.*, 15: 553 - 557 (1997); y Mazor et al., *Nat. Biotechnol.*, 25(5): 563 - 565 (2007)).

Los anticuerpos o proteínas expresados en la superficie se pueden crear mediante la secreción y después la unión (o asociación) de la proteína secretada sobre la superficie celular. La conjugación del anticuerpo o proteína a la membrana celular se puede producir durante la síntesis de proteínas o después de que la proteína se ha secretado de la célula. La conjugación se puede producir mediante enlace covalente, mediante interacciones de unión (p. ej., mediada por miembros de unión específicos) o una combinación de enlaces covalentes y no covalentes. Las proteínas también se pueden acoplar a una célula a través de la creación de un anticuerpo o proteína de fusión de unión a proteína que comprende un primer miembro de unión específica que se une específicamente a una diana de interés condensada con un segundo miembro de unión específica para la expresión sobre una superficie celular (p. ej., en el caso de explotar la unión de la proteína A y un dominio Fc: la proteína A se expresa sobre y se une a una superficie celular y se une a, y localiza, un anticuerpo secretado (o una proteína de interés expresada como una proteína de fusión de Fc)).

En algunos casos, puede ser deseable convertir una proteína expresada en superficie en una proteína que se expresa o desprende de la célula para su caracterización adicional. La conversión se pueden conseguir mediante el uso de un ligador específico que se puede escindir mediante incubación con una proteasa selectiva, tal como el factor X, la trombina o cualquier otro agente proteolítico selectivo. También es posible incluir las secuencias de ácido nucleico que permiten la manipulación genética de la proteína codificada en el vector (es decir, que permiten la escisión de una señal de unión a la superficie del marco de lectura proteico). Dicha manipulación genética se puede conseguir usando un sistema de recombinación. Un "sistema de recombinación" como se usa en el presente documento se refiere a un sistema que permite la recombinación entre un vector y un cromosoma para la incorporación de un gen de interés. Los sistemas de recombinación se conocen en la técnica e incluyen, por ejemplo, sistemas Cre/Lox y sistemas FLP-IN (véase, por ejemplo, Abremski et al., Cell, 32: 1301 - 1311 (1983) y las patentes de EE.UU. N° 4.959.317; 5.654.182; y 5.677.177). Por ejemplo, la inserción de uno o más sitios de restricción únicos, o elementos cre/lox u otros elementos de recombinación que permiten la eliminación selectiva de una señal de unión y la posterior acumulación intracelular (o secreción) de la proteína de interés. Otros ejemplos incluyen la inserción de sitios loxP flanqueantes alrededor de una señal de unión (tal como un dominio transmembrana), lo que permite una eficiente expresión en la superficie celular de una proteína de interés. No obstante, tras la expresión de la cre recombinasa en la célula, la recombinación se produce entre los sitios LoxP, lo que tiene como resultado la pérdida de la señal de unión y, por tanto, que conduce a la liberación o desprendimiento de la proteína de interés.

Un producto génico de interés generado por el procedimiento de la invención se puede someter a detección selectiva de una propiedad deseada (p. ej., un fenotipo seleccionable o mejorado) usando diversos procedimientos fisiológicos, farmacológicos y bioquímicos estándar. Dichos ensayos incluyen, por ejemplo, ensayos bioquímicos tales como ensayos de unión, ensayos de polarización de fluorescencia, ensayos de solubilidad, ensayos de plegamiento, ensayos de termoestabilidad, ensayos de estabilidad proteolítica y ensayos de actividad enzimática (véase, en general, Glickman et al., J. Biomolecular Screening, 7(1): 3 - 10 (2002); Salazar et al., Methods. Mol. Biol., 230: 85 - 97 (2003)), así como un intervalo de ensayos basados en células, incluyendo los ensayos basados en transducción de señal, motilidad, unión de células enteras, citometría de flujo y clasificación celular activada por fluorescencia (FACS). Cuando el producto génico es un anticuerpo, o un fragmento del mismo, el fenotipo/función del anticuerpo, o un fragmento del mismo se pueden analizar adicionalmente usando ensayos reconocidos en la materia (p. ej., ensayos de inmunosorción ligada a enzimas (ELISA), ensayo de inmunosorción puntual ligado a enzimas (ELISPOT), detección en gel y detección de fluorescencia de cadenas IgH mutadas, análisis Scatchard, análisis BIACOR, transferencias de tipo Western, análisis en gel de poliacrilamida (PAGE), radioinmunoensayos etc., que pueden determinar la afinidad de unión, la avidéz de unión etc.).

Las células que expresan una proteína de interés codificada por una biblioteca sintética o semisintética como se describe en el presente documento se pueden enriquecer cualquier ensayo reconocido en la técnica, incluyendo, entre otros, procedimientos de acoplamiento de péptidos a micropartículas.

Muchos FACS y sistemas de detección de alto rendimiento están disponibles comercialmente (véase, por ejemplo, Zymark Corp., Hopkinton, Mass.; Air Technical Industries, Mentor, Ohio; Beckman Instruments Inc., Fullerton, Calif.; Precision Systems, Inc., Natick, Mass.) que permite ejecutar estos ensayos en un modo de alto rendimiento. Estos sistemas normalmente automatizan procedimientos completos, incluyendo pipeteo de todas las muestras y reactivos, incubaciones cronológicas de dispensación de líquidos y lecturas finales de la microplaca en el o los detectores adecuados para el ensayo. Estos sistemas configurables proporcionan un elevado rendimiento y un rápido inicio, así como un grado elevado de flexibilidad y personalización. Los fabricantes de dichos sistemas proporcionan protocolos detallados para varios sistemas de alto rendimiento. Por tanto, por ejemplo, Zymark Corp. Proporciona folletos técnicos que describen sistemas de detección selectiva para detectar la modulación de la transcripción génica, unión al ligando y similares. Ensayos de detección de ejemplo que se pueden usar en el contexto del procedimiento de la invención se describen en la publicación de solicitud de patente de EE.UU. 09/0075378 y las publicaciones de solicitud de patente internacional WO 08/103475 y WO 08/103474.

Una vez que se ha obtenido una población de células de interés, las secuencias de ácido nucleico de interés se pueden rescatar y las correspondientes mutaciones secuenciar e identificar. Por ejemplo, el ARNm total o el ADN plasmídico extracromosómico se pueden amplificar mediante coexpresión del antígeno T de SV40 (J. Virol., 62(10): 3738 - 3746 (1988)) y/o se pueden extraer de las células y usar como molde para la reacción en cadena de la

5 polimerasa (PCR) o PCR-transcriptasa inversa (RT) para clonar la secuencia de ácido nucleico modificada usando los cebadores adecuados. Las secuencias de ácido nucleico mutantes se pueden subclonar en un vector y expresar en *E. coli*. Se puede añadir un marcador (p. ej., marcador de His-6) al extremo carboxi para facilitar la purificación de proteínas usando cromatografía. Los datos resultantes se pueden usar para poblar una base de datos que relaciona sustituciones de aminoácidos específicos con cambios en una o más de las proteínas deseadas. Dichas bases de datos se pueden usar después para recombinar mutaciones favorables o para diseñar la biblioteca de polinucleótidos de la siguiente generación con una diversidad dirigida en regiones recién identificadas de interés, por ejemplo secuencias de ácido nucleico que codifica una porción funcional de una proteína.

10 Cuando el producto génico de interés es un anticuerpo, o fragmento del mismo, el ADN se puede extraer mediante PCR usando cebadores sentido específicos de la región líder de la cadena pesada variable ( $V_H$ ) y la región líder de la cadena ligera variable ( $V_L$ ) y cebadores antisentido específicos de isotipo. Como alternativa, el ARN total de poblaciones celulares clasificadas seleccionadas se puede aislar, someter a RT-PCR usando cebadores sentido específicos de la región líder de la cadena pesada variable ( $V_H$ ) y la región líder de la cadena ligera variable ( $V_L$ ) y cebadores antisentido específicos de isotipo. Los clones se pueden secuenciar usando metodologías estándar y las secuencias resultantes se pueden analizar sobre la frecuencia de inserciones y deleciones de nucleótidos, revisión del receptor y selección del gen V.

15 A continuación, las células se pueden volver a cultivar, reinducir la HMS y volver a someter a detección selectiva en un número de ciclos para efectuar mejoras repetitivas en la función deseada. En cualquier punto, la secuencia de ácido nucleico que codifica el producto génico de interés se puede rescatar y/o secuenciar para monitorizar la mutagénesis en curso.

20 La presente invención proporciona adicionalmente un procedimiento para mutar un organismo no humano que tiene un fenotipo deseado que comprende expresar o inducir la expresión de una proteína AID mutante funcional en el organismo, con lo cual la expresión de la proteína AID mutante funcional induce una mutación dentro del ADN cromosómico del organismo. El organismo es, deseablemente, un procarionta (p. ej., una bacteria) o un eucariota. El eucariota puede ser un invertebrado o un vertebrado, pero es, preferentemente, un vertebrado. Más preferentemente, el organismo es un mamífero. Lo más preferentemente, el organismo es un ratón.

25 Los vectores descritos en el presente documento que comprenden una secuencia de ácido nucleico que codifica una proteína AID mutante funcional pueden usarse en el procedimiento citado anteriormente de mutar un organismo. De hecho, dichos vectores se pueden usar para generar ratones que son transgénicos para una proteína AID mutante funcional usando procedimientos de rutina conocidos en la técnica (véase, por ejemplo, *Methods Mol. Med.*, 99: 255 - 67 (2004)). En una realización, un vector que comprende un ácido nucleico que codifica una proteína AID mutante funcional se puede usar para crear un ratón transgénico en el que el gen de la AID endógena no se altera. En otra realización, un vector que comprende un ácido nucleico que codifica una proteína AID mutante funcional se puede usar para crear un ratón transgénico en el que la secuencia de ácido nucleico que codifica la mutante de AID funcional se inserta en el locus de la AID endógena (es decir, cromosómica) para crear un ratón "knock-in", de modo que se evita la expresión de AID endógena. En determinadas realizaciones, el ratón transgénico comprende una proteína AID mutante funcional cuya expresión se puede regular mediante, por ejemplo, promotores específicos de tejido u otros promotores inducibles (p. ej., doxiciclina o tetraciclina, véase, por ejemplo, *Curr. Opin. Biotechnol.*, 13(5): 448 - 52 (2002)). En otra realización, el organismo comprende al menos una secuencia de ácido nucleico que se ha optimizado por codones para la HMS para incrementar el número de motivos de HMS de acuerdo con los procedimientos descritos anteriormente.

30 Cualquiera que sea el procedimiento usado para generar un ratón transgénico, la expresión de la proteína AID mutante funcional induce una mutación dentro del ADN cromosómico del ratón. Una vez que se ha producido mutagénesis en un organismo de acuerdo con el procedimiento de la invención, la célula o células dentro del organismo se selecciona y/o detecta de forma selectiva, preferentemente, para un fenotipo deseado usando procedimientos conocidos en la técnica y descritos en el presente documento.

35 Los procedimientos de la invención descritos en el presente documento también se pueden usar para generar un animal transgénico no humano que produce un anticuerpo dirigido contra un antígeno de interés, o epítipo del mismo. En un aspecto, los procedimientos de la invención se usan, preferentemente, para generar un ratón transgénico que produce anticuerpos monoclonales. Los procedimientos para generar anticuerpos monoclonales se conocen en la técnica y se describen en, por ejemplo, véase, por ejemplo, Köhler y Milstein, *Eur. J. Immunol.*, 5: 511 - 519 (1976), Harlow y Lane (eds.), *Antibodies: A Laboratory Manual*, CSH Press (1988), y C.A. Janeway et al. (eds.), *Immunobiology*, 5ª Ed., Garland Publishing, New York, NY (2001)).

40 El anticuerpo deseado puede ser cualquier anticuerpo natural o sintético como se describe en el presente documento o cualquier fragmento de unión a antígeno de mismo. Además, el anticuerpo puede ser un anticuerpo no humano, un anticuerpo humanizado o un anticuerpo completamente humano. Preferentemente, el anticuerpo es un anticuerpo humanizado. Las formas "humanizadas" de anticuerpos no humanos (p. ej., murinos) son anticuerpos quiméricos que contienen una secuencia mínima derivada de inmunoglobulina no humana. En su mayor parte, los anticuerpos humanizados son inmunoglobulinas humanas (anticuerpo receptor) en las que residuos de la región hipervariable del receptor son sustituidos por residuos de la región hipervariable de una especie no humana

(anticuerpo donante), tal como ratón, rata, conejo o primate no humano que tiene la especificidad, afinidad y capacidad deseadas. En algunos casos, los residuos de la región armazón (FR) de la inmunoglobulina humana se sustituyen por los correspondientes residuos no humanos. Adicionalmente, los anticuerpos humanizados pueden comprender residuos que no se encuentran en el anticuerpo receptor o en el anticuerpo donante. Estas modificaciones se realizan para refinar adicionalmente el rendimiento de los anticuerpos. Un anticuerpo humanizado puede comprender sustancialmente todo de al menos uno, en algunos casos dos, dominios variables en los que todas, o sustancialmente todas, las regiones hipervariables corresponden a las de una inmunoglobulina no humana y todas, o sustancialmente todas, las FR son las de una secuencia de inmunoglobulina humana. El anticuerpo humanizado opcionalmente también comprenderá al menos una porción de una región constante de inmunoglobulina (Fc), normalmente la de una inmunoglobulina humana. Para más detalles, véase Jones et al., *Nature*, 321: 522 - 525 (1986), Reichmann et al., *Nature*, 332: 323 - 329 (1988), y Presta, *Curr. Op. Struct. Biol.*, 2: 593 - 596 (1992). En otra realización, un anticuerpo monoclonal se puede humanizar mediante injerto de CDR de ratón en un armazón de anticuerpo humano sin interferir sustancialmente con la capacidad del anticuerpo para unirse al antígeno. Los procedimientos de preparación de anticuerpos humanizados son generalmente bien conocidos en la técnica y se pueden aplicar fácilmente a los anticuerpos producidos mediante los procedimientos descritos en el presente documento.

En una realización preferida de la invención, los anticuerpos humanizados o completamente humanos se generan usando ratones transgénicos que comprenden una proteína mutante AID funcional que se han criado con una cepa transgénica de ratones en la que la expresión de genes de anticuerpos endógenos de ratón se suprime y sustituye con eficacia con expresión de genes de anticuerpos humanos. Ejemplos de ratones transgénicos en los que los genes de anticuerpos endógenos se sustituyen con eficacia con genes de anticuerpos humanos incluyen, entre otros, HuMAb-Mouse®, the Kirin TC Mouse™, y KM-Mouse® (véase, por ejemplo, Lonberg N. *Nat. Biotechnol.*, 23(9): 1117 - 25 (2005) y Lonberg N. *Handb. Exp. Pharmacol.*, 181: 69 - 97 (2008)).

Los ejemplos siguientes ilustran adicionalmente la invención, pero no deben interpretarse de ningún modo como limitantes de su alcance.

### Ejemplo 1

Este ejemplo demuestra un procedimiento de detección selectiva de mutadores activos de ADN usando un ensayo de papilación.

Los ensayos de papilación se han usado para la detección selectiva de mutantes de *E. Coli* que son defectuosos en algún aspecto de la reparación del ADN (Nghiem et al., *Proc. Natl. Acad. Sci. USA*, 85: 2709 - 17 (1988) y Ruiz et al., *J. Bacteriol.*, 175: 4985 - 89 (1993)).

Para los ensayos de papilación, los ADNc de AID/APOBEC en el plásmido pTrc99<sup>44</sup> se transformaron en la cepa K12 de *Escherichia coli* CC102  $\Delta(lacproB)_{XIII}$  portador del episoma F' *lacZ* *Z' proAB*<sup>+</sup> en el que el gen *lacZ* porta una mutación de sentido erróneo GAG->GGG en el codón 461 (Cupples et al., *Proc. Natl. Acad. Sci. USA*, 86: 5345 - 49 (1989)) y se siembran en placas con agar MacConkey-lactosa (BD Biosciences) suplementado con ampicilina (100 µg/ml) e isopropil β-D-1-tiogalactopiranosido (IPTG; 1 mM). Las placas se incubaron a 37 °C durante 4 días y las papilas se visualizaron tras 3 días.

La frecuencia de la inversión de los transformantes CC102 [pTrc99-AID/APOBEC] en Lac<sup>+</sup> se determinó sembrando cultivos durante la noche hasta la saturación en medio LB suplementado con ampicilina (100 µg/ml) and IPTG (1 mM) en agar M9+0,2 % de lactosa. Las frecuencias de las mutaciones se midieron determinando la mediana del número de células formadoras de colonias que sobrevivieron a la selección por 10<sup>7</sup> células viables sembradas con cada medio determinado de 12 cultivos independientes. La identidad de las mutaciones se determinó mediante secuenciación de secciones relevantes amplificadas por PCR de 5'-AGAATTCCTGAAGTTCAGATGT (SEC ID N° 79) y 5'-GGAATTCGAAACCGCCAAGAC (SEC ID N° 80).

Las células de *E. Coli* que alojan una mutación de sentido erróneo en *lacZ* da lugar a colonias blancas en placas con agar MacConkey-lactosa: dentro de dichas colonias blancas a menudo se puede distinguir un número pequeño de microcolonias rojas o "papilas" (normalmente 0 - 2 por colonia), que refleja inversores Lac<sup>+</sup> que aparecen de forma espontánea. Los clones bacterianos que exhiben una frecuencia elevada de mutación espontánea se pueden identificar en virtud de un mayor número de papilas.

La cepa CC102 de *E. Coli* porta una mutación de sentido erróneo en el codón 461 de *lacZ*, estando el glutamato sustituido por glicina debido a una mutación de transición de A:T a G:C (Cupples et al., *Proc. Natl. Acad. Sci. USA*, 86:5345 - 49 (1989)). Si la expresión de AID en CC102 iba a incrementar la velocidad de desaminación de citosina en el codón 461, esto podría esperarse que aumente la frecuencia de los inversores de Lac<sup>+</sup>. Los transformantes que expresan AID de CC102 dieron un incremento de la frecuencia de la papilación en placas de agar MacConkey-lactosa (FIG. 1a,b). El número de papilas por colonia, como se analizó tras 6 días de incubación, aumentó de 0-2 por colonia a 8-10, que se correlacionaron con un incremento superior al umbral de la frecuencia de los inversores Lac<sup>+</sup> en cultivos durante la noche, como indican las placas con lactosa mínima. El análisis de la secuencia de seis de estos inversores Lac<sup>+</sup> confirmó que, de hecho, se habían producido mediante inversión en el codón 461. Las desaminasas relacionadas con ADI (APOBEC1 (A1) y APOBEC3G (A3G) también desencadenaron un incremento

de la papilación cuando se expresaron en células CC102 (FIG. 1b).

Este ensayo también se usó para determinar si los mutadores activos podían aislarse de una biblioteca de ADNc esplénica total. Se introdujo una biblioteca de ADNc de bazo humano en células CC102 y se realizó detección selectiva de cincuenta mil colonias para determinar la papilación potenciada. Se identificaron treinta y seis posibles candidatos, que se volvieron a analizar sembrando por raspado en placas de agar MacConkey lactosa. Solo se confirmaron dos colonias con una papilación incrementada. El análisis de la secuencia reveló que portaban ADNc distintos derivados de APOBEC3G. La FIG. 1c representa el ARNm de APOBEC3G de longitud completa de tipo silvestre y los dos ADNc de APOBEC3G obtenidos en la detección selectiva de la biblioteca de ADNc de bazo humano, en la que los residuos de nucleótidos se enumeran respecto a inicio del marco de lectura abierto (+1).

10 Este ejemplo demuestra que se puede usar un ensayo de papilación de *E. Coli* a una detección selectiva de alto rendimiento de mutadores activos.

## Ejemplo 2

Este ejemplo demuestra un ensayo para identificar mutantes de AID.

15 Las bibliotecas de mutantes de AID humana de primera y segunda generación se generaron mediante PCR propensa a error usando Taq polimerasa (2,5 U; Biotline) en 1 ng de ADN molde con 1  $\mu$ M de cebadores directos e inversos (5'-ATGGAATTCATGGACAGCCTCTTG (SEC ID N1 81); 5'-CTGAAGCTTTCAAAGTCCCAAAGTA (SEC ID N° 82)), dNTP 250  $\mu$ M, MgCl<sub>2</sub> 10 mM en tampón Taq a 94° C (2 min), seguido de 30 ciclos de 94° C (30 s), 65° C (30 s) y 72° C (1 min). Las bibliotecas de mutantes de AID humana de tercera generación se generaron usando el kit de mutagénesis aleatoria Genemorph II (Stratagene) en 0,1 ng de ADN molde de acuerdo con las instrucciones del fabricante.

20 Los ensayos de papilación se realizaron como se describe en el ejemplo 1, a excepción de que las placas se incubaron a 37°C durante 3-6 días y las papilas se hicieron visibles tras 3 días y sus números aumentaron hasta el día 7. Para el análisis de la expresión inducible con arabinosa, AID se expresó en el plásmido pBAD30 (Guzman et al., *J. Bacteriol.*, 177: 4121 - 30 (1995)),

25 La frecuencia de inversión de transformantes CC102 [pTrc99-AID] a Lac<sup>+</sup> se determinó como se describe en el ejemplo 1, mientras que la mutación a resistencia a rifampicina (Rif<sup>r</sup>) se evaluó siguiendo la transformación en la cepa KL16 de *E. Coli* (Hfr (PO-45) *relA1 spoT1 thi-1*) y el crecimiento de las colonias en presencia de rifampicina (50  $\mu$ g/ml) y arabinosa (0 % - 0,5 %) Las frecuencias de mutación se midieron como se describe en el ejemplo 1 y la identidad de las mutaciones se determinó mediante secuenciación de secciones relevantes amplificadas por PCR de *lacZ* como se describe en el ejemplo 1 o las secciones relevantes amplificadas por PCR de *rpoB* (5'-TTGGCGAAATGGCGGAAAACC-3' (SEC ID N° 83) y 5'-CACCGACGGATACCACCTGCTG-3' (SEC ID N° 84)).

30 Los resultados mostrados en la FIG. 1d and FIG. 2 demuestran que este ensayo puede identificar mutantes de AID. Un total de sesenta mil colonias de cuatro experimentos de mutagénesis de PCR independientes dio 13 clones que exhibieron un incremento de la papilación en placas de agar MacConkey-lactosa. Nueve de estos mutantes se analizaron después en la cepa KL16 de *E. Coli* retransfectada para determinar la frecuencia con la que las colonias resistentes a rifampicina y las nueve exhibieron un incremento de la frecuencia de mutación en el locus *rpoB*.

35 Los ADNc de la AID de dos de las mutantes de primera generación, es decir Mut1 y Mut7. se sometieron a mutagénesis por OCR y se obtuvieron mutantes de segunda generación exhiben una papilación potenciada (FIG. 2). La papilación elevada exhibida por estos mutantes de segunda generación dificultó discernir visualmente cualquier incremento adicional en la papilación. Con el fin de detectar la potenciación adicional de la actividad de mutación en una tercera ronda de mutación/selección, los ADNc que codifican AID Mut1,1 and Mut7.3 se clonaron en un vector de expresión inducible por arabinosa de forma que el número de papilas obtenidos en los transformantes CC102 se pudieran regular variando la concentración de arabinosa en el medio (FIG. 1e). Una tercera generación de mutantes de AID se obtuvo mediante detección selectiva de papilación con niveles bajos de arabinosa (0,02%), algunos de los cuales dieron una frecuencia de mutación casi 400 veces mayor que la AID de tipo silvestre según a frecuencia de mutación a resistencia a rifampicina (FIG. 2).

40 La FIG. 2 representa la dinastía de mutantes de AID seleccionados mediante la detección selectiva de papilación. Los mutantes obtenidos en tres rondas sucesivas de mutagénesis con mutantes obtenidos de experimentos de mutagénesis por PCR individuales se agruparon como familias de hermanos. Las sustituciones de aminoácidos adicionales introducidas en cada ronda de mutagénesis están indicadas con los números debajo de la sustituciones indicadas, lo que da la frecuencia media de la mutación a Rif<sup>r</sup> respecto al vector. \* indica un truncamiento en el extremo C causada por la introducción de un codón de terminación prematuro en el codón indicado. Los mutantes individuales se enumeran de acuerdo con su origen dinástico: por tanto, por ejemplo, Mut7 (K10E/E156G) es la parental de Mut7.1 (K10E/E156G/F115Y).

45 Varios de los mutantes de tercera generación parecían exhibir toxicidad en *E. Coli* juzgado por un tamaño de colonias menor cuando se cultivan en condiciones inductoras; esto se acompañó de un recuento de células viables reducido en cultivos bacterianos cultivados hasta la saturación. Este resultado se demuestra en la FIG. 1f, que

representa los títulos bacterianos en los transformantes CC102 que expresan mutantes de AID diferentes cultivados hasta la saturación en LB/Amp en condiciones de inducción con IPTG respecto a los títulos obtenidos de cultivos que se han cultivado en ausencia de inducción. Esta toxicidad podría haberse debido a que algunos mutantes altamente formadores de papilación dan frecuencias anormalmente bajas de mutación en Rif<sup>r</sup> (p. ej. Mut7.3.4; FIG. 2), posiblemente regulada por disminución la expresión de AID durante el cultivo durante la noche.

Este ejemplo demuestra que el ensayo de papilación de *E. coli* puede identificar proteínas mutantes AID funcionales que exhiben una mejora de 10 veces la actividad en comparación con la proteína AID de tipo silvestre.

### Ejemplo 3

Este ejemplo demuestra que el ensayo de papilación bacteriana puede identificar puntos calientes para los mutantes de AID que tienen una mayor actividad.

La FIG. 3a compara la secuencia principal de la AID humana (SEC ID N° 2) que contiene mutaciones específicas que confieren aumento de actividad, con mutaciones ascendentes en la secuencia de la ADI de pez globo (Fugu). Se deduce que las mutaciones en los residuos con asterisco confieren una mayor actividad mutadora, ya que constituyen la única diferencia entre al menos un par de secuencias de AID que exhiben una diferencia >2 veces de las frecuencias de mutación en *rpoB*. Los residuos que tienen un subrayado doble indican los sitios en los que se han identificado sustituciones en múltiples mutantes independientes pero en presencia de una o más de otras sustituciones. La caja encima o debajo, los residuos con asterisco o con doble subrayado muestra la identidad de las mutaciones de sustitución y la frecuencia con la que cada sustitución se detectó en el total de nueve bibliotecas independientes. Los residuos en los que la correspondiente posición en la AID de fugu también parecen ser un sitio de mutación ascendente seleccionada, según el hecho de que la única mutación identificada en un mutante ascendente de fugu o que la sustitución se identificó (aunque con otros) en múltiples mutantes ascendentes de fugu, se identifican mediante un único subrayado en negrita. Los motivos de coordinación de cinc (HVE y PCYDC (SEC ID N° 86)) y las regiones de contacto de los polinucleótidos sugeridos (SEC ID N° 87)(Cupples et al., Proc. Natl. Acad. Sci. USA, 86: 5345 - 49 (1989); Conticello et al., Nat. Struct. Mol. Biol., 14: 7 - 9 (2007); y Chen et al., Nature, 452: 116 - 119 (2008)), están resaltadas mediante recuadros.

Aparte de las mutaciones prematuras en los codones de terminación identificadas en tres de los mutantes ascendentes de AID (Mut5, Mut1,3, and Mut1,5), el análisis de las secuencias de los diversos mutantes ascendentes de AID reveló una sorprendente preferencia por determinadas sustituciones de aminoácidos. Por ejemplo, las sustituciones K34E, T82I, y E156G (cada una de las cuales es suficiente por sí misma para aumentar la actividad AID) se seleccionaron en experimentos independientes. Estas mutaciones no se encontraron entre las secuencias de 48 clones aleatorios (es decir, no seleccionados) de las bibliotecas generadas por PCR, en las que se observó un amplio espectro de mutaciones sin indicaciones de cualquier punto caliente mayoritario del propio procedimiento de mutagénesis. Por tanto, la repetida identificación de un número pequeño de sustituciones de aminoácidos sugiere que existe un número limitado de sustituciones de aminoácidos únicas en AID que dan un incremento de la papilación.

Aunque en algunos casos (especialmente en la tercera generación), la multiplicidad de mutaciones introducidas en una única ronda previene la identificación no ambigua de las mutaciones responsables del incremento de la papilación, en muchos casos, la mutación ascendente relevante se puede identificar definitivamente porque constituye la única diferencia entre un par de secuencias de AID de papilación diferente o (algo menos definitivamente), porque se obtuvo de forma independiente en múltiples PCR. Las localizaciones de dichas mutaciones ascendentes se representan en la FIG. 3a, donde se ve que mientras que algunas se localizan alrededor del motivo de coordinación de cinc en las proximidades del sitio catalítico probable (V57A; T82I, otras están en una región equivalente a una porción de APOBEC3 que se ha sugerido previamente que está implicada en la unión de polinucleótidos (F115Y; K120R) (Conticello et al., Nat. Struct. Mol. Biol., 14: 7 - 9 (2007); Chen et al., Nature, 452: 116 - 119 (2008); y Holden et al., Nature, 456: 121 - 124 (2008)), varias se agrupan en regiones cuya función es desconocida.

Este ejemplo demuestra un procedimiento para identificar mutantes de AID que tienen una mayor actividad.

### Ejemplo 4

Este ejemplo demuestra que las mutaciones ascendentes identificadas en la detección selectiva de papilación bacteriana aumentan la actividad específica de AID.

Las proteínas de fusión GST-AID se purificaron de transformantes pOPTG-AID de la cepa Rosetta de *E. coli* DE3) pLysS (el vector pOPTG fue un regalo de Perisic, Cambridge, UK) Las células se cultivaron a 37 °C en 2XTY que contiene 100 µg/ml de ampicilina y nCl<sub>2</sub> 100 nM hasta que el cultivo alcanzó una absorbancia de 0,8 at 600 nm cuando se indujo con IPTG 1 mM durante 16 horas a 18 °C y las células sedimentadas se lisaron después mediante una incubación a 30 minutos en hielo en tampón de lisis (Tris 20 mM, pH 7,4, NaCl 100 mM, 0,1% de triton X-100, DTT 5 Mm, 4 µg/ml de RNase A y cóctel de inhibidores de la proteasa sin EDTA (Roche)), seguido de sonicación. Los lisados celulares se aclararon mediante centrifugación 95.000 g; 1 h) and GST-AID se purificó a partir de estos lisados mediante absorción en glutatión Sepharose (Amersham Pharmacia) at 4° C durante 5 horas y elución

seguido de un extenso lavado con tampón de lisis suplementado con gultatión reducido 50 mM que carece Triton X100. Las muestras eluidas se almacenaron a 4°C durante hasta una semana.

La abundancia de la proteína de fusión GST-AID se monitorizó mediante transferencia de tipo Western ( FIG. 3b). La detección selectiva inicial de los extractos sónicos de un gran número de mutantes ascendentes no reveló ninguna que exhibiera un incremento significativo entre el rendimiento fraccionado de proteínas solubles, según el análisis de tipo Western.

La actividad desaminasa de GST-AID (100 - 400 ng) semipurificada se analizó a 37°C en 10 µl del tampón de reacción (Tris 8 mM, pH 8,0, KCl 8 mM, NaCl 10 mM, EDTA 2,5 mM, ditiotreitól 0,2 mM, 5 µg de RNasa A y 0,4 unidades de uracilo-ADN glifosato (NEB)) con 0,5 pmol de oligodesoxirribonucleótido (fluoresceína-5'-ATATGAATAGAATAGAGGGGTGAGCTGGGGTGAGCTGGGGTGAG-3'-biotina (SEC ID N° 85)). Las reacciones se terminaron a tiempos indicados mediante la adición de un volumen igual de pigmento de carga (formamida, EDTA 0,5 mM) y calentamiento a 98 °C durante 3 minutos. Los oligonucleótidos escindidos resultantes se sometieron a electroforesis en 10% de geles de PAGE UREA y la fuorescencia se detectó con un Typhoon Phosphoimager (Molecular Dynamics). La extensión de la desaminación se determinó a partir de las imágenes escaneadas, que expresan el volumen en píxeles de las bandas de producto escindido (tras la resta del fondo) como un porcentaje del volumen en píxeles combinado del producto y las bandas de sustrajo residuales.

Cuando las proteínas de fusión –GST se generaron a partir de mutantes ascendentes humanos Mut1.1 y Mut7.3.6, se puso de manifiesto un claro incremento de la actividad específica según los ensayos de desaminación *in vitro* realizados en un sustrato de oligonucleótidos (FIG. 3b.c). A partir del análisis de las velocidades iniciales, la actividad de desaminación específica de estos mutantes ascendentes había aumentado unas cinco veces en comparación con el tipo silvestre.

Las mutaciones de transición en uno cualquiera de los 11 pares C-G dentro de *rpoB* puede dar lugar a Rif<sup>r</sup>. La distribución de dichas mutaciones entre las colonias Rif<sup>r</sup> se muestra para los mutantes ascendentes de AID Mut8, 1,1, 1,2, 7,3.5 y 7.3.6 en FIG. 3d. La mayor actividad específica no parece acompañarse de ningún cambio grande en la especificidad diana, ya que el análisis de las mutaciones *rpoB* obtenidas usando varias mutantes de AID humana no reveló ninguna diferencia principal en el espectro de mutaciones (FIG. 3d).

Este ejemplo demuestra que las mutaciones identificadas en la detección selectiva de papilación bacteriana aumentan la actividad específica de AID.

### Ejemplo 5

Este ejemplo describe la generación de mutaciones en la AID de pez globo (fugu) que da lugar a una mayor actividad de AID.

Las bibliotecas que contienen mutantes de AID de fugu se generaron usando el kit de mutagénesis aleatoria Genemorph II (Stratagene) en 0,1 ng de ADN molde de acuerdo con las instrucciones del fabricante. Se realizó una detección selectiva de papilación bacteriana de la biblioteca de mutantes de fugu como se describe en el ejemplo 1. La frecuencia de las mutaciones en Rif respecto a los transformantes de únicamente el vector a 18°C o 37 °C se muestra en la FIG. 4b para la cepa K16 de *E. coli* transformada con plásmidos que codifican AID de fugu de tipo silvestre o como se ha indicado. Se construyeron derivados de Mut4.3 y 4.10 en los que la mutación sin sentido en 190 se había invertido, de modo que da un extremo C de tipo silvestre.

La AID de pez globo (que vive a aproximadamente 26 °C) exhibe poca actividad mutadora bacteriana cuando se analiza a 37 °C, mientras que la actividad mutadora se puede detectar a 18°C (Conticello et al., Mol. Biol. Evol., 22: 367 - 377 (2005)). Los mutantes de AID de fugu que dieron una papilación sólida a 37 °C se identificaron en un ensayo de papilación. Como se muestra en la FIG. 4a, todos los mutantes de primera generación aislados alojaron mutaciones de truncamiento en el extremo C (indicado con un "\*\*\*", en el que "\*\*\*a" y "\*\*\*b" indican diferentes sustituciones de un solo nucleótido en el codón 190, lo que produce el codón de terminación prematura), con los seis mutantes obtenidos que alojan 5 mutaciones de truncamiento distintos. También se identificaron mutaciones causantes de la lectura fuera del marco de la región en C terminal, que se denominan "Ins200a" y "Ins200b" en la FIG. 4a para indicar diferentes mutaciones por inserción de un solo nucleótido en el codón 200.

No obstante, diversas sustituciones de aminoácidos podrían conducir a una papilación potenciada en los mutantes de segunda generación (FIG. 4a), produciéndose varias de estas en las posiciones análogas a las mutaciones ascendentes identificadas en la AID humana (FIG. 3a y 4a). Por tanto, la mutación (C88L) responsable de la actividad incrementada de la AID de fugu Mut1.3 se produce en la posición equivalente a la mutación T821 en la AID humana. De un modo similar, los residuos F121, L124 and L128 en la AID de fugu (Cada uno de los cuales es una diana para la mutación en dos o tres mutantes ascendentes de fugu) se localizan todos en una tira de la AID de fugu correspondiente a 115-121 en la AID humana en la que también se obtuvieron mutaciones ascendentes.

Aunque se detectaron truncamientos en C terminal entre el panel de los mutantes ascendentes de la AID humana y dichos truncamientos se habían mostrado previamente que daban una actividad mutadora mayor en *E. coli* (Barreto et al., Mol. Cell, 12: 501 - 508 (2003); Ta et al, Nat. Immunol., 4: 843 - 848 (2003)), todas las mutantes de primera

generación de la AID de fuga a 37 °C portaban truncamientos en el extremo C. Una explicación plausible para esta observación es que las mutaciones en el extremo C subrayan la mayor estabilidad térmica y que las sustituciones de aminoácidos que dan lugar a un incremento de la papilación en los mutantes ascendentes de fuga de segunda generación podrían no ser discernibles a 37°C en ausencia de una mutación de truncamiento en el extremo C. No obstante, esta sí parece una explicación probable. Las sustituciones C88L y L128P tienen ambas una mayor frecuencia de mutación en Rif<sup>r</sup> como se analiza a 18°C en presencia o ausencia de un truncamiento en el extremo C. No obstante, cuando se analiza a 37°C, estas sustituciones de aminoácidos no dieron ningún incremento discernible en la frecuencia de mutación en ausencia del truncamiento en el extremo C (FIG. 4b).

Este ejemplo demuestra que las mutaciones en la AID de pez globo (fugo) que incrementan su actividad son análogas a determinadas mutaciones identificadas en la AID humana.

### Ejemplo 6

Este ejemplo demuestra un procedimiento de potenciación de la diversificación de anticuerpos en células que usan una proteína mutante de AID funcional de la presente invención.

La mutación somática de la IgV se analizó monitorizando la pérdida e IgM de superficie en células DT40 AID<sup>-/-</sup>  $\phi$ V<sup>-/-</sup> slgM<sup>+</sup> (Teng et al., Immunity, 28: 621 - 629 (2008)) que se han transfectado de forma estable con vectores que codifican AID en base a pExpressPuro2 mediante citometría de flujo. Para cada construcción, el porcentaje de células slgM se monitorizó en 12-24 transfectantes independientes que se habían expandido en selección (0,25 µg/ml puomicina) durante 3 semanas antes de la citometría de flujo.

Las mutaciones en la región IgVλ se caracterizaron mediante secuenciación del ADN genómico que se amplificó mediante PCR de 100.000 sin clasificar o de equivalentes celulares clasificados ((GFP<sup>+</sup>; slgM<sup>+</sup>)(Sale et al., Nature, 412: 921 - 926 (2001)).

Para analizar el cambio de clase se analizó la expresión de superficie de IgG1 mediante citometría de flujo en los linfocitos B que se habían purificado en ratones AID<sup>-/-</sup> y cultivado en presencia de LPS+IL4 (48 h) tras una infección de 24 horas con retrovirus que codifican AID como se ha descrito anteriormente (Di Noia, J. Exp. Med., 204: 3209 - 3219 (2007)). Para facilitar una disminución en la extensión de la sobreexpresión de AID en los linfocitos B transducidos se usó un vector retroviral con una secuencia Kozak mutada como se ha descrito (McBride et al., J. Exp. Med., 205: 2199 - 2206 (2008)). La abundancia de AID en los extractos preparados calentando las células 10<sup>6</sup> en 50 µl de tampón de muestra-DSD reductor se monitorizó tras el SDS/PAGE mediante análisis de transferencia de tipo Western usando antisuero anti-AID de conejo (Abcam); se detectó GFP usando antisuero anti-GFP de cabra conjugado con HRP (Abcam)

Los mutantes 3 (T82I), 8 (K34E, K160E) y 7,3 (K10E, E156G, T82I) se expresaron en un línea celular DT40 de pollo deficiente en AID en la que la mutación somática de la IgV se puede deducir de la frecuencia de la generación de variantes de pérdida de slgM (Arakawa et al., PLoS Biol., 2: E179 (2004)). Tanto Mut3 como Mut7.3 parecieron dar una mutación somática significativamente potenciada según este ensayo de pérdida de slgM (FIG. 5a). Adicionalmente, el análisis de frecuencia reveló que después de un mes de expansión clonal, las células que expresan estas AID mutantes portaban una carga de mutación más alta en el gen IgGλ que las células control que expresan la enzima de tipo silvestre (FIG. 5b). No solo una proporción mayor de secuencias portan mutaciones sino que las que sí portaban mutaciones también eran portadoras de una mayor carga de mutación. Este efecto es particularmente marcado cuando se tiene en cuenta el hecho de que la AID mutante se expresa a una abundancia menor que su homólogo de tipo silvestre en estos transfectantes. Por el contrario, el mutante 8 no dio una mutación somática potenciada, lo que indica que es probable que las sustituciones K34E y/o K160E disminuyan aspectos de la función de la AID en los linfocitos B. Es interesante el hecho de que el polipéptido Mut8 se encuentra a una abundancia mucho más alta en los transfectantes DT40 que son los polipéptidos Mut3 o 7.3. Esto es consistente con las observaciones de otros trabajos (p. ej., Conticello et al., Mol. Cell, 31: 474 - 484 (2008)) de que los mutantes de AID que exhiben una actividad comprometida en la diversificación de los anticuerpos/mutación genómica en las células DT40 tienden a expresarse a una abundancia más alta sin ninguna alteración evidente en la localización intracelular. Una posible explicación de estas diferencias en los niveles de expresión es que, en las transfectantes celulares, existe una selección frente a las células que expresan niveles altos de proteínas AID que son activas en la mutación cromosómica.

Un ensayo basado en la transducción retroviral de las enzimas mutantes en linfocitos B de ratones deficientes en AID se usó para analizar la actividad de la AID mutante en la recombinación por cambio de clase. Con el fin de limitar el grado de sobreexpresión de AID que podría, por otro lado, saturar el ensayo de cambio, el ensayo se realizó usando el virus pMX-Ig convencional así como una variante en la que la AID transducida se expresa a niveles más bajos mediante mutación de la secuencia Kozak (McBride et al., J. Exp. Med., 2005: 2585 - 2594 (2008)). Como se indica en la FIG. 5c, que representa gráficos de citometría de flujo representativos del cambio a IgG1, en la que 'mK' indica cuando la transducción se realizó usando vectores con una secuencia Kozak mutada, Mut7.3 fue más eficaz en la estimulación de la recombinación por cambio de clase que el homólogo de tipo silvestre aunque expresado a niveles menores.

Este ejemplo demuestra el uso de una proteína AID mutante funcional par apotenciar la diversificación de

anticuerpos de acuerdo con el procedimiento de la invención.

### Ejemplo 7

Este ejemplo demuestra que los mutantes de AID aumentan las translocaciones cromosómicas.

5 Un ensayo basado en PCR (Janz et al., Proc. Natl. Acad. Sci. USA., 90: 7361 - 7365 (1993)) se usó para detectar translocaciones c-myc/IgH en linfocitos B. Los linfocitos B de ratones deficientes en AID se transdujeron con retrovirus que expresan AID y se cultivaron en medio que contiene LPS (20 µg/ml) and IL4 (50 ng/ml) como se describe para los ensayos de cambio de clase en el ejemplo 6, sembrando a  $8 \times 10^5$  células/ml en placas de 6 pocillos. El ADN genómico de  $2 \times 10^5$  células que se habían preparado usando DirectPCR (Viatch) de células GFP<sup>+</sup> clasificadas 36 h después de la transducción se sometió a dos rondas de PCR anidada con un sistema de PCR Expand Long Template PCR (Roche), seguido de transferencia de tipo Southern para amplificar y detectar translocaciones der12 c-myc/Igµ y der15 c-myc/Igµ y los productos específicos como se ha descrito (Ramiro et al., Nature, 440: 105 - 109 (2006)). La FIG. 6 (arriba) representa el esquema de la translocación recíproca entre los loci c-myc y IgH e indica los cebadores usados para amplificación por PCR (flechas) y las sondas (P) usadas para la hibridación de transferencia de tipo Southern.

15 Los linfocitos B de ratones deficientes en AID se transdujeron retroviralmente para la expresión de AID y se cultivaron durante 1-2 días *in vitro*. La AID Mut7.3 dio lugar a una proporción significativamente mayor de cultivos que contenían translocaciones c-myc/IgH que la enzima de tipo silvestre (FIG. 6 (parte inferior)).

Este ejemplo demuestra un procedimiento de incremento de las translocaciones cromosómicas usando una proteína mutante AID funcional.

### 20 Ejemplo 8

Este ejemplo demuestra que las secuencias de ácido nucleico que codifican mutantes de AID con mayor actividad están más cerca de las secuencias de ácido nucleico de las desaminasas APOBEC3 que la AID de tipo silvestre.

25 Se realizó una alineación Web LOGO (Crooks et al., Genome Research, 14: 1188 - 1190 (2004)) (FIG. 7), que representa la conservación de aminoácidos alrededor de los sitios principales de la mutación ascendente e AID y las regiones homólogas en los dominios Z1, Z2 y Z3 de las APOBEC3 de mamífero (vaca, oveja, cerdo, perro, pecarí, caballo, gato, perro, ratón, rata, ser humano y macaco: los números de acceso de las secuencias se proporcionan en la FIG. 8). Cualquier secuencia con una identidad de aminoácidos superior al 90% con cualquier otra secuencia se desechó de la generación de los perfiles LOGO. Las mutaciones ascendentes de AID se muestran en el recuadro encima de los residuos numerados. Las flechas en la parte inferior de la alineación destacan los residuos homólogos en las APOBEC3.

30 La FIG. 7 ilustra que las proteínas de la familia APOBEC3 están en evolución rápida y están presentes en múltiples copias de animales superiores: sus dominios de coordinación de cinc se pueden clasificar mediante homología de secuencia en uno de tres subagrupamientos (Z1, Z2 y Z3) (Conticello et al., Mol. Biol. Evol., 22: 367 - 377 (2005)). La alineación de las secuencias de AID con las de las APOBEC3 reveló que la mayoría de las mutaciones ascendentes seleccionadas con frecuencia en la AID humana sirvió para llevar la secuencia de AID más cerca a la de sus parientes APOBEC3 (FIG. 7). De hecho, mientras que a mutación ascendente de AID en los sustitutos F115, el aminoácido preferido en la posición correspondiente en los dominios Z2 de APOBEC3 (Y), las mutaciones ascendentes en K34, T82 and E156 sustituyen todas al aminoácido preferido en la correspondiente posición en los dominios Z1 de APOBEC3. Es interesante el hecho de que estos dominios Z1 que se encontró que eran los más catalíticamente activos de los dominios de APOBEC3 (LaRue et al., J. Virol., 83: 494 - 497 (2009)). Por tanto, parece que mientras que la desaminación de la actividad de AID puede aumentarse artificialmente mediante mutación ascendente específica, dichas mutaciones ascendentes pueden haberse contraseleccionado durante la evolución de AID pero no durante la evolución de APOBEC3

45 Este ejemplo demuestra que las secuencias de ácido nucleico que codifican mutantes de AID con mayor actividad están más cerca de las secuencias de ácido nucleico de las desaminasas APOBEC3 que la AID de tipo silvestre.

### Ejemplo 9

Este ejemplo compara mutaciones ascendentes de ADI humana (SEC ID N° 2) y de pez globo (fugu) (SEC ID N° 13).

50 Las secuencias primarias de AID humana y fugu se alinean usando ClustalW2 (p. ej., Larkin et al., Bioinformatics, 23: 2947 - 2948 (2007)) (FIG. 9). Las mutaciones ascendentes de AID human están indicadas por un asterisco o doble subrayado como se describe en el Ejemplo 3 (FIG. 3a). Las mutaciones ascendentes de AID de fugu están indicadas por un ("^") que se ha identificado porque constituyen la única mutación en un mutante ascendente de fugo o porque el residuo se mutó en múltiples mutantes ascendentes de fugu. La naturaleza de las sustituciones está indicada en recuadros por encima o por debajo de los residuos destacados como en la FIG. 3a. Los motivos de coordinación de cinc (HVE y PCYDC) y las regiones del contacto de polinucleótidos sugerido (FCEDRKA) están en un recuadro.

Este ejemplo comparó las mutaciones ascendentes de AID humana y de pez globo (fugu).

### Ejemplo 10

Este ejemplo describe un procedimiento de generación de un mutante de AID funcional que comprende sustituir una secuencia de aminoácidos de una proteína AID de tipo silvestre con una secuencia de aminoácidos correspondiente de un homólogo de AID.

Los mutantes de AID humana en los que los residuos de aminoácidos 115-123 se han sustituido por regiones equivalentes de APOBEC3C (AID/3C), APOBEC3F (AID/3F), and APOBEC3G (AID/3G) se clonaron en un plásmido de expresión bacteriana. La actividad mutadora de estas secuencias de AID modificadas se analizó mediante monitorización de la crecencia con la que dieron colonias resistentes a rifampicina tras la transformación en *E. coli*. Específicamente, la cepa KL16 de *E. coli* [Hfr (PO-45) *relA1 spoT1 thi-1*] transformada con los plásmidos pTrc99/AID se cultivó durante la noche hasta saturación en medio LB suplementado con ampicilina (100 µg ml<sup>-1</sup>) e isopropil β-D-1-tiogalactopiranosido (IPTG; 1 mM), y se sembró en agar LB con baja concentración de sales que contiene ampicilina (100 µg ml<sup>-1</sup>) y rifampicina (50 µg ml<sup>-1</sup>). Las frecuencias de las mutaciones se midieron determinando la mediana del número de células formadoras de colonias que sobrevivieron a la selección por 10<sup>7</sup> células viables sembradas con cada medio determinado de 12 cultivos independientes. La identidad de las mutaciones se determinó mediante secuenciación de la sección relevante de *rpoB* (normalmente de 25 - 200 clones individuales) tras la amplificación por PCR usando los oligonucleótidos 5'-TTGGCGAAATGGCGGAAAACC ° 88) y 5'-CACCGACGGATACCACCTGCTG-3' (SEC ID N° 89). Mientras que las proteínas AID/3C and AID/3F retenían una buena actividad mutadora, AID/3G dio colonias resistentes a rifampicina a una frecuencia indistinguible sobre el fondo. La resistencia a rifampicina es conferida por una de un número limitado de mutaciones en *rpoB* con la naturaleza de las mutaciones obtenidas y da información sobre la especificidad de la diana de la desaminasa (Harris et al., *Mol. Cell.*, 10: 1247 - 1253 (2002)). La AID de tipo silvestre prefiere desaminar los residuos C en la posición 1576 de *rpoB* (C1576), que tiene un residuo de purina (G) flanqueante en 5'. Por el contrario, las variantes de AID en las que los residuos 115-123 se han sustituido por las correspondientes regiones de APOBEC3C/F/G mostraron preferencias (como las propias APOBEC3) por las pirimidinas en la posición -1. Por tanto, AID/3C and AID/3F exhibieron un cambio en el espectro de las mutaciones *rpoB* a favor de las dianas con una a 5'-T (C1535, C1565 y C1592), mientras que los transformantes AID/3G casi únicamente estaban dirigidos a C1691, que tiene un C en 5.

Estos resultados de este ejemplo demuestran que la sustitución de los residuos de aminoácidos 115-123 de la AID humana con las correspondientes secuencias de las proteínas APOBEC3 altera la actividad específica de AID.

### Ejemplo 11

Este ejemplo describe un procedimiento de generación de un mutante de AID funcional que comprende sustituir una secuencia de aminoácidos de una proteína AID de tipo silvestre con una secuencia de aminoácidos correspondiente de un homólogo de AID.

Aunque el ejemplo 10 demuestra que la actividad mutadora de AID/3G es suficiente para dar un cambio en la distribución de las mutaciones *rpoB* observadas en *E. coli* resistente a rifampicina, la actividad mutadora del mutante AID/3G es considerablemente menor que la de la AID de tipo silvestre, ya que no da una frecuencia total de mutación en resistencia a rifampicina que es superior al nivel de fondo. Para mejorar la actividad mutadora del mutante AID/3G se generaron dos mutantes ascendentes AID/3G (designados /3G and AID2/3G) en los que se introdujeron tres aminoácidos adicionales (es decir, AID1: K10E, T82I, E156G; AID2: K34E, E156G, R157T) en estas proteínas. Estos mutantes ascendentes AID/3G parecían conservar la preferencia de la proteína AID/3G por un residuo C flanqueante en 5' determinado mediante el espectro de mutación *rpoB*.

También se generaron variantes de AID (designadas AID\*, AID\*/3F, AID1\*/3G, etc.) en las que la porción en C-terminal de AID (que incluye su secuencia de exportación nuclear) se había delecionado. El truncamiento en C-terminal no da un efecto detectable sobre el sitio diana mutacional de AID preferente en ensayos de mutación bacteriana.

Para analizar la especificidad bioquímica por la diana de las AID mutantes anteriores con mayor detalle, las diversas enzimas de AID se purificaron parcialmente a partir de extractos de *E. coli* como proteínas de fusión-GST recombinantes y se usaron para desaminar el ADN diana de lacZ monocatenario en el contexto del ensayo dúplex M13 (Bebenek and Kunkel, *Methods Enzymol.*, 262: 217 - 232 (1995); Pham et al., *Nature*, 424: 103 - 107 (2003)). En este ensayo, la GST-AID recombinante se incubó con ADN de M13lacZ dúplex con huecos, que después se transforma en *E. coli*.

El análisis de 30 - 50 clones de M13lacZ mutados en cada experimento dio bases de datos de 471-685 mutaciones, todas las cuales fueron transiciones en los pares C:G. En el caso de AID1, el 74 % de las mutaciones en C estaba en los sitios flanqueados por una purina en 5'. Por el contrario, los mutantes de AID portadores de segmentos transplantados de las proteínas APOBEC3 mostraron un cambio hacia una preferencia por una pirimidina flanqueante que era especialmente marcada en el caso de las proteínas AID/3C and AID/3G (85% y 77% de pirimidina, respectivamente). Este cambio en la preferencia del nucleótido flanqueante se acompañó de un cambio en la distribución de las butacones a lo largo de lacZ. Dado que para la mayoría de las variantes de AID, las

secuencias mutadas portaban una media de 10-16 mutaciones de transición sobre la tira de 475 nucleótidos del sustrato monocatenario analizado, las mutaciones observadas reflejaron considerablemente la preferencia intrínseca del proceso mutacional sin una desviación extensa en virtud de la selección para la inactivación de lacZ.

5 Los resultados de este ejemplo confirman que la sustitución de los residuos de aminoácidos 115-123 de la AID humana con las correspondientes secuencias de las proteínas APOBEC3 altera la actividad específica de AID.

### Ejemplo 12

Este ejemplo demuestra que las proteínas AID mutantes exhiben espectros de mutación alterados en linfocitos B.

10 Con el fin de determinar si cambiar la especificidad catalítica de AID tiene como resultado una alteración en la distribución de las sustituciones de aminoácidos introducidas durante la HMS en linfocitos B, las AID mutantes descritas en los ejemplos 10 y 11 se expresaron en la línea de linfocitos B DT40 de pollo con delección de  $\psi$ V deficientes en AID. En la línea de linfocitos B DT40, las mutaciones están en gran medida restringidas a sustituciones de aminoácidos en los pares C:G con poca contribución de la hipermutación desencadenada por polimerasa  $\eta$  (Arakawa et al PLoS Biol., 2: E179 (2004); Di Noia and Neuberger, Nature, 419: 43 - 48 (2002); Sale et al., Nature, 412: 921 - 926 (2001)), lo que significa que las mutaciones en C:G pueden deberse en gran medida a los efectos directos de la AID en lugar de ser posiblemente una consecuencia de una segunda fase de la creación de mutación. La frecuencia de HMS en la IgV se puede deducir de la frecuencia de la generación de variantes de pérdida de sIgM Buerstedde et al., EMBO J., 9: 921 - 927 (1990); Sale et al., Nature, 412: 921 - 926 (2001)). Este ensayo demostró que tanto AID/3C como AID/3F son eficaces en la HMS. De hecho, AID/3C es más potente que la enzima de tipo silvestre, especialmente cuando se tiene en cuenta la menor abundancia del polipéptido AID/3C en los extractos de linfocitos B. La baja abundancia de AID/3C fue evidente en múltiples transfectantes independientes. La razón de esta baja expresión puede reflejar la citotoxicidad de la excesiva actividad de desaminasa de ADN.

15 En contraste con los mutantes AID/3C y AID/3F, el mutante AID1/3G si solo una frecuencia muy baja de variantes de pérdida de sIgM. No obstante, esta frecuencia se potenció considerablemente delecionando la porción C-terminal de la AID.

25 Para caracterizar el espectro de hipermutación del gen IgV en los transfectantes de linfocitos B de DT40 que expresan las diversas proteínas AID modificadas, el segmento IgV $\lambda$  de múltiples transfectantes independientes de cada construcción de expresión se amplificó por PCR y se secuenció después de ocho semanas de expansión clonal. Los resultados revelaron que las modificaciones del sitio activo de la AID tuvo como resultado una alteración sustancial en el espectro de la mutación IgV $\lambda$ . Por tanto, AID/3C y AID1\*/3G estaban dirigidos principalmente a los residuos en C con un residuo de pirimidina flanqueante en 5' (68 % and 75 %, respectivamente, en contraste con la enzima de tipo silvestre en la que solo el 19% de las mutaciones están dirigidas a los residuos C con una pirimidina flanqueante en 5'. Este cambio significativo en el espectro de mutaciones es evidente tanto en los conjuntos de datos compuestos como en cada uno de los clones independientes. En contraste, AID/3F mantuvo la preferencia de la enzima parental por un residuo de purina flanqueante pero, como se encuentra en el ensayo *in vitro* en el sustrato de lacZ dúplex con huevos (Ejemplo 11), existe un cambio hacia una preferencia por una guanina flanqueante en lugar de adenina.

30 El cambio en la focalización mutacional según la naturaleza del nucleótido flanqueante en 5' se correlacionó ampliamente con un espectro mutacional alterado, determinado por la distribución de sustituciones de nucleótidos a lo largo del segmento IgV $\lambda$ . Por tanto, por ejemplo, los puntos calientes de la mutación IgV $\lambda$  se encontraron en distintas localizaciones cuando se compara la AID de tipo silvestre con AID1\*/3G. Con la AID de tipo silvestre (así como en el mutante ascendente AID1), fueron evidentes los puntos calientes dentro de CDR1, hacia el lado 5' de la CDR2 y también dentro de CDR3, adecuándose principalmente estos puntos calientes a una consenso WRC como se ha observado anteriormente (Arakawa et al., PLoS Biol., 2: E179 (2004); Sale et al., Nature, 412: 921 - 926 (2001); Saribasak et al., J. Immunol., 176: 365 - 371 (2006); Wang et al., Nat Struct Mol Biol., 16: 769 - 76 (2009)). En contraste con ello, las mutaciones IgV $\lambda$  obtenidas usando AID1\*/3G mostraron menor agrupación en CDR1 y CDR3, con un enfoque en los puntos calientes con una pirimidina flanqueante en 5' y que se localizan en las regiones (FR1 y FR3) que están relativamente libres en la enzima de tipo silvestre.

35 Los resultados de este ejemplo demuestran que el cambio del sitio activo de la AID modifica el espectro de mutaciones que se obtiene mediante la desaminación del ADN *in vitro* y mediante la hipermutación de anticuerpos en transfectantes de linfocitos B.

### Ejemplo 13

Este ejemplo describe la identificación de "puntos calientes" de mutantes de AID.

40 Aunque las modificaciones de los sitios activos en AID/3C and AID1\*/3G tuvieron como resultado un cambio desde una preferencia por una purina flanqueante en 5' a una pirimidina flanqueante en 5' en ensayos de mutación tanto *in vivo* (DT40 IgV $\lambda$ ) como *in vitro* (dúplex con huevos lacZ), la naturaleza del cambio en los dos ensayos no fue equivalente. Por tanto, para AID/3G, aunque T es la pirimidina flanqueante de elección en el espectro de mutación IgV $\lambda$  en DT40, se prefiere una C flanqueante de en el ensayo *in vitro*. Esta discrepancia se debe sustancialmente al

efecto de desviación de unos pocos puntos calientes principales en el espectro de mutación *IgVλ* en DT40, lo que sugiere que algún aspecto de la hipermutación en los linfocitos B podría dar lugar a la creación de puntos calientes dominantes que no se recapitulan en el ensayo dúplex con huecos.

5 Para confirmar esto, el ensayo de mutación dúplex con huecos se realizó en una secuencia diana *IgVλ* (en lugar de *lacZ*) y el espectro de mutación *in vitro* resultante se comparó con el observado en la hebra de ADN de *IgVλ* (no transcrita) equivalente en linfocitos B de DT40. Se observaron diferencias significativas en el enfoque mutacional. Estas diferencias fueron evidentes de un modo similar si las secuencias altamente mutadas se excluyeron de las bases de datos usadas para deducir los patrones del enfoque mutacional *in vitro*.

10 Para encontrar si las diferencias en los efectos del enfoque refleja el hecho de que la mutación *in vivo* probablemente se produzca en la transcripción de un ADN bicatenario, mientras que el ensayo dúplex con huecos usa un ADN monocatenario diana, se usó enfoque mutacional, usando un ensayo descrito por Bransteitter et al., J. Biol. Chem., 279: 51612 - 21(2004). Este ensayo de enfoque mutacional implica incubar la AID recombinante con ADN bicatenario al mismo tiempo que el gen diana (*lacZ*) dentro del sustrato se está transcribiendo desde un promotor de T7 polimerasa ligado. En este ensayo, AID1\*/3G dirigió claramente de la AID de tipo silvestre, prefiriendo todavía una primidina en 5' y, especialmente, 5'-C en lugar de la 5'-T observada en los linfocitos B DT40. Con el fin de evaluar el enfoque mutacional dentro de un sustrato *IgVλ* en un ensayo acoplado a transcripción *in vitro*, el ensayo ligado a T7 se modificó para crear un sustrato en el que mutaciones no seleccionadas en segmentos cortos de *IgVλ* podría puntuarse en clones que han sufrido inactivación mutacional de un gen indicador de GFP estrechamente unida. No obstante, en dichos ensayos se encontró que, como en el ensayo dúplex con huecos, la dominancia relativa de los puntos calientes principales en *IgVλ*, posiciones 141 (AID de tipo silvestre) o 252 (con AID1\*/30), observados durante la hipermutación en células DT40 no se recapturó. De hecho, el enfoque mutacional en el ensayo ligado a transcripción pareció ser más similar al obtenido en el ensayo dúplex con huecos que al patrón de enfoque mutacional observado en los linfocitos B DT40. Por tanto, ni el ensayo *in vitro* recapturó completamente el patrón de dominancia del punto caliente de *IgV* observado en linfocitos B.

25 Los resultados de este ejemplo demuestran que los linfocitos B que expresan proteínas AID modificadas dan un uso de puntos calientes alterado.

#### Ejemplo 14

Este ejemplo describe los efectos de transferir la mutación Mut7.3 a AID canina y AID humana.

30 La función de AID en las células HEK293-c18 se midió secuenciando un molde de anticuerpo coexpresado. Las células se contransfectaron con tres vectores episomales que contenían marcadores de selección únicos, uno que expresa la cadena pesada de un anticuerpo con selección con puromicina, uno que expresa la cadena ligera de un anticuerpo con selección con higromicina y uno que expresa AID con selección con blasticidina. Después de la transfección, las células siempre se cultivaron con puromicina e higromicina pero se trataron de forma diferencial con blasticidina. Para las células "pulsadas" con AID, no se añadió blasticidina al cultivo y se repitió la transfección transitoria de un vector de AID cada semana del experimento. Para las células con AID "estable" se añadió blasticidina al medio de cultivo y para las células Estables + pulsadas" se cultivaron con blasticidina además de la transfección con un vector de AID cada semana. En estos experimentos se examinaron tres mutantes de AID diferentes: AID canina ("MutE"), AID canina que contiene Mut7.3 ("Mut 7,3 E") y AID humana que contiene Mut7.3 ("humana 7.3") (SEC ID N°: 88 - 93 y las Figuras 10a y 10b) y dos construcciones de vectores diferentes se analizaron para determinar la expresión de AID (es decir, un vector IRES y un vector pEpi). Con el vector IRES, la expresión tanto de AID como de blasticidina estaba controlada por el mismo promotor con un elemento IRES entre los genes. En los vectores pEpi la expresión de blasticidina estaba controlada por un promotor distinto.

45 Después de aproximadamente un mes de cultivo, las regiones variables de la cadena pesada se recuperaron mediante PCR para secuenciación. Se secuenciaron noventa y cuatro moldes para cada experimento de transfección celular separada con una media de 88 secuencias completas retornadas por experimento. Los cromatogramas de secuenciación se examinaron para verificar la calidad de las mutaciones observadas y la frecuencia de la mutación se calculó dividiendo el número de mutaciones por el número total de nucleótidos secuenciados, y, después, dividiendo por el número de días en cultivo. El tiempo de duplicación de las células HEK293-c18 es de aproximadamente 24 horas, por lo que los días en cultivo se usaron para normalizar la tasa de mutación por generación.

50 No se observaron diferencias significativas en la frecuencia de mutaciones entre los grupos pulsado, estable o estable más pulsado para cada vector de AID. Además, no había diferencias significativas entre los vectores IRES y pEpi para MutE AID, ni diferencias significativas entre Mut 7.3E y 7.3 humana. No obstante, la diferencia en la frecuencia de mutación para Mut 7.3E en pEpi fue estadísticamente significativa ( $p=0,0003$ ) de Mut 7.3E en IRES

55 Este ejemplo demuestra que Mut7.3 se puede traducir en AID canina y AID humana.

El uso de los términos "un" y "una/uno" y "el/la" y referencias similares en el contexto de la descripción de la invención (especialmente en el contexto de las reivindicaciones siguientes) se deberá interpretar que abarca tanto el singular como el plural, a menos que se indique lo contrario en el presente documento o esté claramente

5        contraindicado por el contexto. Los términos "que comprende", "que tiene", "incluyendo" y "que contiene" se deben interpretar como términos abiertos (es decir, que significan incluyendo, aunque sin limitaciones", a menos que se indique lo contrario. Los intervalos de valores citados en el presente documento están destinados simplemente a servir como método abreviado de hacer referencia individualmente a cada valor por separado que entra dentro del intervalo, a menos que se indique lo contrario en el presente documento, y cada valor distinto se incorpora en la memoria descriptiva como si se hubiera citado individualmente en el presente documento. Todos los procedimientos descritos en el presente documento se pueden realizar en cualquier orden adecuado, a menos que se indique lo contrario en el presente documento o que, de otro modo, claramente se contradiga en el contexto. El uso de cualquiera y todos los ejemplos, o términos ilustrativos (p. ej., "tal como") proporcionados en el presente documento está destinado simplemente a iluminar mejor la invención y no impone una limitación del alcance de la invención a menos que se reivindique de otro modo. Ninguna expresión en la memoria descriptiva se debe interpretar como indicativa de ningún elemento no reivindicado esencial en la práctica de la invención.

10       Realizaciones preferidas de la presente invención se describen en el presente documento, incluido el mejor modo conocido por los inventores para llevar a cabo la invención. Variaciones de estas realizaciones preferidas pueden ser evidentes para los expertos en la técnica a la luz de la lectura de la descripción anterior. Los inventores esperan que los expertos usen dichas variaciones según sea adecuado y los inventores pretenden que la invención se ponga en práctica de otro modo distinto a lo descrito específicamente en el presente documento. De acuerdo con esto, la presente invención incluye todas las modificaciones y equivalentes de la materia objeto citada en las reivindicaciones adjuntas según lo permitido por la legislación aplicable. Además, cualquier combinación de los elementos descritos anteriormente de todas las posibles variaciones de la misma están comprendidos en de la invención, a menos que se indique lo contrario en el presente documento o el contexto lo contradiga claramente.

**Listado de secuencias**

- 25        <110> MEDICAL RESEARCH COUNCIL
- <120> MUTANTES DE LA CITIDINA DESAMINASA INDUCIDA POR ACTIVACIÓN (AID) Y PROCEDIMIENTOS DE USO
- <130> P040484WO
- 30        <140> PCT/IB2010/000958
- <141> 2010 - 04 - 05
- <150> US 61/166.349
- 35        <151> 2009 - 04 - 03
- <160> 93
- <170> PatentIn versión 3,5
- 40        <210> 1
- <211> 198
- <212> PRT
- <213> Homo sapiens
- 45        <400> 1

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160  
 Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

5 <210> 2  
 <211> 198  
 <212> PRT  
 <213> Homo sapiens

10 <400> 2

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- 5 <210> 3
- <211> 198
- <212> PRT
- <213> Canis lupus

10 <400> 3

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Lys Gln Arg Lys Phe Leu Tyr His Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
 35 40 45

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Tyr Pro Asn Leu Ser Leu Arg Ile Phe Ala Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Lys Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- <210> 4
- <211> 198
- 5 <212> PRT
- <213> Mus musculus
- <400> 4



ES 2 498 765 T3

Met Asp Ser Leu Leu Met Lys Gln Lys Lys Phe Leu Tyr His Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
 35 40 45

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Glu  
 85 90 95

Phe Leu Arg Trp Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Gly Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Thr Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Ile Leu Gly Leu  
 195

<210> 6  
 <211> 199  
 <212> PRT  
 <213> Bos taurus

<400> 6

Met Asp Ser Leu Leu Lys Lys Gln Arg Gln Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

5

10

ES 2 498 765 T3

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Pro Thr Ser Phe Ser Leu Asp Phe Gly His  
 35 40 45

Leu Arg Asn Lys Ala Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Tyr Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Asp Lys Glu Arg Lys Ala Glu Pro Glu Gly Leu Arg  
 115 120 125

Arg Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp  
 130 135 140

Tyr Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe  
 145 150 155 160

Lys Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln  
 165 170 175

Leu Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp  
 180 185 190

Ala Phe Arg Thr Leu Gly Leu  
 195

<210> 7  
 <211> 198  
 <212> PRT  
 <213> Gallus gallus

5

<400> 7

Met Asp Ser Leu Leu Met Lys Arg Lys Leu Phe Leu Tyr Asn Phe Lys  
 1 5 10 15

Asn Leu Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Cys Ser Leu Asp Phe Gly Tyr  
 35 40 45

10

ES 2 498 765 T3

Leu Arg Asn Lys Met Gly Cys His Val Glu Val Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Ala Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Ile Thr Trp  
 65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Ala Tyr Pro Asn Leu Thr Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Ala Gln Ile Ala Ile Met Thr Phe Lys Asp Phe  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Lys Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val His Leu Ser Arg Lys Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Lys Thr Leu Gly Leu  
 195

<210> 8  
 <211> 198  
 5 <212> PRT  
 <213> Sus scrofa

<400> 8

Met Asp Ser Leu Leu Met Lys Gln Arg Gln Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
 35 40 45

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

10

ES 2 498 765 T3

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asn  
85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
100 105 110

Leu Tyr Phe Cys Asp Gly Tyr Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Arg Ser Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Thr Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

- <210> 9
- <211> 198
- <212> PRT
- <213> Pan troglodytes

5

<400> 9

Met Asp Ser Leu Leu Met Asn Arg Lys Lys Phe Leu Tyr Gln Phe Lys  
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

10

ES 2 498 765 T3

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 10  
 <211> 198  
 <212> PRT  
 <213> Macaca mulatta

<400> 10

5

10

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr His Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
 35 40 45

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

ES 2 498 765 T3

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- <210> 11
- <211> 239
- 5 <212> PRT
- <213> Equus caballus
- <400> 11

Met Ser Glu Val Leu Thr Thr Arg Pro Ala Val Pro Asp Ala Val Pro  
 1 5 10 15

Pro Val Pro Asn Leu Leu Ile His Thr Ala Gln Glu Glu Asn Asp Leu  
 20 25 30

Cys Ser Phe Ile Phe Leu Pro Pro Leu Ile His Ser Leu Leu Met Lys  
 35 40 45

Gln Arg Lys Phe Leu Tyr His Phe Lys Asn Val Arg Trp Ala Lys Gly  
 50 55 60

Arg His Glu Thr Tyr Leu Cys Tyr Val Val Lys Arg Arg Asp Ser Ala  
 65 70 75 80

Thr Ser Phe Ser Leu Asp Phe Gly His Leu Arg Asn Lys Ser Gly Cys  
 85 90 95

His Val Glu Leu Leu Phe Leu Arg Tyr Ile Ser Asp Trp Asp Leu Asp  
 100 105 110

Pro Gly Arg Cys Tyr Arg Val Thr Trp Phe Thr Ser Trp Ser Pro Cys  
 115 120 125

Tyr Asp Cys Ala Arg His Val Ala Asp Phe Leu Arg Gly Tyr Pro Asn  
 130 135 140

10

ES 2 498 765 T3

Leu Ser Leu Arg Ile Phe Ala Ala Arg Leu Tyr Phe Cys Glu Asp Arg  
145 150 155 160

Lys Ala Glu Pro Glu Gly Leu Arg Arg Leu His Arg Ala Gly Val Gln  
165 170 175

Ile Ala Ile Met Thr Phe Lys Asp Tyr Phe Tyr Cys Trp Asn Thr Phe  
180 185 190

Val Glu Asn Arg Glu Arg Thr Phe Lys Ala Trp Glu Gly Leu His Glu  
195 200 205

Asn Ser Val Arg Leu Ser Arg Gln Leu Arg Arg Ile Leu Leu Pro Leu  
210 215 220

Tyr Glu Val Asp Asp Leu Arg Asp Ala Phe Arg Thr Leu Gly Leu  
225 230 235

<210> 12  
<211> 201  
5 <212> PRT  
<213> Xenopus sp.  
  
<400> 12

Met Thr Met Asp Ser Met Leu Leu Lys Arg Asn Lys Phe Ile Tyr His  
1 5 10 15

Tyr Lys Asn Leu Arg Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys  
20 25 30

Tyr Ile Val Lys Arg Arg Tyr Ser Ser Val Ser Cys Ala Leu Asp Phe  
35 40 45

Gly Tyr Leu Arg Asn Arg Asn Gly Cys His Ala Glu Met Leu Phe Leu  
50 55 60

Arg Tyr Leu Ser Ile Trp Val Gly His Asp Pro His Arg Asn Tyr Arg  
65 70 75 80

Val Thr Trp Phe Ser Ser Trp Ser Pro Cys Tyr Asp Cys Ala Lys Arg  
85 90 95

Thr Leu Glu Phe Leu Lys Gly His Pro Asn Phe Ser Leu Arg Ile Phe  
100 105 110

Ser Ala Arg Leu Tyr Phe Cys Glu Glu Arg Asn Ala Glu Pro Glu Gly  
115 120 125

10

ES 2 498 765 T3

Leu Arg Lys Leu Gln Lys Ala Gly Val Arg Leu Ser Val Met Ser Tyr  
 130 135 140

Lys Asp Tyr Phe Tyr Cys Trp Asn Thr Phe Val Glu Thr Arg Glu Ser  
 145 150 155 160

Gly Phe Glu Ala Trp Asp Gly Leu His Glu Asn Ser Val Arg Leu Ala  
 165 170 175

Arg Lys Leu Arg Arg Ile Leu Gln Pro Pro Tyr Asp Met Glu Asp Leu  
 180 185 190

Arg Glu Val Phe Val Leu Leu Gly Leu  
 195 200

<210> 13

<211> 204

5 <212> PRT

<213> Takifugu rubripes

<400> 13

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160  
 Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala Trp Asp Leu  
 180 185 190

Glu Asp Leu Arg Asp Ala Leu Lys Leu Leu Gly Phe  
 195 200

10

ES 2 498 765 T3

<210> 14  
 <211> 210  
 <212> PRT  
 5 <213> Danio rerio

<400> 14

Met Ile Cys Lys Leu Asp Ser Val Leu Met Thr Gln Lys Lys Phe Ile  
 1 5 10 15  
 Phe His Tyr Lys Asn Val Arg Trp Ala Arg Gly Arg His Glu Thr Tyr  
 20 25 30  
 Leu Cys Phe Val Val Lys Arg Arg Ile Gly Pro Asp Ser Leu Ser Phe  
 35 40 45  
 Asp Phe Gly His Leu Arg Asn Arg Ser Gly Cys His Val Glu Leu Leu  
 50 55 60  
 Phe Leu Arg His Leu Gly Ala Leu Cys Pro Gly Leu Ser Ala Ser Ser  
 65 70 75 80  
 Val Asp Gly Ala Arg Leu Cys Tyr Ser Val Thr Trp Phe Cys Ser Trp  
 85 90 95  
 Ser Pro Cys Ser Lys Cys Ala Gln Gln Leu Ala His Phe Leu Ser Gln  
 100 105 110  
 Thr Pro Asn Leu Arg Leu Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys  
 115 120 125  
 Asp Glu Glu Asp Ser Val Glu Arg Glu Gly Leu Arg His Leu Lys Arg  
 130 135 140  
 Ala Gly Val Gln Ile Ser Val Met Thr Tyr Lys Asp Phe Phe Tyr Cys  
 145 150 155 160  
 Trp Gln Thr Phe Val Ala Arg Arg Glu Arg Ser Phe Lys Ala Trp Asp  
 165 170 175  
 Gly Leu His Glu Asn Ser Val Arg Leu Val Arg Lys Leu Asn Arg Ile  
 180 185 190  
 Leu Gln Pro Cys Glu Thr Glu Asp Leu Arg Asp Val Phe Ala Leu Leu  
 195 200 205

Gly Leu  
 210

10

<210> 15  
 <211> 198  
 <212> PRT  
 15 <213> Secuencia artificial

<220>  
 <223> AID mutant Mut1

20 <400> 15

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro, Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175  
 Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

5 <210> 16  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

10 <220>  
 <223> mutante de AID Mutl.1

<400> 16

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Thr Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- <210> 17
- 5 <211> 198
- <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> AID mutant Mut1,1.1
- <400> 17

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Met Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Gly Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Thr Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Ile Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- <210> 18
- 5 <211> 198
- <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> AID mutant Mut1,1.2
- <400> 18

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Ile Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Tyr Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Asp Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Thr Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Pro Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Val  
 195

- <210> 19
- <211> 198
- 5 <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> AID mutant Mut1,1.3
- <400> 19

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Lys Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Gly  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Thr Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Met Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- <210> 20
- <211> 198
- 5 <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> AID mutant Mut1,1.4
- <400> 20

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Phe Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Thr Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Arg Leu  
 195

<210> 21  
 <211> 198  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> AID mutant Mut1,2  
 10 <400> 21

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 22  
 <211> 189  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> AID mutant Mut1,3

<400> 22

5

10

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Lys Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu  
 180 185

<210> 23  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> AID mutant Mut1,4

<400> 23

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

ES 2 498 765 T3

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 24  
 <211> 189  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> AID mutant Mut1,5

10

<400> 24

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr



ES 2 498 765 T3

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

<210> 26

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut3

10

<400> 26

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

ES 2 498 765 T3

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

<210> 27

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut4

10

<400> 27

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
20 25 30

Val Lys Gly Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

ES 2 498 765 T3

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Leu Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- <210> 28
- <211> 183
- 5 <212> PRT
- <213> Secuencia artificial
- <220>
- <223> Mutante de AID Mut5
- 10 <400> 28

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg



ES 2 498 765 T3

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Pro Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 30

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut7

10

<400> 30

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

ES 2 498 765 T3

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

<210> 31

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut7.1

10

<400> 31

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
100 105 110

Leu Tyr Tyr Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
145 150 155 160  
Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

15

ES 2 498 765 T3

<210> 32  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Mutante de AID Mut7.2

<400> 32

10

```

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys
 1          5          10          15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val
 20          25          30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr
 35          40          45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr
 50          55          60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp
 65          70          75          80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp
 85          90          95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg
100          105          110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg
115          120          125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr
130          135          140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys
145          150          155          160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu
165          170          175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala
180          185          190

Phe Arg Thr Leu Gly Leu
195
    
```

<210> 33  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

15

<220>  
 <223> Mutante de AID Mut7.3

20

<400> 33

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

- <210> 34
- 5 <211> 198
- <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> Mutante de AID Mut7.3.1
- <400> 34

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Ser Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Cys Asp Ser Ala Thr Ser Phe Ser Arg Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Ser Asp Cys Ala Arg Leu Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Glu Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190  
 Phe Arg Thr Leu Gly Leu  
 195

5 <210> 35  
 <211> 180  
 <212> PRT  
 <213> Secuencia artificial

10 <220>  
 <223> Mutante de AID Mut7.3.2

<400> 35

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Thr Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Ile Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Glu  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu  
 180

- <210> 36
- <211> 198
- 5 <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> Mutante de AID Mut7.3.3
- <400> 36

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Tyr Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Tyr Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Gly Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 37  
 <211> 198  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Mut7.3.4.

10 <400> 37

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Trp Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Tyr Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Trp Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Asn Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

5 <210> 38  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

10 <220>  
 <223> Mutante de AID Mut7.3.5

<400> 38

ES 2 498 765 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Ser Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Met Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Arg Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Ile Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 39  
 <211> 198  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Mut7.3.6

10 <400> 39

Met Asp Ser Leu Leu Met Asn Arg Ser Glu Phe Leu Tyr Gln Phe Lys



ES 2 498 765 T3

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Tyr Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Leu Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 41  
 <211> 198  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Mut7.3.8

10 <400> 41

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Ala Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

ES 2 498 765 T3

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg Arg Val Ala Asp  
 85 90 95

Phe Leu Arg Glu Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Glu  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Thr  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 42  
 <211> 198  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Mut7.3.9

10 <400> 42

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

ES 2 498 765 T3

Leu Arg Asn Lys Asn Gly Cys His Ala Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 43

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut7.4

10

<400> 43

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr



ES 2 498 765 T3

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Leu Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

<210> 45

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut7.6

10

<400> 45

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

ES 2 498 765 T3

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ser Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Tyr Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 46  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Mut7.7

<400> 46

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

ES 2 498 765 T3

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 47  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Mutante de AID Mut7.8

10

<400> 47

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg

ES 2 498 765 T3

100 105 110

Leu Tyr Phe Cys Glu Asp Arg Arg Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

5 <210> 48  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

10 <220>  
 <223> Mutante de AID Mut8  
 <400> 48

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

ES 2 498 765 T3

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Glu  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 49  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Mutante de AID Mut9

10

<400> 49

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Glu Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

ES 2 498 765 T3

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu  
 195

<210> 50

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut10

10

<400> 50

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Pro Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Arg Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

ES 2 498 765 T3

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

<210> 51

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Mut11

10

<400> 51

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys



ES 2 498 765 T3

<210> 53  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Mutante de AID Mut13

<400> 53

10

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15  
 Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30  
 Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45  
 Leu Arg Asn Lys Asn Gly Cys His Ala Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60  
 Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80  
 Tyr Thr Ser Arg Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95  
 Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110  
 Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125  
 Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140  
 Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160  
 Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175  
 Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190  
 Phe Arg Thr Leu Gly Leu  
 195

15

<210> 54  
 <211> 189  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Fugu1

20

<400> 54

ES 2 498 765 T3

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Leu  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 55

<211> 189

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu2

10

<400> 55

ES 2 498 765 T3

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Arg Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 56  
 <211> 228  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Fugu3

10 <400> 56

ES 2 498 765 T3

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Ser Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Arg Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Glu Leu Asn Arg Ile Leu Gln Ala Trp Asp Leu  
 180 185 190

Asp Asp Leu Arg Asp Ala Leu Asn Pro Pro Asp Ser Glu Asp Pro Leu  
 195 200 205

Glu Ser Thr Cys Arg His Ala Ser Leu Ala Val Leu Ala Asp Glu Arg  
 210 215 220

Arg Phe Ser Ala  
 225

<210> 57  
 <211> 189  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Fugu4

<400> 57

ES 2 498 765 T3

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 58

<211> 189

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu4,1

10

<400> 58

ES 2 498 765 T3

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys His Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 59  
 <211> 189  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Fugu4,2

10 <400> 59

Met Leu Leu Pro Arg Asn Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

ES 2 498 765 T3

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Tyr Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 60  
 <211> 189  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Mutante de AID Fugu4,3

10

<400> 60

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

ES 2 498 765 T3

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Leu Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 61

<211> 189

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu4,4

10

<400> 61

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

ES 2 498 765 T3

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Tyr Cys Asp Leu Glu Asp Ser Leu  
115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Asp Leu His Gln Asn Ser  
165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
180 185

<210> 62

<211> 189

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu4,5

10

<400> 62

Met Leu Leu Pro Ser Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
50 55 60

Leu Cys Pro Gly Leu Trp Gly Asn Gly Ala Ala Gly Glu Lys Arg Leu  
65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu

ES 2 498 765 T3

	100		105		110
Arg	Ile Phe Val Ser Arg Leu Tyr Tyr Cys Asp Leu Glu Asp Ser Leu				
	115		120		125
Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser					
	130		135		140
Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp					
	145		150		155
Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser					
		165		170	175
Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala					
	180		185		

<210> 63  
 <211> 189  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Mutante de AID Fugu4,6

10

<400> 63

Met	Leu	Leu	Pro	Arg	Lys	Lys	Phe	Ile	Tyr	His	Tyr	Lys	Asn	Val	Arg
1				5					10					15	
Trp	Ala	Arg	Gly	Arg	His	Glu	Thr	Tyr	Leu	Cys	Phe	Val	Val	Lys	Arg
			20					25					30		
Arg	Val	Gly	Pro	Asp	Thr	Leu	Thr	Leu	Asp	Phe	Gly	His	Leu	Arg	Asn
		35					40					45			
Arg	Ser	Gly	Cys	His	Val	Glu	Leu	Leu	Phe	Leu	Arg	Tyr	Leu	Gly	Ala
	50					55					60				
Leu	Cys	Pro	Gly	Leu	Trp	Gly	Tyr	Gly	Ala	Ala	Gly	Glu	Lys	Arg	Leu
65					70					75					80
Ser	Tyr	Ser	Val	Thr	Trp	Phe	Cys	Ser	Trp	Ser	Pro	Cys	Val	Asn	Cys
				85					90					95	
Ser	Ile	Gln	Leu	Cys	Gln	Phe	Leu	Asn	Asn	Thr	Pro	Asn	Leu	Arg	Leu
			100					105					110		
Arg	Ile	Phe	Val	Ser	Arg	Leu	Tyr	Phe	Cys	Asp	Pro	Glu	Asp	Ser	Leu
		115					120					125			

ES 2 498 765 T3

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 64

<211> 189

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu4,7

10

<400> 64

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Ile Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Gln Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

15

<210> 65

<211> 189

<212> PRT

ES 2 498 765 T3

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu4,8

5

<400> 65

```

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg
 1           5           10           15
Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg
          20           25           30
Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn
      35           40           45
Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala
      50           55           60
Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu
 65           70           75
Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys
          85           90           95
Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu
          100          105          110
Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Met Glu Asp Ser Leu
          115          120          125
Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser
      130          135          140
Val Met Ser Tyr Lys Asp Tyr Val Tyr Cys Trp Gln Lys Phe Val Asp
 145          150          155          160
Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser
          165          170          175
Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala
          180          185

```

10 <210> 66  
 <211> 189  
 <212> PRT  
 <213> Secuencia artificial

15 <220>  
 <223> Mutante de AID Fugu4,9  
 <400> 66

ES 2 498 765 T3

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Pro  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 67  
 <211> 189  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID Fugu4,10

10 <400> 67

ES 2 498 765 T3

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Pro  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Glu Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala  
 180 185

<210> 68

<211> 207

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu5

10

<400> 68

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

ES 2 498 765 T3

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg  
 20 25 30

Arg Val Gly Pro Asp Thr Leu Thr Phe Asp Phe Gly His Leu Arg Asn  
 35 40 45

Arg Arg Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr Leu Gly Ala  
 50 55 60

Leu Cys Pro Gly Leu Trp Gly Tyr Gly Ala Ala Gly Glu Lys Arg Leu  
 65 70 75 80

Ser Tyr Ser Val Thr Trp Phe Cys Ser Trp Ser Pro Cys Val Asn Cys  
 85 90 95

Ser Ile Gln Leu Cys Gln Phe Leu Asn Asn Thr Pro Asn Leu Arg Leu  
 100 105 110

Arg Ile Phe Val Ser Arg Leu Tyr Phe Cys Asp Leu Glu Asp Ser Leu  
 115 120 125

Glu Arg Glu Gly Leu Arg Met Leu Thr Lys Ala Gly Val Arg Ile Ser  
 130 135 140

Val Met Ser Tyr Lys Asp Tyr Phe Tyr Cys Trp Gln Lys Phe Val Asp  
 145 150 155 160

Cys Lys Lys Ser Asn Phe Lys Ala Trp Glu Gly Leu His Gln Asn Ser  
 165 170 175

Val Arg Leu Thr Arg Lys Leu Asn Arg Ile Leu Gln Ala Trp Asp Leu  
 180 185 190

Glu Asp Leu Arg Asp Thr Leu Ser Pro Arg Ile Leu Arg Ile Leu  
 195 200 205

<210> 69

<211> 191

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID Fugu6

10

<400> 69

Met Leu Leu Pro Arg Lys Lys Phe Ile Tyr His Tyr Lys Asn Val Arg  
 1 5 10 15

Trp Ala Arg Gly Arg His Glu Thr Tyr Leu Cys Phe Val Val Lys Arg

ES 2 498 765 T3

	20		25		30														
Arg	Val	Gly	Pro	Asp	Thr	Leu	Thr	Ser	Asp	Phe	Gly	His	Leu	Arg	Asn				
		35					40					45							
Arg	Ser	Gly	Cys	His	Val	Glu	Leu	Leu	Phe	Leu	Arg	Tyr	Leu	Gly	Ala				
	50					55					60								
Leu	Cys	Pro	Gly	Leu	Trp	Gly	Tyr	Gly	Ala	Ala	Gly	Glu	Lys	Arg	Leu				
65					70				75						80				
Ser	Tyr	Ser	Val	Thr	Trp	Phe	Cys	Ser	Trp	Ser	Pro	Cys	Val	Asn	Cys				
				85					90					95					
Ser	Ile	Gln	Leu	Cys	Gln	Phe	Leu	Asn	Asn	Thr	Pro	Asn	Leu	Arg	Leu				
			100					105					110						
Arg	Ile	Phe	Val	Ser	Arg	Leu	Tyr	Phe	Cys	Asp	Leu	Glu	Asn	Ser	Leu				
		115					120					125							
Glu	Arg	Glu	Gly	Leu	Arg	Met	Leu	Thr	Lys	Ala	Gly	Val	Arg	Ile	Ser				
	130					135					140								
Val	Met	Ser	Tyr	Lys	Asp	Tyr	Phe	Tyr	Cys	Trp	Gln	Lys	Phe	Val	Asp				
145					150					155					160				
Cys	Lys	Lys	Ser	Asn	Phe	Lys	Ala	Trp	Glu	Glu	Leu	His	Gln	Asn	Ser				
				165					170					175					
Val	Arg	Leu	Thr	Arg	Lys	Leu	Asn	Arg	Ile	Leu	Gln	Ala	Trp	Asp					
			180					185						190					

<210> 70  
 <211> 198  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> AB350 w/o FLAG/NLS

10 <400> 70

Met	Asp	Ser	Leu	Leu	Met	Lys	Gln	Arg	Lys	Phe	Leu	Tyr	His	Phe	Lys
1				5					10					15	
Asn	Val	Arg	Trp	Ala	Lys	Gly	Arg	His	Glu	Thr	Tyr	Leu	Cys	Tyr	Val
			20					25					30		
Val	Lys	Arg	Arg	Asp	Ser	Ala	Thr	Ser	Phe	Ser	Leu	Asp	Phe	Gly	His
		35					40					45			

ES 2 498 765 T3

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

Phe Leu Arg Gly Tyr Pro Asn Leu Ser Leu Arg Ile Phe Ala Ala Arg  
100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Lys Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Ala  
195

<210> 71

<211> 198

5 <212> PRT

<213> Secuencia artificial

<220>

<223> AB273 w/o FLAG/NLS

10

<400> 71

Met Asp Ser Leu Leu Met Lys Gln Arg Lys Phe Leu Tyr His Phe Lys  
1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
35 40 45

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
50 55 60

ES 2 498 765 T3

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
85 90 95

Phe Leu Arg Gly Tyr Pro Asn Leu Ser Leu Arg Ile Phe Ala Ala Arg  
100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Lys Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Glu Glu Leu Arg Glu Ala  
180 185 190

Phe Arg Ile Leu Gly Ala  
195

<210> 72  
<211> 214  
5 <212> PRT  
<213> Secuencia artificial

<220>  
<223> AB350

10 <400> 72

Met Asp Tyr Lys Asp Asp Asp Asp Lys Gly Pro Lys Lys Lys Arg Lys  
1 5 10 15

Val Asp Ser Leu Leu Met Lys Gln Arg Lys Phe Leu Tyr His Phe Lys  
20 25 30

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
35 40 45

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
50 55 60

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
65 70 75 80

ES 2 498 765 T3

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 85 90 95

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 100 105 110

Phe Leu Arg Gly Tyr Pro Asn Leu Ser Leu Arg Ile Phe Ala Ala Arg  
 115 120 125

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 130 135 140

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 145 150 155 160

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Lys Thr Phe Lys  
 165 170 175

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 180 185 190

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 195 200 205

Phe Arg Thr Leu Gly Ala  
 210

<210> 73  
 <211> 214  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> AB273

10 <400> 73

Met Asp Tyr Lys Asp Asp Asp Asp Lys Gly Pro Lys Lys Lys Arg Lys  
 1 5 10 15

Val Asp Ser Leu Leu Met Lys Gln Arg Lys Phe Leu Tyr His Phe Lys  
 20 25 30

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
 35 40 45

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
 50 55 60

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr



ES 2 498 765 T3

<220>  
<223> AB273

5 <400> 75  
 atggactaca aagatgacga tgataaaggt ccaaagaaga agagaaaagg agactctctc 60  
 ctcatgaagc agagaaagtt tctctaccac ttcaagaacg tcagatgggc caaggggaga 120  
 catgagacct atctctgtta cgtcgtcaag aggagagact cagccacctc tttctccctc 180  
 gactttgggc atctccgga caagtctggg tgtcatgtcg aactcctctt cctccgctat 240  
 atctcagact gggacctcga ccccgggaga tgctatagag tcacttgggt tacctcttgg 300  
 tccccctggt atgactgctc cagacatgct gccgacttcc tcagggggta tcccaatctc 360  
 tccctccgca tattcgccgc ccgactctat ttttgtgagg acaggaaagc cgagcccgag 420  
 gggctcagga gactccaccg ggccggggtc cagatcgcca tcatgacatt taaggactat 480  
 ttctattggt ggaatacatt tgtcgagaat cgggagaaga ctttcaaagc ctggggaggg 540  
 ctccatgaga actctgtcag actctctagg cagctcagga gaatcctctt ccccctctat 600  
 gaggtcgaag aactcagaga agccttccgg atcctcgggg cttga 645

10 <210> 76  
 <211> 7  
 <212> PRT  
 <213> Secuencia artificial

15 <220>  
 <223> NLS consenso  
 <400> 76

Pro Lys Lys Lys Arg Lys Val  
 1 5

20 <210> 77  
 <211> 16  
 <212> PRT  
 <213> Secuencia artificial

25 <220>  
 <223> NLS bipartita

30 <400> 77  
 Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys  
 1 5 10 15

35 <210> 78  
 <211> 5  
 <212> PRT  
 <213> Secuencia artificial

40 <220>  
 <223> Motivo NLS  
 <400> 78

Lys Ile Pro Ile Lys  
 1 5

45 <210> 79  
 <211> 22

<212> ADN  
 <213> Secuencia artificial  
  
 <220>  
 5 <223> LacZ cebador directo  
  
 <400> 79  
  
 10                   agaattcctg aagttcagat gt           22  
  
 <210> 80  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial  
 15  
 <220>  
 <223> LacZ cebador inverso  
  
 <400> 80  
 20                   ggaattcgaa accgccaaga c           21  
  
 <210> 81  
 <211> 24  
 25 <212> ADN  
 <213> Secuencia artificial  
  
 <220>  
 <223> hAID cebador directo  
 30  
 <400> 81  
  
 35                   atggaattca tggacagcct cttg           24  
  
 <210> 82  
 <211> 25  
 <212> ADN  
 <213> Secuencia artificial  
 40  
 <220>  
 <223> hAID cebador inverso  
  
 <400> 82  
 45                   ctgaagcttt caaagtcca aagta           25  
  
 <210> 83  
 <211> 21  
 <212> ADN  
 50 <213> Secuencia artificial  
  
 <220>  
 <223> rpoB cebador directo  
 55 <400> 83  
  
 60                   ttggcgaaat ggcggaaaac c           21  
  
 <210> 84  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial  
 65  
 <220>  
 <223> rpoB cebador inverso

ES 2 498 765 T3

<400> 84  
 caccgacgga taccacctgc tg 22

5 <210> 85  
 <211> 44  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> Sustrato de desaminasa (5'-FITC; 3'-biotina)

<400> 85

15 atatgaatag aatagagggg tgagctgggg tgagctgggg tgag 44

<210> 86  
 <211> 5  
 <212> PRT  
 <213> Secuencia artificial

20 <220>  
 <223> Motivo de coordinación de cinc

25 <400> 86

Pro Cys Tyr Asp Cys  
 1 5

<210> 87  
 <211> 7  
 <212> PRT  
 <213> Secuencia artificial

30 <220>  
 <223> Motivo de coordinación de cinc

35 <400> 87

Phe Cys Glu Asp Arg Lys Ala  
 1 5

40 <210> 88  
 <211> 645  
 <212> ADN  
 <213> Secuencia artificial

45 <220>  
 <223> Mutante de AID canina

50 <400> 88

ES 2 498 765 T3

```

atggactaca aagatgacga tgataaaggt ccaaagaaga agagaaaggt agactctctc      60
ctcatgaagc agagaaagtt tctctaccac ttcaagaacg tcagatgggc caaggggaga      120
catgagacct atctctgtta cgtcgtcaag aggagagact cagccacctc tttctccctc      180
gactttgggc atctccggaa caagtctggg tgtcatgtcg aactcctctt cctccgctat      240
atctcagact gggacctcga ccccgggaga tgctatagag tcacttggtt tacctcttgg      300
tccccctggt atgactgctc cagacatgtc gccgacttcc tcagggggta tcccaatctc      360
tccctccgca tattcgccgc ccgactctat ttttgtgagg acaggaaagc cgagcccagag      420
gggctcagga gactccaccg ggccggggtc cagatcgcca tcatgacatt taaggactat      480
ttctattggt ggaatacatt tgtcagagaat cgggagaaga ctttcaaagc ctgggagggg      540
ctccatgaga actctgtcag actctctagg cagctcagga gaatcctcct ccccctctat      600
gaggtcgaag aactcagaga agccttccgg atcctcgggg cttga                          645

```

<210> 89  
 <211> 645  
 <212> ADN  
 <213> Secuencia artificial

5

<220>  
 <223> Mutante de AID canina

10

<400> 89

```

atggactaca aagatgacga tgataaaggt ccaaagaaga agagaaaggt agactctctc      60
ctcatgaagc agagagaggt tctctaccac ttcaagaacg tcagatgggc caaggggaga      120
catgagacct atctctgtta cgtcgtcaag aggagagact cagccacctc tttctccctc      180
gactttgggc atctccggaa caagtctggg tgtcatgtcg aactcctctt cctccgctat      240
atctcagact gggacctcga ccccgggaga tgctatagag tcacttggtt tatctcttgg      300
tccccctggt atgactgctc cagacatgtc gccgacttcc tcagggggta tcccaatctc      360
tccctccgca tattcgccgc ccgactctat ttttgtgagg acaggaaagc cgagcccagag      420
gggctcagga gactccaccg ggccggggtc cagatcgcca tcatgacatt taaggactat      480
ttctattggt ggaatacatt tgtcagagaat cgggggaaga ctttcaaagc ctgggagggg      540
ctccatgaga actctgtcag actctctagg cagctcagga gaatcctcct ccccctctat      600
gaggtcgaag aactcagaga agccttccgg atcctcgggg cttga                          645

```

15 <210> 90  
 <211> 597  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> Mutante de AID humana

<400> 90

ES 2 498 765 T3

```

atggacagcc tttgatgaa cggaggagg tttctttacc aattcaaaaa tgtccgctgg      60
gctaagggtc ggcgtgagac ctacctgtgc tacgtagtga agaggcgtga cagtgtaca      120
tccttttcac tggactttgg ttatcttcgc aataagaacg gctgccacgt ggaattgctc      180
ttcctccgct acatctcgga ctgggaccta gaccctggcc gctgctaccg cgtcacctgg      240
ttcatctcct ggagcccctg ctacgactgt gcccgacatg tggccgactt tctgcgaggg      300
aaccccaacc tcagtctgag gatcttcacc gcgcgcctct acttctgtga ggaccgcaag      360
gctgagcccg agyggctgcg gcggctgcac cgcgcggggg tgcaaatagc catcatgacc      420
ttcaaagatt atttttactg ctggaatact tttgtagaaa accacggaag aactttcaaa      480
gcctgggaag ggtgcatga aaattcagtt cgtctctcca gacagcttcg gcgcacctct      540
ttgccctgtg atgaggttga tgacttacga gacgcatttc gtactttggg acttttga      597

```

<210> 91

<211> 214

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Mutante de AID canina

<400> 91

```

Met Asp Tyr Lys Asp Asp Asp Asp Lys Gly Pro Lys Lys Lys Arg Lys
 1          5          10
Val Asp Ser Leu Leu Met Lys Gln Arg Lys Phe Leu Tyr His Phe Lys
 20        25        30
Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val
 35        40        45
Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His
 50        55        60
Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr
 65        70        75
Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp
 85        90        95
Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp
100       105       110
Phe Leu Arg Gly Tyr Pro Asn Leu Ser Leu Arg Ile Phe Ala Ala Arg
115       120       125

```

ES 2 498 765 T3

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 130 135 140

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 145 150 155 160

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Lys Thr Phe Lys  
 165 170 175

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 180 185 190

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Glu Glu Leu Arg Glu Ala  
 195 200 205

Phe Arg Ile Leu Gly Ala  
 210

<210> 92

<211> 214

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Mutante de AID canina

10

<400> 92

Met Asp Tyr Lys Asp Asp Asp Asp Lys Gly Pro Lys Lys Lys Arg Lys  
 1 5 10 15

Val Asp Ser Leu Leu Met Lys Gln Arg Glu Phe Leu Tyr His Phe Lys  
 20 25 30

Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val  
 35 40 45

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly His  
 50 55 60

Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 65 70 75 80

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 85 90 95

Phe Ile Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 100 105 110

Phe Leu Arg Gly Tyr Pro Asn Leu Ser Leu Arg Ile Phe Ala Ala Arg

ES 2 498 765 T3

115 120 125

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 130 135 140

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 145 150 155 160

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Gly Lys Thr Phe Lys  
 165 170 175

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 180 185 190

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Glu Glu Leu Arg Glu Ala  
 195 200 205

Phe Arg Ile Leu Gly Ala  
 210

<210> 93  
 <211> 198  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Mutante de AID humana

<400> 93

Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Ile Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

ES 2 498 765 T3

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys  
145 150 155 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
180 185 190

Phe Arg Thr Leu Gly Leu  
195

## REIVINDICACIONES

- 5 1. Una molécula de ácido nucleico aislada o purificada que comprende una secuencia de nucleótidos que codifica una citidina desaminasa inducida por activación (AID) mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácido seleccionada de:
- (a) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuo 34, 82 y 156,
  - (b) en al menos una sustitución de aminoácidos en el residuo 10 y al menos una sustitución de aminoácidos en el residuo 156,
  - 10 (c) en al menos una sustitución de aminoácidos en el residuo 35 y al menos una sustitución de aminoácidos en el residuo 145,
  - (d) en al menos una sustitución de aminoácidos en el residuo 34 y al menos una sustitución de aminoácidos en el residuo 160,
  - 15 (e) en al menos una sustitución de aminoácidos en el residuo 43 y al menos una sustitución de aminoácidos en el residuo 120,
  - (f) al menos dos sustituciones de aminoácidos en las que al menos una sustitución está en el residuo 57 y al menos una sustitución está en el residuo 81 o 145,
  - (g) en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 82,
  - 20 (h) en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 34,
  - (i) en al menos una sustitución de aminoácidos en el residuo 156 y al menos una sustitución de aminoácidos en el residuo 157,
  - 25 (j) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuo 10, 82 y 156,
  - (k) en al menos una sustitución de aminoácidos en el residuo 115, y
  - (l) en al menos una sustitución de aminoácidos en el residuo 120,
- en la que la proteína AID mutante funcional tiene al menos una mejora de 10 veces la actividad en comparación con la proteína AID humana en un ensayo de papilación bacteriana.
- 30 2. La molécula de ácido nucleico aislada o purificada de la reivindicación 1, en la que la proteína AID mutante funcional comprende una sustitución de aminoácidos en el residuo 34.
3. La molécula de ácido nucleico aislada o purificada de la reivindicación 2, en la que la sustitución de aminoácidos en el residuo 34 es K34E o K34D.
- 35 4. La molécula de ácido nucleico aislada o purificada de una cualquiera de las reivindicaciones 1-3, en la que la proteína AID mutante funcional comprende una sustitución de aminoácidos en el residuo 82.
5. La molécula de ácido nucleico aislada o purificada de la reivindicación 4, en la que la sustitución de aminoácidos en el residuo 82 es T82I o T82L.
6. La molécula de ácido nucleico aislada o purificada de una cualquiera de las reivindicaciones 1-5, en la que la proteína AID mutante funcional comprende una sustitución de aminoácidos en el residuo 156.
- 40 7. La molécula de ácido nucleico aislada o purificada de la reivindicación 6, en el que la sustitución de aminoácidos en el residuo 156 es E156G o E156A.
8. La molécula de ácido nucleico aislada o purificada de una cualquiera de las reivindicaciones 2-7, en la que la proteína citidina desaminasa inductiva por activación (AID) mutante funcional comprende además al menos una sustitución de aminoácidos en el residuo 157.
- 45 9. La molécula de ácido nucleico aislada o purificada de la reivindicación 8, en la que la sustitución de aminoácidos en el residuo 157 es treonina (T) o lisina (K).
10. La molécula de ácido nucleico aislada o purificada de la reivindicación 8 o la reivindicación 9,
- en la que la proteína AID mutante funcional comprende además uno o más de los siguientes:
- 50 (a) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuo 9, 38 y 180,
  - (b) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuo 42, 115, 132, 183, y 198,
  - (c) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuo 9, 96 y 181, y
  - 55 (d) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en el residuo

13 y el residuo 197,

11. La molécula de ácido nucleico aislada o purificada de la reivindicación 10, en la que

- (a) la sustitución del aminoácido en el residuo 9 es metionina (M) o lisina (K),
- (b) la sustitución del aminoácido en el residuo 13 es fenilalanina (F) o triptófano (W),
- 5 (c) la sustitución de aminoácido en el residuo 38 es glicina (G) o alanina (A).
- (d) la sustitución de aminoácido en el residuo 42 es isoleucina (I) o leucina (L).
- (e) la sustitución de aminoácido en el residuo 96 es glicina (G) o alanina (A).
- (f) la sustitución del aminoácido en el residuo 115 es tirosina (Y) o triptófano (W),
- (g) la sustitución de aminoácido en el residuo 132 es ácido glutámico (E) o ácido aspártico (D),
- 10 (h) la sustitución de aminoácido en el residuo 180 es isoleucina (I) o alanina (A).
- (i) la sustitución del aminoácido en el residuo 181 es metionina (M) o valina (V),
- (j) la sustitución de aminoácido en el residuo 183 es isoleucina (I) o prolina (P).
- (k) la sustitución del aminoácido en el residuo 197 es arginina (R) o lisina (K), y
- (l) la sustitución de aminoácido en el residuo 198 es valina (V) o leucina (L).

15 12. La molécula de ácido nucleico aislada o purificada de la reivindicación 1, en el que la proteína AID mutante funcional comprende una sustitución de aminoácidos en el residuo 10 y una sustitución de aminoácido en el residuo 156.

20 13. La molécula de ácido nucleico aislada o purificada de la reivindicación 12, en la que el aminoácido en el residuo 10 está sustituido con ácido glutámico (E) o ácido aspártico (D) y el aminoácido en el residuo 156 está sustituido con glicina (G) o alanina (A).

14. La molécula de ácido nucleico aislada o purificada de la reivindicación 12 o la reivindicación 13, en la que la proteína AID mutante funcional comprende al menos una sustitución adicional de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 13, 34, 82, 95, 115, 120, 134 y 145.

15. La molécula de ácido nucleico aislada o purificada de la reivindicación 14, en la que

- 25 (a) la sustitución del aminoácido en el residuo 13 es fenilalanina (F) o triptófano (W),
- (b) la sustitución de aminoácido en el residuo 34 es ácido glutámico (E) o ácido aspártico (D),
- (c) la sustitución de aminoácido en el residuo 82 es isoleucina (I) o leucina (L).
- (d) la sustitución de aminoácido en el residuo 95 es serina (S) o leucina (L).
- 30 (e) la sustitución del aminoácido en el residuo 115 es tirosina (Y) o triptófano (W),
- (f) la sustitución del aminoácido en el residuo 120 es arginina (R) o asparagina (N), y
- (g) la sustitución de aminoácido en el residuo 145 es leucina (I) o isoleucina (I).

35 16. Una molécula de ácido nucleico aislada o purificada que comprende una secuencia de nucleótidos que codifica una citidina desaminasa inducida por activación (AID) mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácido seleccionada de:

- (a) una sustitución de aminoácidos en el residuo 35 y una sustitución de aminoácidos en el residuo 145,
- (b) una sustitución de aminoácidos en el residuo 34 y una sustitución de aminoácidos en el residuo 160,
- (c) una sustitución de aminoácidos en el residuo 43 y una sustitución de aminoácidos en el residuo 120,
- 40 (d) al menos dos sustituciones de aminoácidos en las que una sustitución está en el residuo 57 y una sustitución está en el residuo 81 o 145,
- (e) una sustitución de aminoácidos en el residuo 156 y una sustitución de aminoácidos en el residuo 82,
- (f) una sustitución de aminoácidos en el residuo 156 y una sustitución de aminoácidos en el residuo 34, y
- (g) una sustitución de aminoácidos en el residuo 156 y una sustitución de aminoácidos en el residuo 157.

17. La molécula de ácido nucleico aislada o purificada de la reivindicación 16, en la que

- 45 (a) el aminoácido en el residuo 35 está sustituido con glicina (G) o alanina (A) y el aminoácido en el residuo 145 está sustituido con leucina (L) o isoleucina (I).
- (b) el aminoácido en el residuo 34 está sustituido con ácido glutámico (E) o ácido aspártico (D) y el aminoácido en el residuo 160 está sustituido con ácido glutámico (E) o ácido aspártico (D).
- 50 (c) el aminoácido en el residuo 43 está sustituido con prolina (P) y el aminoácido en el residuo 120 está sustituido con arginina (R),
- (d) el aminoácido en el residuo 156 está sustituido con glicina (G) o alanina (A) y el aminoácido en el residuo 82 está sustituido con leucina (L) o isoleucina (I),
- (e) el aminoácido en el residuo 156 está sustituido con glicina (G) o alanina (A) y el aminoácido en el residuo 34 está sustituido con ácido glutámico (E) o ácido aspártico (D), y
- 55 (f) el aminoácido en el residuo 156 está sustituido con glicina (G) o alanina (A) y el aminoácido en el residuo 157 está sustituido con lisina (K) o asparagina (N),

18. La molécula de ácido nucleico aislada o purificada de la reivindicación 16, en la que al menos dos sustituciones

de aminoácidos están en los residuos 57 y 145.

19. La molécula de ácido nucleico aislada o purificada de la reivindicación 18, en la que el aminoácido en el residuo 57 está sustituido con glicina (G) o alanina (A) y el aminoácido en el residuo 145 está sustituido con leucina (L) o isoleucina (I).

5 20. La molécula de ácido nucleico aislada o purificada de la reivindicación 16, en la que al menos dos sustituciones de aminoácidos están en los residuos 57 y 81.

21. La molécula de ácido nucleico aislada o purificada de la reivindicación 20, en la que el aminoácido en el residuo 57 está sustituido con glicina (G) o alanina (A) y el aminoácido en el residuo 81 está sustituido con tirosina (T) o triptófano (W).

10 22. Una molécula de ácido nucleico aislada o purificada que comprende una secuencia de nucleótidos que codifica una citidina desaminasa inducida por activación (AID) mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de una proteína AID humana (SEC ID N° 1 o SEC ID N° 2) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 10, 82 y 156.

15 23. La molécula de ácido nucleico aislada o purificada de la reivindicación 22, en la que la proteína AID mutante funcional comprende adicionalmente uno o más de los siguientes:

(a) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 9, 36, 44, 88, 93, y 142,

(b) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 66, 104 y 160,

20 (c) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 15, 115 y 185,

(d) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 9, 30, 34, 100, y 184,

25 (e) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 34, 35, 59, 120, y 157,

(f) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 9, 74, 77, 118, 157, y 188,

(g) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en el residuo 53 y el residuo 145, y

30 (h) en al menos una sustitución de aminoácidos en un residuo seleccionado del grupo que consiste en los residuos 18, 93, 100, 160, y 192.

24. La molécula de ácido nucleico aislada o purificada de la reivindicación 22 o la reivindicación 23,

en las que

(a) la sustitución del aminoácido en el residuo 9 es serina (S), metionina (M) o triptófano (W),

35 (b) la sustitución de aminoácido en el residuo 10 es ácido glutámico (E) o ácido aspártico (D),

(c) la sustitución de aminoácido en el residuo 15 es tirosina (Y) o leucina (L).

(d) la sustitución de aminoácido en el residuo 18 es alanina (A) o leucina (L).

(e) la sustitución del aminoácido en el residuo 30 es tirosina (Y) o serina (S),

40 (f) la sustitución de aminoácido en el residuo 34 es ácido glutámico (E) o ácido aspártico (D),

(g) la sustitución del aminoácido en el residuo 35 es serina (S) o lisina (K),

(h) la sustitución de aminoácido en el residuo 36 es cisteína (C),

(i) la sustitución del aminoácido en el residuo 44 es arginina (R) o lisina (K),

(j) la sustitución del aminoácido en el residuo 53 es tirosina (Y) o glutamina (Q),

45 (k) la sustitución del aminoácido en el residuo 59 es metionina (M) o alanina (A),

(l) la sustitución de aminoácido en el residuo 66 es treonina (T) o alanina (A),

(m) la sustitución del aminoácido en el residuo 74 es histidina (H) o lisina (K),

(n) la sustitución del aminoácido en el residuo 77 es serina (S) o lisina (K),

(o) la sustitución de aminoácido en el residuo 82 es isoleucina (I) o leucina (L).

(p) la sustitución del aminoácido en el residuo 88 es serina (S) o treonina (T),

50 (q) la sustitución del aminoácido en el residuo 93 es leucina (L), arginina (R) o lisina (K),

(r) la sustitución de aminoácido en el residuo 100 es ácido glutámico (E), triptófano (W) o fenilalanina (F),

(s) la sustitución de aminoácido en el residuo 104 es isoleucina (I) o alanina (A).

(t) la sustitución de aminoácido en el residuo 115 es tirosina (Y) o leucina (L).

(u) la sustitución del aminoácido en el residuo 118 es ácido glutámico (E) o valina (V),

55 (v) la sustitución de aminoácido en el residuo 120 es arginina (R) o leucina (L).

(w) la sustitución de aminoácido en el residuo 142 es ácido glutámico (E) o ácido aspártico (D),

(x) la sustitución de aminoácido en el residuo 145 es leucina (L) o tirosina (Y).

(y) la sustitución de aminoácido en el residuo 156 es glicina (G) o alanina (A),

(z) la sustitución del aminoácido en el residuo 157 es glicina (G) o lisina (K),

- (aa) la sustitución de aminoácido en el residuo 160 es ácido glutámico (E) o ácido aspártico (D),  
 (bb) la sustitución del aminoácido en el residuo 184 es asparagina (N) o glutamina (Q),  
 (cc) la sustitución de aminoácido en el residuo 185 es glicina (G) o ácido aspártico (D).  
 (dd) la sustitución de aminoácido en el residuo 188 es glicina (G) o ácido glutámico (E), y  
 (ee) la sustitución del aminoácido en el residuo 192 es treonina (T) o serina (S).
- 5 25. La molécula de ácido nucleico aislada o purificada de la reivindicación 1, en el que la proteína AID mutante funcional comprende uno o más de los siguientes:
- (a) una sustitución de aminoácidos en el residuo 115, y  
 (b) una sustitución de aminoácidos en el residuo 120.
- 10 26. La molécula de ácido nucleico aislada o purificada de la reivindicación 25, en la que el aminoácido en el residuo 115 está sustituido por tirosina (Y), triptófano (W) o leucina (L).
27. La molécula de ácido nucleico aislada o purificada de la reivindicación 25 o la reivindicación 26, en la que el aminoácido en el residuo 120 está sustituido por arginina (R), asparagina (N) o leucina (L).
- 15 28. La molécula de ácido nucleico aislada o purificada de una cualquiera de las reivindicaciones 1-27, en la que la proteína AID mutante funcional comprende un truncamiento en C-terminal, una proteína de fusión o un codón de iniciación en o distal al residuo 181.
29. La molécula de ácido nucleico aislada o purificada de una cualquiera de las reivindicaciones 1-28, en la que la proteína AID mutante funcional comprende adicionalmente una sustitución de aminoácidos seleccionada del grupo que consiste en M6K, N7K, R8Q, R9K, K10Q, K10L, Q14H, Q14N, R25H, A39P, F42C, Y48H, N52S, N52A, N52M, D67A, G100A, G100W, E118D, V135A, A138G, Y145F, H156R, R171H, S173T, Q175K, R158K, R194K T195I, L198A, L198F, inserción de K después de 118, y D119E.
- 20 30. La molécula de ácido nucleico aislada o purificada de una cualquiera de las reivindicaciones 1-29, en la que la secuencia de ácido nucleico aislada o purificada se ha optimizado por codones para reducir el número de motivos de hipermutación somática (HMS).
- 25 31. Una molécula de ácido nucleico aislada o purificada que comprende una secuencia de nucleótidos que codifica una citidina desaminasa inducida por activación (AID) mutante funcional cuya secuencia de aminoácidos difiere de la secuencia de aminoácidos de (a) una proteína AID canina (SEC ID N° 3), (b) una proteína AID murina (SEC ID N° 4), (c) una proteína AID de rata (SEC ID N° 5), (d) proteína AID bovina (SEC ID N° 6), o (e) una proteína AID de pollo (SEC ID N° 7) en al menos una sustitución de aminoácidos, en la que un residuo seleccionado del grupo que consiste en el residuo 34, el residuo 82 y el residuo 156, en la que la proteína AID mutante funcional tiene al menos una mejora de 10 veces la actividad en comparación con la proteína AID humana en un ensayo de papilación bacteriana.
- 30 32. Un vector de expresión que comprende una molécula de ácido nucleico de cualquiera de las reivindicaciones 1-31.
- 35 33. Una célula eucariota o una célula procariota aislada que comprende la molécula de ácido nucleico de cualquiera de las reivindicaciones 1-31 o el vector de la reivindicación 32.
34. Un animal transgénico no humano que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1-33.
- 40 35. El animal transgénico de la reivindicación 34, en el que el animal se selecciona del grupo que consiste en un ratón, un conejo, un camélido, una cabra y una rata.
36. Un procedimiento de preparación de un producto génico que tiene una propiedad deseada, en el que el procedimiento comprende expresar un ácido nucleico que codifica el producto génico en una población de células, en la que la población de células expresa, o puede ser inducida a que exprese, una proteína citidina desaminasa inducida por activación (AID) mutante funcional codificada por la molécula de ácido nucleico de cualquiera de las reivindicaciones 1-31, con lo cual la expresión de la proteína AID mutante funcional induce una mutación en el ácido nucleico que codifica el producto génico.
- 45 37. El procedimiento de la reivindicación 36, en el que el procedimiento comprende la etapa de seleccionar una célula o células dentro de la población que expresa la secuencia de ácido nucleico mutada que codifica el producto génico que tiene la propiedad deseada.
- 50 38. El procedimiento de la reivindicación 36 o la reivindicación 37, en el que la célula es una célula de mamífero.
39. El procedimiento de una cualquiera de las reivindicaciones 36 a 38, en el que la célula es un linfocito B o un derivado de linfocito B.
40. El procedimiento de una cualquiera de las reivindicaciones 36 a 39, en el que la célula comprende al menos una

secuencia de ácido nucleico que se ha optimizado por codones para la HMS para aumentar el número de motivos de HMS.

- 5 41. Un procedimiento para mutar un organismo no humano que tiene un fenotipo deseado que comprende expresar o inducir la expresión de una proteína citidina desaminasa inducida por activación (AID) mutante funcional codificada por la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1-31 en el organismo, con lo cual la expresión de la proteína AID mutante funcional induce una mutación dentro del ADN cromosómico del organismo.
42. El procedimiento de la reivindicación 41, en el que el procedimiento comprende la etapa de seleccionar una célula o células dentro del organismo que expresa el fenotipo deseado.
- 10 43. El procedimiento de la reivindicación 41 o la reivindicación 42, en el que el organismo comprende al menos una secuencia de ácido nucleico que se ha optimizado por codones para la HMS para aumentar el número de motivos de HMS.

FIG. 1

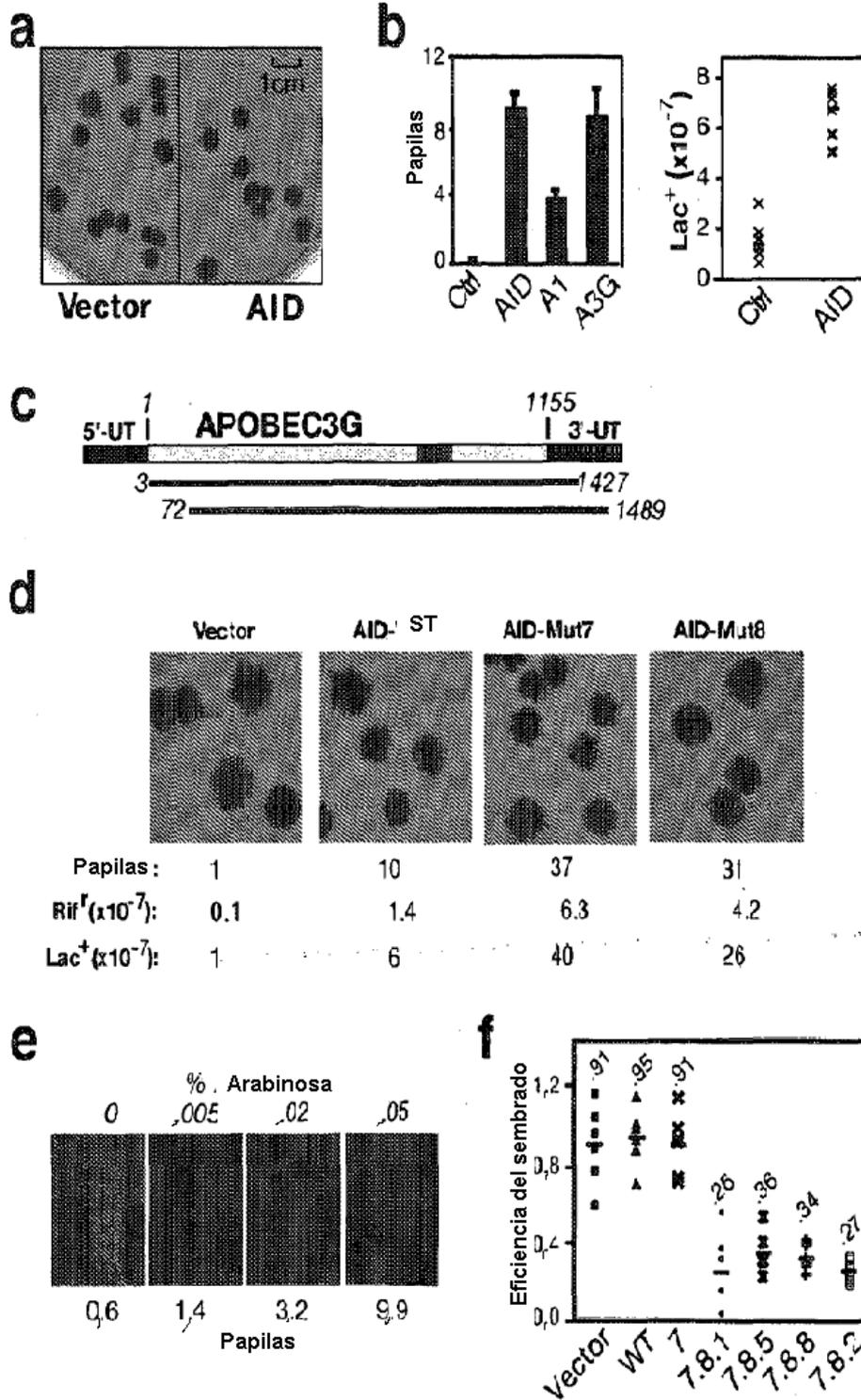


FIG. 2

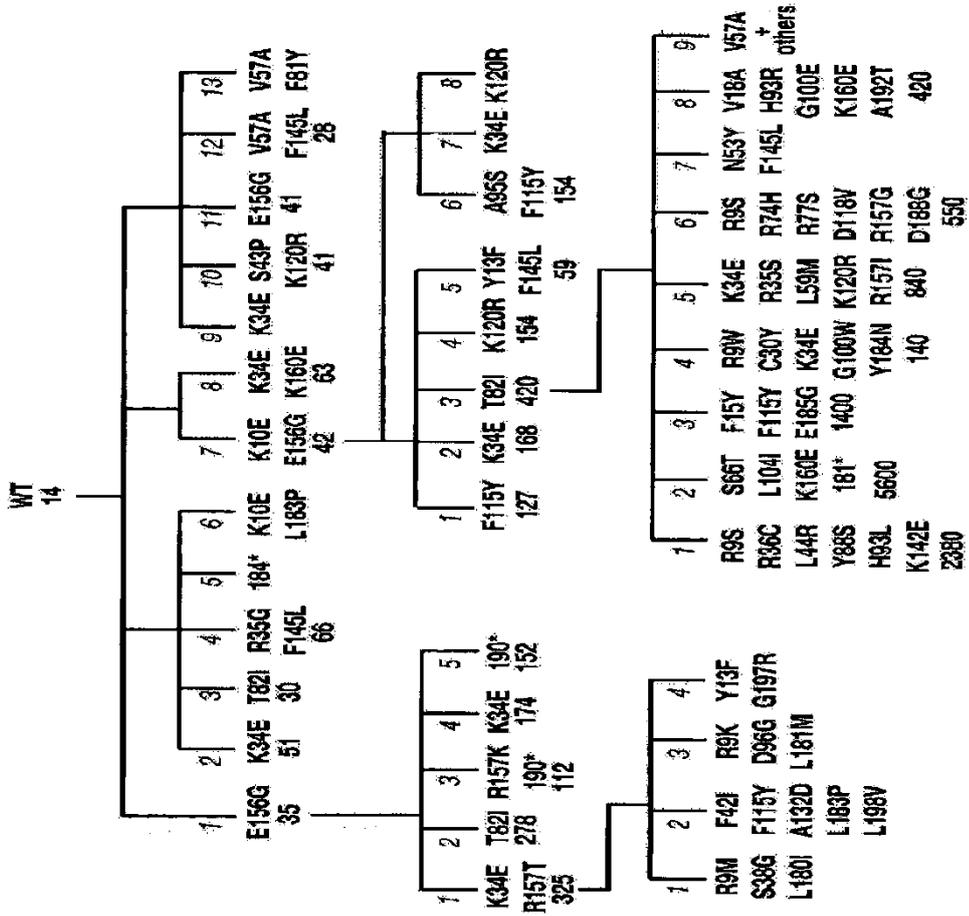
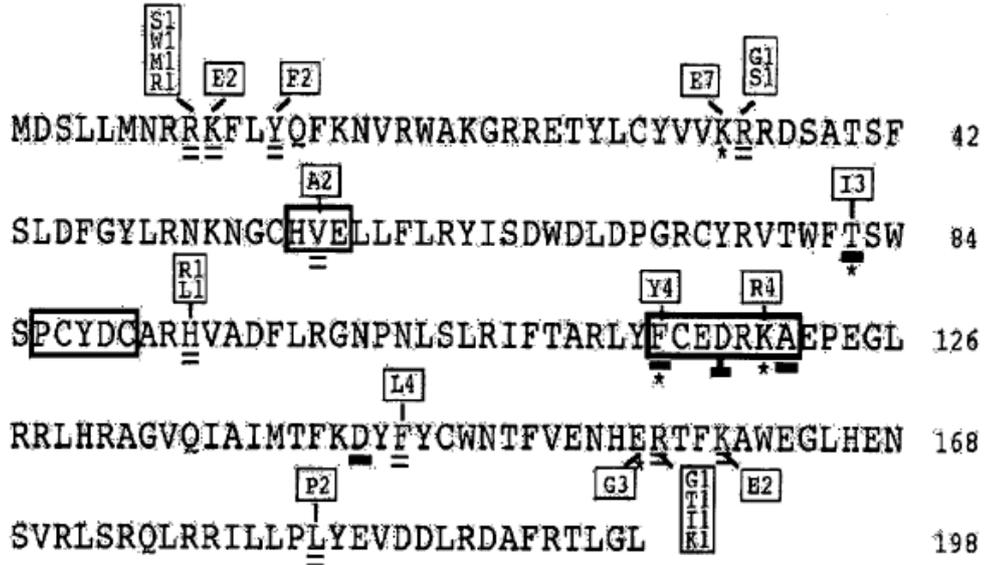
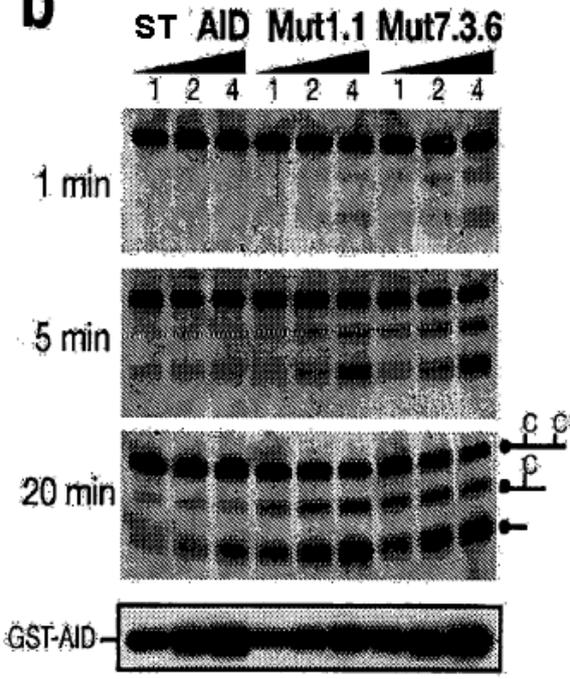


FIGURA 3

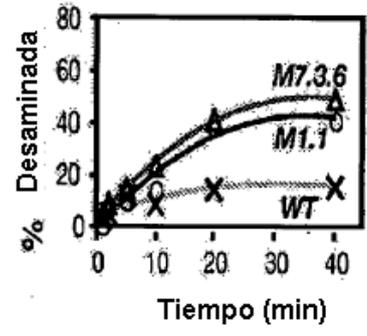
a



b



c



d

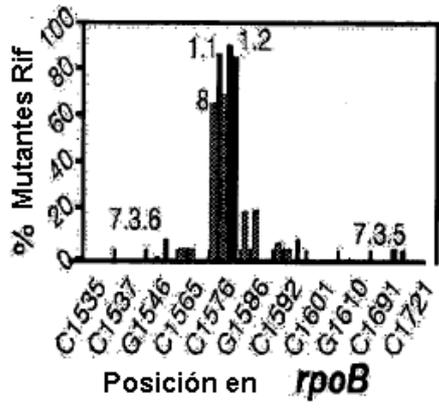


FIG. 4

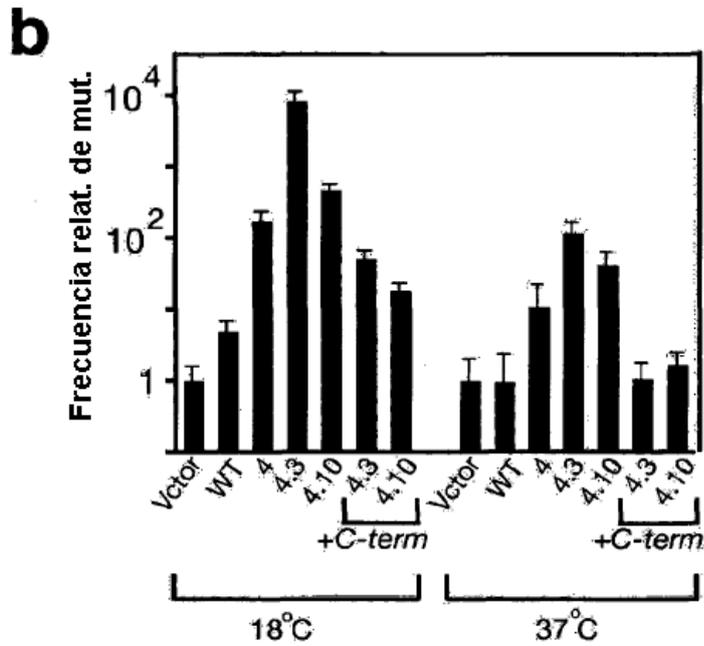
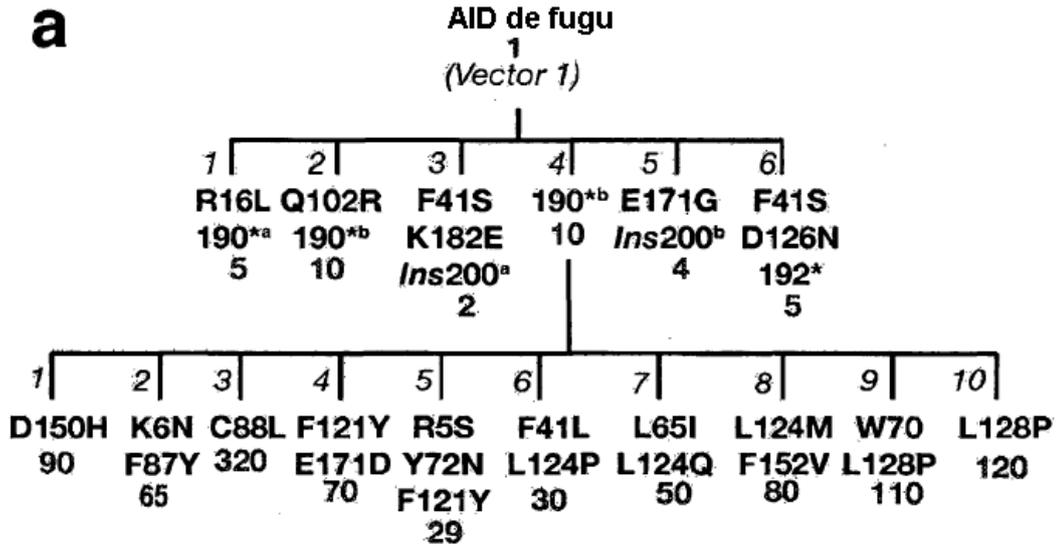


FIG. 5

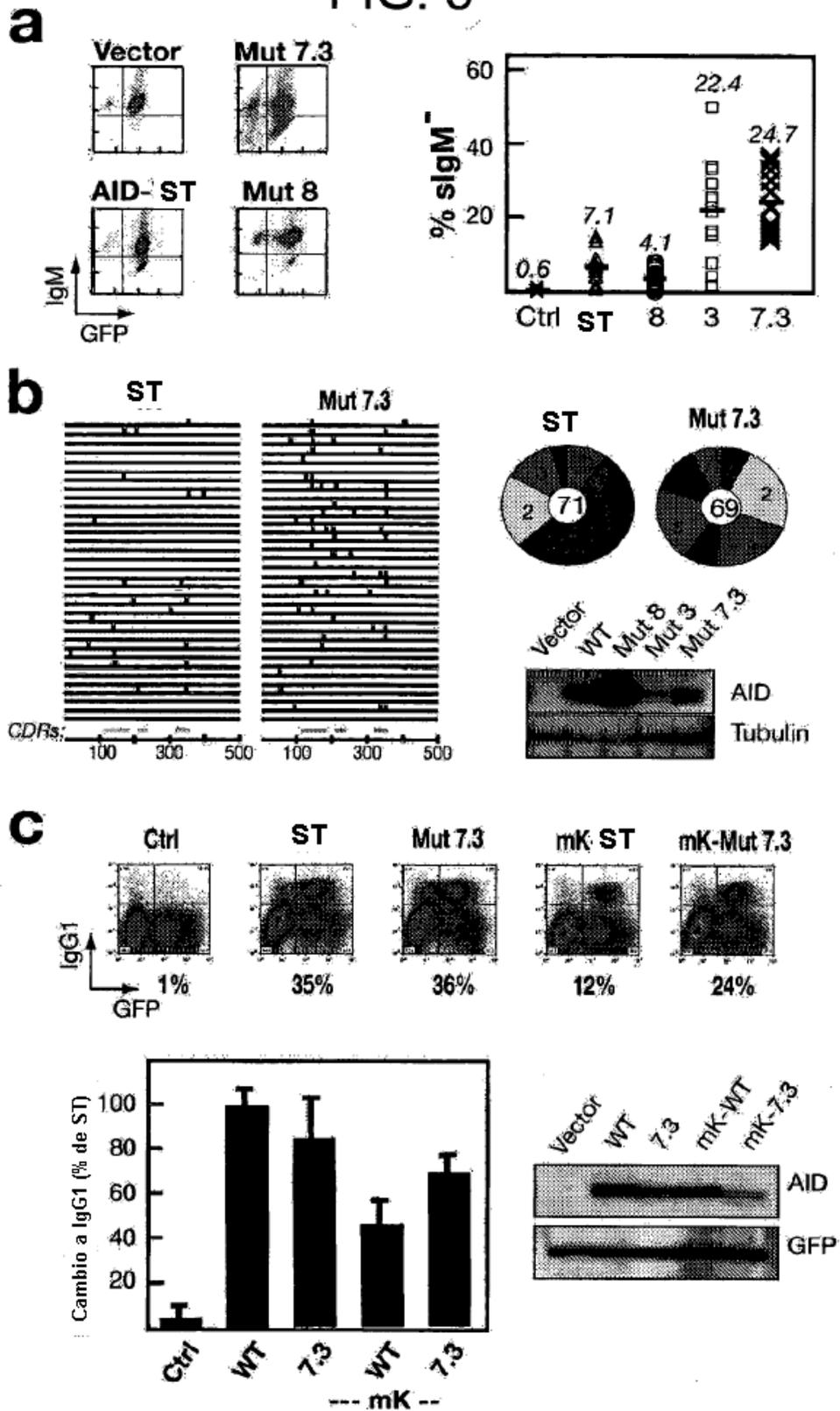


FIG. 6

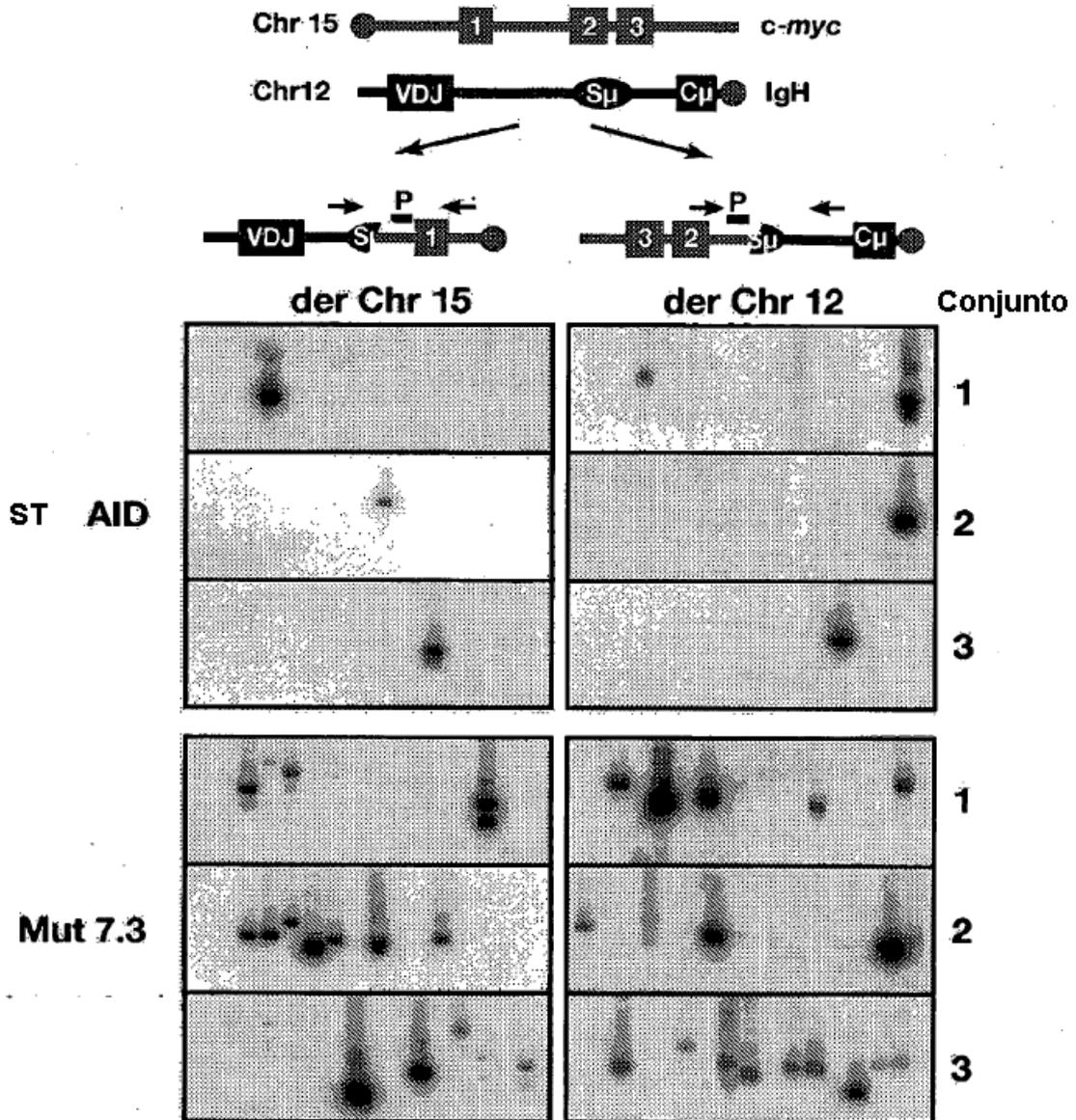


FIG. 7

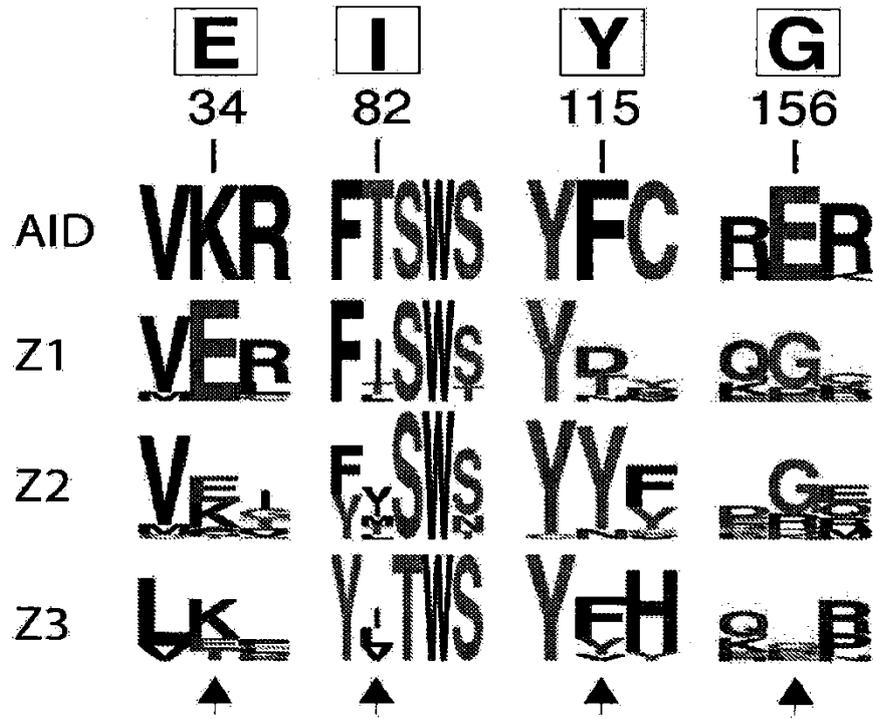


FIG. 8

Especie	Nombre común	Proteína	Número de acceso/ ID ens,
Homo sapiens	Humana	A3A (A3Z1)	NM_145699
Homo sapiens	Humana	A3B (A3Z2-Z1)	NM_004900
Homo sapiens	Humana	A3C (A3Z2)	NM_014508
Homo sapiens	Humana	A3DE (A3Z2-Z2)	NM_152426
Homo sapiens	Humana	A3F (A3Z2-Z2)	NM_145298
Homo sapiens	Humana	A3G (A3Z2-Z1)	NM_021822
Homo sapiens	Humana	A3H (A3Z3)	NM_181773
Macaca mulatta	Macaco	A3A (A3Z1)	ENSMMUG00000019046
Macaca mulatta	Macaco	A3B (Z2Z1)	XM_001117049
			XM_001117028
Macaca mulatta	Macaco	A3C (A3Z2)	NM_001114359
Macaca mulatta	Macaco	A3DE (A3Z2-Z2)	XM_001094328
Macaca mulatta	Macaco	A3F (A3Z2-Z2)	NM_001042373
Macaca mulatta	Macaco	A3G (A3Z2-Z1)	XM_001094452
Macaca mulatta	Macaco	A3H (A3Z3)	XM_001096739
Bos taurus	Vaca	A3Z1	EU864534
Bos taurus	Vaca	A3Z2	EU864535
Bos taurus	Vaca	A3Z3	EU864536
Ovis aries	Oveja	A3Z1	EU864541
Ovis aries	Oveja	A3Z2	EU864542
Ovis aries	Oveja	A3Z3	EU864543
Sus scrofa	Cerdo	A3Z2	EU864539
Sus scrofa	Cerdo	A3Z3	EU864540
Tayssu tajacu	Pecarí	A3Z2-Z3	EU864537
Equus caballus	Caballo	A3Z1	XM_001499871
Equus caballus	Caballo	A3Z2	XM_001501833
Equus caballus	Caballo	A3Z3	XM_001501833
Felis catus	Gato	A3Z2-Z3	EF173021
Canis lupus	Perro	A3Z1	XM_847690
Canis lupus	Perro	A3Z2	AACN010393938
Canis lupus	Perro	A3Z3	XM_538369
Mus musculus	Ratón	A3Z2-Z3	NM_030255
Rattus norvegicus	Rata	A3Z2-Z3	NM_001033703
Homo sapiens	Humana	AID	NM_020661
Macaca mulatta	Macaco	AID	XM_001113641
Bos taurus	Vaca	AID	NM_001038682
Ovis aries	Oveja	AID	EE793762
Sus scrofa	Cerdo	AID	BP157753
Tayssu tajacu	Pecarí	AID	EU864538
Equus caballus	Caballo	AID	XM_001493186
Felis catus	Gato	AID	ENSFCAG00000006052
Canis lupus	Perro	AID	NM_001003380
Mus musculus	Ratón	AID	NM_009645
Rattus norvegicus	Rata	AID	XM_001060382

FIG. 9

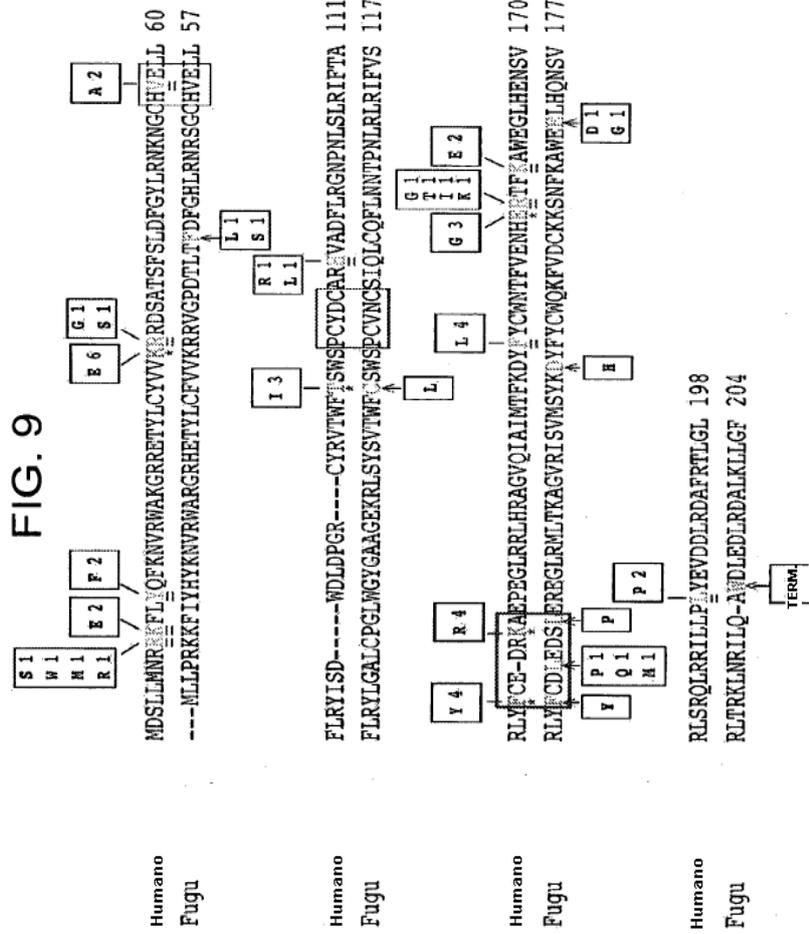


FIG. 10a

MutE  
 ATGGACTACAAAGATGACGATGATAAAGGTCCAAAGAAGAAGAGAAAAGGTAGACTCTCTCCTCAT  
 GAAGCAGAGAAAAGTT 80  
 Mut 7.3E  
 ATGGACTACAAAGATGACGATGATAAAGGTCCAAAGAAGAAGAGAAAAGGTAGACTCTCTCCTCAT  
 GAAGCAGAGAGAGTT 80  
 Human 7.3 -----  
 ATGGACAGCCTCTTGATGAACCGGAGGGAGTT 32

MutE  
 TCTCTACCAC TTC AAGAACGT CAGATGGGCCAAGGGGAGACATGAGACCTATCTCTGTTACGTCG  
 TCAAGAGGAGAGACT 160  
 Mut 7.3E  
 TCTCTACCAC TTC AAGAACGT CAGATGGGCCAAGGGGAGACATGAGACCTATCTCTGTTACGTCG  
 TCAAGAGGAGAGACT 160  
 Human 7.3  
 TCTTTACCAATTC AAAAATGTCCGCTGGGCTAAGGGTTCGGCGTGAGACCTACCTGTGCTACGTAG  
 TGAAGAGGCGTGACA 112

MutE  
 CAGCCACCTCTTTCTCCCTCGACTTTGGGCATCTCCGGAACAAGTCTGGGTGTCATGTCGAACTC  
 CTCTTCCTCCGCTAT 240  
 Mut 7.3E  
 CAGCCACCTCTTTCTCCCTCGACTTTGGGCATCTCCGGAACAAGTCTGGGTGTCATGTCGAACTC  
 CTCTTCCTCCGCTAT 240  
 Human 7.3  
 GTGCTACATCCTTTTCACTGGACTTTGGTTATCTTCGCAATAAGAACGGCTGCCACGTGGAATTG  
 CTCTTCCTCCGCTAC 192

MutE  
 ATCTCAGACTGGGACCTCGACCCCGGGAGATGCTATAGAGTCACTTGGTTTACCTCTTGGTCCCC  
 CTGTTATGACTGCCG 320  
 Mut 7.3E  
 ATCTCAGACTGGGACCTCGACCCCGGGAGATGCTATAGAGTCACTTGGTTTATCTCTTGGTCCCC  
 CTGTTATGACTGCCG 320  
 Human 7.3  
 ATCTCGEACTGGGACCTAGACCCTGGCCGCTGCTACCGCGTCACTGGTTTCACTCTCTGGAGCCC  
 CTGCTACGACTGTGC 272



MutE  
 CAGACATGTCGCCGACTTCCTCAGGGGGTATCCCAATCTCTCCCTCCGCATATTCGCCGCCCGAC  
 TCTATTTTTGTGAGG 400  
 Mut 7.3E  
 CAGACATGTCGCCGACTTCCTCAGGGGGTATCCCAATCTCTCCCTCCGCATATTCGCCGCCCGAC  
 TCTATTTTTGTGAGG 400  
 Human 7.3  
 CCGACATGTGGCCGACTTTCTGCGAGGGAAACCCCAACCTCAGTCTGAGGATCTTCACCGCGCGCC  
 TCTACTTCTGTGAGG 352

FIG. 10a (continuación)

MutE  
 ACAGGAAAGCCGAGCCCGAGGGGCTCAGGAGACTCCACC GGGCCGG GGTCCAGATCGCCATCATG  
 ACATTTAAGGACTAT 480  
 Mut 7.3E  
 ACAGGAAAGCCGAGCCCGAGGGGCTCAGGAGACTCCACC GGGCCGGGGTCCAGATCGCCATCATG  
 ACATTTAAGGACTAT 480  
 Human 7.3  
 ACCGCAAGGCTGAGCCCGAGGGGCTGCGGCGGCTGCACC GCGCCGGGGTGCAAATAGCCATCATG  
 ACCTTCAAAGATTAT 432

MutE  
 TTCTATTGTTGGAATACATTTGTCGAGAATCGGGAGAAGACTTTCAAAGCCTGGGAGGGGCTCCA  
 TGAGAACTCTGTCAG 560  
 Mut 7.3E  
 TTCTATTGTTGGAATACATTTGTCGAGAATCGGGGAAGACTTTCAAAGCCTGGGAGGGGCTCCA  
 TGAGAACTCTGTCAG 560  
 Human 7.3  
 TTTTACTGCTGGAATACTTTTGTAGAAAACACGGAAGAAGACTTTCAAAGCCTGGGAAGGGCTGCA  
 TGAAAATTCAGTTCG 512

MutE  
 ACTCTCTAGGCAGCTCAGGAGAAATCCTCCTCCCCCTCTATGAGGTCGAAGAACTCAGAGAAGCCT  
 TCCGGATCCTCGGGG 640  
 Mut 7.3E  
 ACTCTCTAGGCAGCTCAGGAGAAATCCTCCTCCCCCTCTATGAGGTCGAAGAACTCAGAGAAGCCT  
 TCCGGATCCTCGGGG 640  
 Human 7.3  
 TCTCTCCAGACAGCTTCGGCGCATCCTTTTGCCCCCTGTATGAGGTTGATGACTTACGAGACGCAT  
 TTCGTACTTTGGGAC 592

MutE	CTTGA	645	(SEC ID N° 88)
Mut 7.3E	CTTGA	645	(SEC ID N° 89)
Humano 7.3	TTTGA	597	(SEC ID N° 90)

FIG.10b

MutE MDYKDDDDKGPKKKRVDSLLMKQRKFLYHFKNVRWAKGRHETYL CYVVKRRDSATSFSLDFGHL  
 RNKSGCHVLLFLRY 238

Mut 7.3E MDYKDDDDKGPKKKRVDSLLMKQREFLYHFKNVRWAKGRHETYL CYVVKRRDSATSFSLDFGHL  
 RNKSGCHVLLFLRY 238

Humano 7.3 -----  
 MDSLLMNRREFLYQFKNVRWAKGRRETYL CYVVKRRDSATSFSLDFGYLRNKNKSGCHVLLFLRY  
 190

MutE ISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFAARLYFCEDRKAEP EGLRRL  
 HRAGVQIAIMTFKDY 478

Mut 7.3E ISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFAARLYFCEDRKAEP EGLRRL  
 HRAGVQIAIMTFKDY 478

Human 7.3 ISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYFCEDRKAEP EGLRRL  
 HRAGVQIAIMTFKDY 430

MutE FYCWNTFVENREKTFKAW EGLHENS VRLSRQLRRILL PLYEVEELREAFRILGA.  
 643 (SEC ID N° 91)

Mut 7.3 FYCWNTFVENRGKTFKAW EGLHENS VRLSRQLRRILL PLYEVEELREAFRILGA.  
 643 (SEC ID N° 92)

Human 7.3 FYCWNTFVENHGRTFKAW EGLHENS VRLSRQLRRILL PLYEVDLDRDAFRITLGL.  
 595 (SEC ID N° 93)