

(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **2 440 777**

(51) Int. Cl.:

**C07K 16/00** (2006.01)  
**C07K 16/24** (2006.01)  
**C12N 15/63** (2006.01)  
**C12N 15/10** (2006.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **21.12.2005 E 05855447 (8)**

(97) Fecha y número de publicación de la concesión europea: **23.10.2013 EP 1839120**

---

(54) Título: **Vectores basados en anticuerpo anti-il-12, células huéspedes y métodos de producción y usos**

(30) Prioridad:

**21.12.2004 US 637936 P**

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

**30.01.2014**

(73) Titular/es:

**JANSSEN BIOTECH, INC. (100.0%)**  
800/850 Ridgeview Drive  
Horsham, PA 19044 , US

(72) Inventor/es:

**LU, JIN;**  
**NESSPOR, THOMAS;**  
**SCALLON, BERNARD y**  
**SNYDER, LINDA**

(74) Agente/Representante:

**IZQUIERDO FACES, José**

**ES 2 440 777 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

Vectores basados en anticuerpo anti-il-12, células huéspedes y métodos de producción y usos

5 **Campo de la invención**

La presente invención se refiere a vectores y plásmidos que dirigen expresiones de un anticuerpo, células huéspedes y métodos para hacer uso de los mismos, incluyendo secuencias potenciadoras y promotoras del vector específico y su interacción con los factores de transcripción de célula huésped.

10

**Antecedentes**

Las moléculas de anticuerpo consisten en una combinación de dos polipéptidos de cadena pesada (P) y dos de cadena ligera (L). Cada cadena pesada y ligera comprende una región constante que contiene las regiones CL, CH1, región bisagra, CH2 y CH3, y una región variable que contiene las regiones hipervariables (regiones determinante del complemento (RDCs)); las RDCs controlan las características de enlace de antígeno del anticuerpo. Las dos cadenas pesadas se unen y las cadenas ligeras en una estructura en forma de Y por medio de puentes disulfuros de tal manera que las regiones variables de las cadenas ligeras (V.<sub>sub</sub> L.) y cadenas pesadas (V.<sub>sub</sub> P.) están situadas unas junto a otras.

20

Para generar anticuerpos, se han usado técnicas convencionales de hibridoma en las que los clones de las células híbridas que expresan genes que codifican las cadenas ligeras y pesadas de una molécula de anticuerpo se obtienen mediante inmunización con una molécula de anticuerpo. Esta técnica necesita la fusión de células de origen linfocítico, que contienen los genes para la formación de anticuerpos y células que forman líneas inmortales. Las células que tienen los genes en cuestión se obtienen generalmente mediante creación arbitraria de bibliotecas de células circulantes, y la filtración de los hibridomas con una reacción antígeno-anticuerpo después de que los clones del hibridoma se hayan multiplicado y cultivado. Esta técnica puede ser incierta y laboriosa con una producción limitada de anticuerpos, y se limita a aplicaciones para la producción de anticuerpos no humanos (por ejemplo, ratón).

30

Además, los anticuerpos monoclonales y sus fragmentos pueden expresarse en varios sistemas huéspedes, tales como levadura *E. coli*, y células huéspedes de mamíferos. En general, un vector de expresión de mamífero contendrá (1) elementos reguladores, normalmente en forma de secuencias promotoras y potenciadoras virales y caracterizado por una amplio rango de huésped y tejido; (2) una secuencia "policlonaje", que facilita la inserción de un fragmento de ADN en el vector plásmido; y (3) las secuencias responsables para el empalme de intrón y la poliadenilación de transcripciones de ARN. Esta región continua del sitio promotor-policlonaje-poliadenilación es referida comúnmente como la unidad de transcripción. El vector probablemente también tendrá (4) un gen o genes marcadores seleccionables (por ejemplo, el gen beta-lactamasa), que a menudo confieren resistencia a un antibiótico (tal como ampicilina), que permite la selección de transformadores positivos iniciales en *E. coli*; y (5) secuencias que facilitan la réplica del vector en huéspedes bacterianos y mamíferos.

40

A diferencia de la mayoría de los genes que se transcriben a partir de secuencias continuas genómicas de ADN, los genes de anticuerpo se montan a partir de segmentos de gen que pueden separarse comúnmente en la línea del germen. En particular, los genes de cadena pesada se forman mediante recombinación de tres segmentos genómicos que codifican las regiones variables (V), de diversidad (D) y unión (U)/constante (C) del anticuerpo. Los genes de cadena ligera funcional se forman uniendo dos segmentos de gen; uno codifica la región V y el otro codifica la región U/C. Los locus de la cadena pesada y la cadena ligera .kappa. contienen muchos segmentos de gen V (los cálculos aproximados varían entre 100s y 1000s) estimados para abarcar más de 1000 kb. El locus .lambda. es, en cambio, mucho más pequeño y ha demostrado abarcar aproximadamente 300 kb en el cromosoma 16 en el ratón. Consiste en cuatro segmentos de gen de unión/constantes y dos segmentos de gen variables. La recombinación que da como resultado los genes funcionales ocurre predominantemente entre elementos V.<sub>sub</sub>1 y U.<sub>sub</sub>1/C.<sub>sub</sub>1 o U.<sub>sub</sub>3/C.<sub>sub</sub>3 o entre elementos V.<sub>sub</sub>2 y U.<sub>sub</sub>2/C.<sub>sub</sub>2 (U.<sub>sub</sub>4 / C.<sub>sub</sub>4 es un pseudógeno), aunque las recombinaciones entre V.<sub>sub</sub>2 y U.<sub>sub</sub>3/C.<sub>sub</sub>3 o U.<sub>sub</sub>1/C.<sub>sub</sub>1 se ven muy raramente.

55

Un ejemplo de un vector de expresión de mamífero es CDM8. La unidad de transcripción de CDM8 está compuesta por un promotor químérico (el promotor constitutivo AD169 de citomegalovirus humano fusionado con el promotor de polimerasa T7 ARN), una región de policlonaje y el empalme de antígeno de tumor pequeño (t) SV40 y las señales de poliadenilación de región temprana derivadas de pSV2. El promotor de citomegalovirus humano (HCMV) se expresa en una variedad de tipos de células de mamíferos, mientras que el promotor de polimerasa de ARN dependiente de ADN bacteriófago T7 puede moverse a la transcripción/traslación in vitro libre de células de inserciones clonadas. Esta fusión promotora particular permite que los experimentos iniciales se realicen dentro de los confines del tipo de célula de mamífero huésped, mientras que el análisis y utilización adicional de la inserción clonada puede realizarse potencialmente en un sistema de transcripción/traslación in vitro libre de células. El promotor HCMV constitutivamente expresado también se ha utilizado en otros vectores de expresión mamífera además de CDM8. Los orígenes de réplica en CDM8 incluyen (1).pi.VX (permitiendo, por ejemplo, réplica en *E. coli*) (2) origen SV40 (por ejemplo, permitiendo la réplica en una variedad de tipos de célula COS) (3) origen de poliomia

(por ejemplo, permitiendo réplica en fibroblastos de ratón transformados por virus de polioma) y (4) el origen de bacteriófago M13 (por ejemplo, permitiendo la generación de plantilla de hebra sencilla para análisis de secuencia de ADN y/o mutagénesis de oligonucleótido dirigida al sitio).

5 Además, CDM8 tiene el gen supF para la selección en *E. coli*. En este sistema de selección antibiótica, una construcción de plásmido basada en CDM8 se transforma a cepa *E. coli* especializada que contiene un episoma que transporta genes que codifican la resistencia a los antibióticas, ampicilina y tetracilina. Sin embargo, ambos genes contienen mutaciones de punto de terminación de cadena (codón "sin sentido") que inactivan el fenotipo de resistencia. El producto de gen supF, un TARN supresor sin sentido, restaura el fenotipo resistente para cada 10 antibiótico. Por lo tanto, la selección se basa en el crecimiento de la cepa *E. coli* que transporta el episomal especializado en el medio que contiene ampicilina y tetracilina. Las colonias que muestran este fenotipo se transforman supuestamente con la construcción de plásmido basada en CDM8.

15 El vector CDM8 es compatible con líneas celulares COS así como líneas celulares transformadas con el virus polioma. Las líneas celulares COS son células CV1 de mono verde africano transformadas con un virus mutante SV40 defectuoso de origen. Las células COS producen el antígeno T grande, que se requiere en trans para promover la réplica de SV40 o construcciones de plásmido, tales como CDM8, que contienen las respectivas 20 secuencias cis-actoras que inician la réplica viral. Por lo tanto, las células COS transfectadas con una construcción basada en CDM8 soportará la réplica del plásmido, dando como resultado un mayor número de copia de plásmido y un sobreexpresión temporal del gen de interés.

25 El principal uso de CDM8 es la clonación de expresión de cADN y la sobreproducción de proteínas específicas en un sistema de expresión in vitro de mamíferos. La clonación de expresión toma varias formas dependiendo del modo de detección utilizado para identificar el cADN de interés; sin embargo, la etapa inicial consiste en aislar mARN y sintetizar copias de ácido deoxirribonucleico de doble hebra de la población de mARN (cADNs). Estos cADNs deber estar eficientemente ligados a un plásmido o vector de clonación de ADN bacteriófago y transferirse al huésped apropiado antes de la filtración y el análisis de la biblioteca. El vector CDM8 contiene dos 30 sitios de restricción BstXI, haciéndolo tratable para el procedimiento del enlace "adaptador" de ligar cADNs al vector, es decir, el uso de fragmentos de ADN con extremo desafilado en un extremo (y por lo tanto compatibles con la ligadura con el cADN de extremo desafilado) pero que contienen un saliente no palindrómico (extremo engomado) en el otro extremo (por ejemplo, compatible con ligadura con vector de ADN digerido con BstXI, pero no con otros cADNs).

35 Otro ejemplo de vector de expresión de mamífero es pCMX. Este vector contiene: (1) el promotor temprano inmediato de HCMV, (2) una secuencia empalme/poliadenilación SV40 ARN, (3) una resistencia a un antibiótico, tal como el gen beta-lactamasa que confiere resistencia a la ampicilina antibiótica. El vector pCMX también puede usarse para la expresión temporal de una secuencia de ADN clonada en células COS transfectadas.

40 El control de la transcripción de ambos genes de cadena pesada y ligera .kappa. reorganizados depende de la actividad del promotor específico del tejido corriente arriba de la región V y del potenciador específico de tejido situado en el intrón J-C. Estos elementos actúan sinérgicamente. También, se ha identificado un segundo potenciador específico de célula B en el locus de cadena ligera .kappa. Este potenciador adicional está situado 9kb corriente debajo de C.sub..kappa.

45 Tal sistema de huésped de mamífero usado para producir anticuerpos es una célula huésped de mieloma de ratón que se ha transfectado con ADN clonado que codifica el anticuerpo deseado. Tales "anticuerpos monoclonales recombinantes" a menudo son distintos de los anticuerpos monoclonales derivados de hibridoma para los que el ADN no se ha clonado y para los que las células que producen el anticuerpo monoclonal se derivan inmortalizando una célula que produce un anticuerpo monoclonal natural aislada de un animal. Los genes de inmunoglobulina de cadena pesada y ligera (Ig) que se expresan en células de hibridoma están bajo el control del promotor endógeno natural que siempre se ha unido con la secuencia de la región variable particular que se expresa en oposición al promotor contenido en el vector recombinante.

55 En producción recombinante, la secuencia de anticuerpo monoclonal que se clonará debe ligarse en un vector apropiado después del tratamiento de enzima de restricción del vector. Esta tarea puede ser difícil e imprecisa como el proceso de incorporar la secuencia o secuencias de nucleótido de anticuerpo en un vector de expresión o plásmido es complejo.

60 Sin embargo, al clonar las secuencias de ADN de anticuerpo monoclonal antes de preparar los anticuerpos monoclonales derivados de célula transfectada, pueden usarse métodos de ADN recombinante para sustituir el promotor endógeno natural para un gen Ig por cualquier promotor preferido. Una razón principal para cambiar un promotor es realizar niveles más altos de producción de anticuerpo monoclonal.

65 Las secuencias promotoras, en conjunto con las secuencia potenciadoras corriente abajo, son responsables de dirigir la transcripción (es decir, síntesis de ARN) de los genes de cadena pesada y ligera en las células transfectadas al enlazarse con proteínas nucleares especializadas llamadas factores de transcripción. Es

5 evidente que hay menos sitios para el enlace del factor de transcripción en un promotor Ig que en un potenciador Ig; sin embargo, el hecho de que hay variabilidad de secuencia entre promotores pero solamente una única copia de la secuencia promotora hace que sea muy probable que haya variabilidad funcional entre promotores Ig. Un promotor puede ser “fuerte”, es decir, eficiente en enlazar una combinación favorable de factores de transcripción y lleva a altos niveles de síntesis de ARN de anticuerpo monoclonal, mientras que otro promotor puede ser “débil”, debido a que tiene una secuencia diferente de ADN. Ya que cada uno de los más de 200 genes HC de región variable y los más de 200 genes de LC de región variable en un repertorio Ig tiene su propio promotor naturalmente unido, y es probable que dos promotores no tengan secuencias idénticas, es probable que los muchos promotores diferentes Ig varíen significativamente con respecto a cómo de bien llevan la transcripción.

10 10 Los promotores Ig son solamente funcionales en células huésped de tipo linfoide, tales como células T y B (y células de mieloma), debido a su requisito para factores de transcripción específicos de gen Ig (por ejemplo, Oct-2 y OBF-1) no expresados en otros tipos de células. Además, incluso los factores de transcripción específicos de células pueden expresarse solamente en fases particulares de diferenciación celular de tal manera que la expresión óptima puede depender de unir el estado de diferenciación de la línea celular huésped con los motivos apropiados de secuencia en los promotores de gen Ig. Aunque la especificidad de célula huésped de promotores Ig puede verse como una desventaja menor para la expresión del anticuerpo monoclonal en una célula huésped no linfoide, la gran selección de promotores HC y LC permite una posibilidad de identificar y quizás optimizar más los promotores fuertes que pueden incorporarse a los vectores específicos de célula linfoide.

15 20 La expresión de anticuerpos monoclonales detrás de un promotor fuerte aumenta las posibilidades de identificar líneas celulares altamente productoras y obtener mayores producciones de anticuerpos monoclonales. Como consecuencia, los vectores Ig con promotores fuertes son muy deseables para expresar cualquier anticuerpo monoclonal de interés. Además, los vectores con sitios de clonación únicos de ADN corriente debajo de los promotores fuertes tendrían una conveniencia añadida.

25 30 Por consiguiente, hay una necesidad de nuevos vectores y plásmidos útiles para la expresión de anticuerpos que simplifiquen las técnicas de ligadura y permitan la customización de secuencias promotoras y potenciadoras con el fin de aumentar la producción de anticuerpos.

35 30 **Resumen de la invención**  
40 La presente invención se refiere a vectores y plásmidos de expresión recombinante que comprenden sitios de restricción para la clonación de varios anticuerpos. En otra realización, la divulgación proporciona la expresión de secuencias de control en el vector, tales como secuencias potenciadoras y promotoras, que pueden customizarse en relación con el gen de anticuerpo que se clonará y transcribirá y el tipo de célula huésped que se usará, con el fin de llevar la transcripción eficientemente. La presente invención también comprende células huésped aisladas, por ejemplo, células de mamíferos y no mamíferos, que contienen tal vector o plásmido. La divulgación proporciona además métodos para producir un anticuerpo mediante cultivo, en un medio adecuado, una célula huésped que contiene un vector de expresión recombinante de la invención de tal manera que se produzca el anticuerpo.

45 45 En otra realización, la divulgación comprende un método para identificar, modular y/o determinar la interacción entre los factores de transcripción de célula huésped y las secuencias promotoras y potenciadoras de un vector de expresión. Esta interacción impulsa el proceso de transcripción. Los factores de transcripción y las secuencias promotoras y potenciadoras pueden customizarse para mejorar su afinidad para enlazarse entre sí, lo que puede aumentar la producción y eficiencia del proceso de transcripción.

50 55 **Descripción de las figuras**  
50 La Fig. 1A es una representación esquemática del mapa vector para el plásmido p1560.  
55 La Fig. 1B es una representación esquemática del mapa vector para el plásmido p1558.

55 **Descripción de la invención**  
60 55 Las siguientes definiciones se exponen para ilustrar y definir el significado y alcance de varios términos usados para describir la invención en el presente documento.  
60 Una “actividad” o “actividad biológica”, y una actividad funcional de un polipéptido se refieren a una actividad ejercida por una proteína o polipéptido in vivo, in situ o in vitro, de acuerdo con técnicas estándares. Tales actividades pueden ser una actividad directa, tal como una asociación con una actividad enzimática en una segunda proteína, o una actividad indirecta, tal como un proceso celular mediado por la interacción de la proteína con una segunda proteína o una serie de interacciones como en la señalización intracelular o la cascada de coagulación.

65 65 Un “anticuerpo” incluye cualquier polipéptido o péptido que contiene molécula que comprende al menos una parte de una molécula de inmunoglobulina, tal como aunque sin limitar, al menos una región determinante de la

complementariedad (RDC) de una cadena pesada o ligera o una parte de enlace de ligando de la misma, una región variables de cadena pesada o cadena ligera, una región constante de cadena pesada o cadena ligera, una región marco, o una parte, fragmento o variante de la misma. El término "anticuerpo" pretende además abarcar anticuerpos, fragmentos de digestión, partes especificadas y variantes de los mismos, incluyendo miméticos de anticuerpo o que comprende partes de anticuerpos que imitan la estructura y/o función de un anticuerpo o fragmento especificado o parte del mismo, incluyendo anticuerpos de cadena sencilla y fragmentos de los mismos. Por ejemplo, los fragmentos de anticuerpo incluyen, aunque no se limitan a, Fab (por ejemplo, por digestión de papaína), Fab' (por ejemplo, por digestión de pepsina y reducción parcial) y F(ab')2 (por ejemplo, por digestión de pepsina, reducción parcial y reagregación), fragmentos de Fv o scFv (por ejemplo, mediante técnicas de biología molecular), están incluidas en la invención (véase, por ejemplo, Colligan, et al., eds., *Current Protocols in Immunology*, John Wiley & Sons, Inc., NY (1994-2001); Colligan et al., *Current Protocols in Polypeptide Science*, John Wiley & Sons, NY, NY (1997-2001)).

Moléculas "químéricas" o de " fusión" son ácidos nucleicos o polipéptidos que se crean combinando uno o más polinucleótidos (o sus partes) con secuencias adicionales de ácido nucleico. Tales secuencias combinadas pueden introducirse en un vector apropiado y expresarse para dar lugar a un polipéptido químérico o de fusión.

"Complemento de" o "complementario a" una secuencia de ácido nucleico de la invención se refiere a una molécula de polinucleótido que tiene una secuencia base complementaria y una orientación inversa en comparación con un primer polinucleótido.

"Fragmento" es un polipéptido variante que tiene una secuencia de aminoácido que es completamente la misma como parte pero no toda de cualquier secuencia de aminoácido de un polipéptido o un polinucleótido variante que tiene una secuencia de ácido nucleico que es completamente la misma como parte pero no toda de cualquier secuencia de ácido nucleico de cualquier polipéptido. Los fragmentos pueden incluir, por ejemplo, polipéptidos de truncamiento que tienen una parte de una secuencia de aminoácido, o de variantes de la misma, tales como una serie continua de residuos que incluyen una secuencia de aminoácido heteróloga amino- y/o carboxi-terminal. Las formas de degradación de los polipéptidos producidos por o en una célula huésped también se incluyen. Otros fragmentos ejemplares se caracterizan por los atributos estructurales o funcionales, tales como fragmentos que comprenden hélice alfa o regiones que forman hélice alfa, lámina beta o regiones que forman lámina beta, giro o regiones que forman giro, bobina o regiones que forman bobina, regiones hidrofílicas, regiones hidrofóbicas, regiones alfa-anfipáticas, regiones beta-anfipáticas, regiones flexibles, regiones que forman superficie, regiones que se enlazan con sustrato, regiones extracelulares, y regiones de alto índice antigenico.

Más fragmentos ejemplares incluyen un polipéptido aislado que comprende una secuencia de aminoácido que tiene al menos 15, 20, 30, 40, 50 o 100 aminoácidos contiguos de una secuencia de aminoácido de longitud completa, o un polipéptido aislado que comprende una secuencia de aminoácido que tiene al menos 15, 20, 30, 40, 50 o 100 aminoácidos contiguos truncados o eliminados de la secuencia de aminoácido de longitud completa. Los fragmentos también incluyen polinucleótidos aislados que tienen tamaños o características similares.

"Identidad", como se conoce en la técnica, es una relación entre dos o más secuencias de polipéptido o dos o más secuencias de polinucleótido, como se determina comparando las secuencias. En la técnica, "identidad" también significa el grado de relación secuencial entre secuencias e polipéptido o polinucleótido, como lo determina la coincidencia entre series de tales secuencias. "Identidad" y "similitud" pueden calcularse fácilmente mediante métodos conocidos, incluyendo, aunque sin limitar, a aquellos descritos en *Computational Molecular Biology*, Lesk, A. M., ed., Oxford University Press, Nueva York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D. W., ed., Academic Pres, Nueva York, 1993; *Computer Analysis of Sequence Data*, Parte I, Griffin, A. M., y Griffin, H. G., eds., Humana Press, Nueva Jersey, 1994; *Sequence Analysis in Moleculr Biologay*, von Heinje, G., Academic Press, 1987; y *Sequence Analysis Primer*, Gribskov, M., y Devereus, J. Eds., M Stockton Press, Nueva York, 1991; y Carillo, H., y Lipman, D., Siam J. Applied Math., 48:1073 81988). Además, los valores para la identidad porcentual pueden obtenerse a partir de alineaciones de aminoácidos y secuencias nucleótidos generadas usando los ajustes estándares para el componente AlignX de Vector NTI Suite 8.0 (Informax, Frederick, MD).

Los métodos preferentes para determinar la identidad están diseñados para dar la mayor coincidencia entre las secuencias analizadas. Los métodos para determinar identidad y similitud se codifican en programas de ordenador públicamente disponibles. Los métodos para programas de ordenador preferentes para determinar identidad y similitud entre dos secuencias incluyen, aunque no se limitan a, el paquete de programa GCG (Devereux, J., et al., *Nucleic Acids Research* 12(1):387 (1984)); BLASTN y FASTA (Altschul, S. F. et al., *J. Molec. Biol.* 215:403-410 (1990). El programa BLAST X está públicamente disponible en NCBI y otras fuentes (Manual BLAST, Altschul, S., et al., NCBINLM NIH Besteda, Md. 20894: Altschul, S. et al., *J. Mol. Biol.* 215:403-410 (1990). El algoritmo bien conocido de Smith Waterman también puede usarse para determinar identidad.

Los parámetros preferentes para la comparación de secuencia de polipéptidos incluyen los siguientes:  
(1) Algoritmo: Needleman y Wunsch, *J. Mol Biol.* 48:443-453 (1970)  
Matriz de comparación: BLOSSUM62 de Hentikoff y Hentikoff, *Proc. Natl. Acad. Sci. USA.* 89:10915-10919 (1992)  
Penalización en la puntuación del alineamiento por cada hueco: 12

Penalización en la puntuación del alineamiento por cada hueco en longitud: 4

Un programa útil con estos parámetros está públicamente disponible como el programa "gap" de Genetics Computer Group, Madison Wis. Los parámetros anteriormente mencionados son los parámetros estándares para las comparaciones de secuencia de péptido (junto con las faltas de penalización para los huecos finales).

- 5 Los parámetros preferentes para la comparación de polinucleótidos incluyen los siguientes:  
 (1) Algoritmo: Needleman y Wunsch, J. Mol Biol. 48:443-453 (1970)  
 Matriz de comparación: coincidencias=+10, sin coincidencia=0
- 10 Penalización en la puntuación del alineamiento por cada hueco: 50  
 Penalización en la puntuación del alineamiento por cada hueco en longitud: 3  
 Disponible como: el programa "gap" de Genetics Computer Group, Madison Wis. Estos son parámetros estándares para comparaciones de secuencia de ácido nucleico.
- 15 A modo de ejemplo, una secuencia de polinucleótido puede ser idéntica a otra secuencia, que es 100% idéntica, o puede incluir hasta ciertos números enteros de alteraciones de nucleótidos en comparación con la secuencia referencia. Tales alteraciones se seleccionan del grupo consistente en al menos una eliminación de nucleótido, sustitución, incluyendo transición o transversión, o inserción, y donde las alteraciones ocurren en las posiciones terminales 5' o 3' de la secuencia de nucleótido referencia o en cualquier sitio entre esas posiciones terminales, intercaladas individualmente entre los nucleótidos en la secuencia referencia o en uno o más grupos contiguos en la secuencia referencia. El número de alteraciones de nucleótidos se determina multiplicando el número total de nucleótidos en la secuencia por el porcentaje numérico de la respectiva identidad porcentual (dividida entre 100) y restando el producto del número total de nucleótidos en la secuencia, o:  

$$n.\text{sub.}n.\text{Itorsim.}x.\text{sub.}n-(x.\text{sub.}n.y),$$
 donde  $n.\text{sub.}n$  es el número de alteraciones de nucleótido,  $x.\text{sub.}n$  es el número total de nucleótidos en la secuencia, e y es, por ejemplo 0,70 para 70%, 0,80 para 80%, 0,85 para 85%, 0,90 para 90%, 0,95 para 95%, etc., donde cualquier producto no entero de  $x.\text{sub.}n$  e y se redondea hacia abajo al número entero más cercano para restar de  $x.\text{sub.}n$ .
- 20 30 Las alteraciones de una secuencia de polinucleótido que codifica un polipéptido pueden crear mutaciones sin sentido, de sentido erróneo o de cambio de estructura en esta secuencia codificadora y por lo tanto alterar el polipéptido codificado por el polinucleótido seguido de tales alteraciones. Similarmente, una secuencia de polipéptido puede ser idéntica a una secuencia de referencia, es decir 100% idéntica, o puede incluir hasta cierto número entero de alteraciones de aminoácido en comparación con la secuencia de referencia de tal manera que la identidad porcentual sea inferior al 100%. Tales alteraciones se seleccionan del grupo consistente en al menos una eliminación de aminoácido, sustitución, incluyendo sustitución conservadora y no conservadora, o inserción, y donde las alteraciones pueden ocurrir en la posición amino- o carboxi-terminal de la secuencia de polipéptido de referencia o en cualquier sitio entre esas posiciones terminales, intercaladas individualmente entre los aminoácidos en la secuencia referencia o en uno o más grupos contiguos en la secuencia referencia. El número de alteraciones de aminoácidos para una identidad% dada se determina multiplicando el número total de aminoácidos en la secuencia por el porcentaje numérico de la respectiva identidad porcentual (dividida entre 100) y después restando ese producto del número total de aminoácidos en la secuencia, o:  

$$n.\text{sub.}a.\text{Itorsim.}x.\text{sub.}a-(x.\text{sub.}a.y),$$
 donde  $n.\text{sub.}a$  es el número de alteraciones de aminoácido,  $x.\text{sub.}a$  es el número total de aminoácidos en la secuencia, e y es, por ejemplo 0,70 para 70%, 0,80 para 80%, 0,85 para 85%, etc., donde cualquier producto no entero de  $x.\text{sub.}a$  e y se redondea hacia abajo al número entero más cercano para restar de  $x.\text{sub.}a$ .
- 25 35 40 45 50 55 60 65 "Ácidos nucleicos" son polímeros de nucleótidos, donde un nucleótido comprende una base unida a un azúcar cuyos azúcares a su vez se unen entre sí mediante al menos una molécula bivalente, tal como un ácido fosfórico. En ácidos nucleicos que ocurren de manera natural, el azúcar es 2'-deoxirribosa (ADN) o ribosa (ARN). Los poli- u oligonucleótidos no naturales contienen bases modificadas, azúcares, o molécula de enlace, pero se entiende que generalmente imitan la naturaleza complementaria de los ácidos nucleicos que ocurren de manera natural a los que se han diseñado. Un ejemplo de oligonucleótido no natural es una composición de molécula anti-sentido que tiene una red central de fosforotiorato. Un "oligonucleótido" generalmente se refiere a ácidos nucleicos que tienen menos de 30 nucleótidos.
- Una "polipéptido" es un polímero de residuos de aminoácidos unidos por enlaces peptídicos, y un péptido se refiere generalmente a polímeros de aminoácido de 12 o menos residuos. Los enlaces peptídicos pueden producirse naturalmente como los dirigidos por la plantilla de ácido nucleico o sintéticamente por métodos bien conocidos en la técnica.
- Una "proteína" es una macromolécula que comprende una o más cadenas de polipéptido. Una proteína puede además comprender grupos sustituyentes unidos a los grupos laterales de los aminoácidos implicados en la formación de los enlaces peptídicos. Típicamente, las proteínas formadas por expresión de célula eucariótica también contienen carbohidratos. Las proteínas se definen en el presente documento en términos de su secuencia o red central de aminoácido y los sustituyentes no se especifican, sean conocidos o no.

5 El término "receptor" denota una molécula que tiene actividad biológica resultante de la interacción con un ligando específico o pareja de enlace. Los receptores unidos a la membrana celular se caracterizan por un dominio de enlace de ligando extracelular, uno o más dominios que cubren la membrana o transmembrana, y un dominio efector intracelular que está típicamente implicado en transducción de señal. En enlace de ligando a los receptores de la membrana celular provoca cambios en el dominio extracelular que se comunican con la membrana celular, interacción directa o indirecta con una o más proteínas intracelulares, y altera las proteínas celulares, tales como actividad enzimática, forma celular o perfil de expresión de gen. Los receptores también pueden estar liberados a la superficie celular y pueden ser citosólicos, nucleares o liberarse de la célula juntos. Los receptores asociados no celulares se llaman receptores solubles.

10 10 Todas las publicaciones o patentes citadas en el presente documento desvelan el estado de la técnica en el momento de la presente invención y/o proporcionan la descripción y capacidad de la presente invención. Las publicaciones se refieren a cualquier publicación científica o de patente, o cualquier otra información disponible en cualquier formato de medio, incluyendo formatos grabados, electrónicos o impresos. Se citan las siguientes referencias: Ausubel et al., ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., NY, NY (1987-2001); Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, 2<sup>a</sup> Edición, Cold Spring Harbor, NY (1989); Harlow y Lane, *antibodies, a Laboratory Manual*, Cold Spring Harbor, NY (1989); Colligan, et al., eds., *Current Protocols in Immunology*, John Wiley & Sons, Inc., NY (1994-2001); Colligan et al., *Current Protocols in Protein Science*, John Wiley & Sons, NY, NY, (1997-2001).

15 20 20 **Vectores de Expresión Recombinante y Células Huéspedes**

25 La invención proporciona vectores, preferentemente, vectores de expresión, que contienen un ácido nucleico que codifica un polipéptido específicos, por ejemplo, el anticuerpo anti-IL-12, o puede usarse para obtener plásmidos que contienen varios genes de anticuerpo HC o LC o partes de los mismos. Como se usa en el presente documento, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico con el que se ha unido. Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN circular de doble hebra al que pueden ligarse segmentos adicionales de ADN. Otro tipo de vector es un vector viral, donde segmentos adicionales de ADN pueden ligarse al genoma viral.

30 30 Ciertos vectores son capaces de réplica autónoma en una célula huésped en la que se introducen (por ejemplo, vectores bacterianos que tiene un origen bacteriano de réplica y vectores episómico de mamífero). Otros vectores (por ejemplo, vectores no episómico de mamífero) están integrados en el genoma de una célula huésped tras la introducción en la célula huésped, y de este modo se replican junto con el genoma huésped. Además, ciertos vectores, por ejemplo, vectores de expresión, son capaces de dirigir la expresión de genes a los que operativamente se unen. En general, los vectores de expresión de utilidad en técnicas de ADN recombinante tienen a menudo forma de plásmidos (vectores). Sin embargo, la invención pretende incluir otras formas de vectores de expresión, tales como vectores virales (por ejemplo, retrovirus defectuosos de réplica, adenovirus y virus adeno-asociados), que sirven funciones equivalentes.

40 40 Los vectores de expresión recombinantes de la invención comprenden un ácido nucleico en una forma adecuada para la expresión del ácido nucleico en una célula huésped. Esto significa que los vectores de expresión recombinantes incluyen una o más secuencias reguladoras, seleccionadas en base a las células huéspedes que se usarán para la expresión, que están operativamente unidas con la secuencia de ácido nucleico que se expresará. En un vector de expresión recombinante, "operativamente unido" pretende significar que la secuencia de nucleótido de interés está unida a la secuencia o secuencias reguladoras de manera que permite la expresión de la secuencia de nucleótido (por ejemplo, en un sistema de transcripción/traslación *in vitro* o en una célula huésped cuando el vector se introduce en la célula huésped). Además, la secuencia reguladora se optimiza en base a las características de la célula huésped, es decir, factores de transcripción.

50 50 50 El término "secuencia reguladora" pretende incluir promotores, potenciadores y otros elementos de control de expresión (por ejemplo, señales de poliadenilación). Tales secuencias reguladoras se describen, por ejemplo, en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990). Las secuencias reguladoras incluyen aquellas que dirigen expresión constitutiva de una secuencia de nucleótido en muchos tipos de células huéspedes y aquellas que dirigen expresión de la secuencia de nucleótido solamente en ciertas células huéspedes (por ejemplo, secuencias reguladoras específicas de tejido). Aquellos expertos en la técnica apreciarán que el diseño del vector de expresión puede depender de factores tales como la elección de la célula huésped que se transformará, el nivel de expresión de proteína deseada, y similares. Los vectores de expresión de la invención pueden introducirse en las células huéspedes para producir de ese modo proteínas o péptidos, incluyendo proteínas o péptidos de fusión o quiméricos, codificados por ácidos nucleicos como se describe en el presente documento.

60 60 60 Los vectores de expresión recombinantes de la invención pueden estar diseñados para la expresión de un polipéptido en células procarióticas (por ejemplo, *E. coli*) o eucarióticas (por ejemplo, células de insectos (usando vectores de expresión de baculovirus), células de levadura o células de mamíferos). Las células huéspedes adecuadas se analizan con más detalle en Goeddel, supra. Alternativamente, el vector de expresión puede

transcribirse y trasladarse in vitro, por ejemplo usando secuencias reguladoras de promotor T7 y polimerasa T7.

La expresión de proteínas en procariotas se realiza a menudo en *E. coli* con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de proteínas de fusión o de no fusión. Los vectores de fusión añaden un número de aminoácidos a una proteína codificada en los mismos, normalmente a la terminal amino de la proteína recombinante. Tales vectores de fusión típicamente sirven tres fines: (1) aumentar la expresión de proteína recombinante; (2) aumentar la solubilidad de la proteína recombinante; y (3) ayudar en la purificación de proteína recombinante actuando como un ligando en la purificación por afinidad. A menudo, en vectores de expresión de fusión, se introduce un sitio de segmentación proteolítica en la unión de la fracción de fusión y la proteína recombinante para permitir la separación de la proteína recombinante de la fracción de fusión después de la purificación de la proteína de fusión. Tales enzimas, y sus secuencias semejantes de reconocimiento, incluyen Factor Xa, trombina y enterokinasa. Los vectores de expresión de fusión típicos incluyen pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) Gene 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) y pRIT5 (Pharmacia, Piscataway, N.J) que fusionan glutationa S-transferasa (GST), proteína de enlace maltosa E, o proteína A, respectivamente, con la proteína recombinante diana.

Para ayudar en la purificación por afinidad, varios polipéptidos con etiqueta y sus respectivos anticuerpos son bien conocidos en la técnica. Ejemplos incluyen etiquetas poli-histidina (poli-his) o poli-histidina-glicina (poli-his-gli); el polipéptido con etiqueta de gripe HA y su anticuerpo 12CA5 (Field et al., Mol. Cell. Biol, 8:2159-2165 (1988)); la etiqueta c-myc y los anticuerpos 8F9, 3C7, 6E10, G4, 7 y 9E10 en los mismos (Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)); y la etiqueta de glicoproteína D de virus Herpes Simplex (gD) y su anticuerpo (Paborsky et al., Protein Enzymeering 3(6):547-553 (1990)). Otros polipéptidos con etiqueta incluyen el péptido Flag (Hopp et al., Bio Technology, 6:1204-1210 (1988)); el péptido de epítope KT3 (Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)); y la etiqueta de péptido de proteína 10 de gen T7 (Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)). Una etiqueta preferente es la etiqueta FLAG.

Ejemplos de vectores de expresión inducibles adecuados de *E.coli* de no fusión incluyen pTrc (Amann et al., (1988) Gene 69:301-315) y pET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990) 60-89). La expresión de gen diana del vector pTrc depende de la transcripción de polimerasa de ARN huésped de un promotor de fusión del híbrido trp-lac. La expresión del gen diana del vector pET 11d depende de la transcripción de un promotor de fusión T7 gn10-lac mediado por la polimerasa de ARN viral co-expresada (T7 gn1). Esta polimerasa viral es suministrada por la cepa huésped BL21 (DE3) o HMS174(DE3) de un prófago residente λ que alberga un gen T7 gn1 bajo el control transcripcional del promotor lacUV 5.

Una estrategia para maximizar la expresión de proteína recombinante en *E. coli* es expresar la proteína en la bacteria huésped con una capacidad disminuida para partir proteolíticamente la proteína recombinante (Gottesman, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990) 119-128). Otra estrategia es alterar la secuencia de ácido nucleico para insertarla en un vector de expresión de manera que los codones individuales para cada aminoácido sean aquellos preferentemente utilizados en *E. coli* (Wada et al. (1992) Nucleic Acids Res. 20:2111-2118). Tal alteración de secuencias de ácido nucleico de la invención puede realizarse mediante técnicas estándares de síntesis de ADN.

En otra realización, el vector de expresión es un vector de expresión de levadura. Ejemplos de vectores para expresión en levadura *S. cerevisiae* incluyen pYEPSec1 (Baldari et al. (1987) EMBO J. 6:229-234), pMFA (Kurjan y Herskowitz, (1982) Cell 30:933-943), pJRY88 (Schultz et al. (1987) Gene 54:114-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), y pPicZ (Invitrogen Corp, San Diego, Calif.).

Alternativamente, el vector de expresión es un vector de expresión de baculovirus. Los vectores de baculovirus disponibles para la expresión de proteínas en células de insectos cultivadas (por ejemplo, células Sf 9) incluyen la serie pAc (Smith et al., (1983) Mol. Cell Biol. 3:2156-2165) y la serie pVL (Lucklow y Summers (1989) Virology 170:31-39).

En otra realización más, un ácido nucleico de la invención se expresa en células mamíferas usando un vector de expresión de mamífero. Ejemplos de vectores de expresión de mamíferos incluyen pCDM8 (Seed (1987) Nature 329:840) y pMT2PC (Kaufman et al. (1987) EMBO J. 6:187:195). Cuando se usan en células mamíferas, las funciones de control del vector de expresión a menudo las proporcionan los elementos reguladores virales. Por ejemplo, los promotores comúnmente usados se derivan de poliomavirus, Adenovirus 2, citomegalovirus y Virus del Simio 40. Para otros sistemas de expresión adecuados para célula procarióticas y eucariótica, véase capítulos 16 y 17 de Sambrook et al., supra.

En otra realización, el vector de expresión de mamífero recombinante es capaz de dirigir la expresión del ácido nucleico, preferentemente en un tipo de célula particular, tales como células de linfoma (por ejemplo, células de mieloma de ratón). En tipos específicos de células, los elementos reguladores específicos de tejido se usan para expresar el ácido nucleico. Los elementos reguladores específicos de tejido son bien conocidos en la técnica. Ejemplos no limitativos de promotores adecuados específicos de tejido incluyen el promotor de albúmina (específico

5 del hígado; Pinkert et al. (1987) Genes DEv. 1:268-277), promotores específicos de linfoide (Calame y Eaton (1988) Adv. Immunol. 43:235-275), en particular promotores de receptores de célula T (Winoto y Baltimore (1989) EMBO J. 8:729-733), e inmunoglobulinas (Banerji et al. (1983) Cell 33:729-740; Queen y Baltimore (1983) Cell 33:741-748), promotores neuro-específicos (por ejemplo, promotor de neurofilamento, Byrne y Ruddle (1989) Proc. Natl. Acad. Sci. UsA 86:5473-5477), promotores específicos de páncreas (Edlund et al., (1985) Science 230:912-916), y promotores específicos de glándula mamaria (por ejemplo, promotor de suero de leche; Patente de Estados Unidos Nº 4.873.316 y Publicación de Solicitud Europea Nº 264.166). Los promotores regulados con desarrollo también se incluyen, por ejemplo, en los promotores hox murinos (Kessel y Gruss (1990) Science 249:374-379) y el promotor  $\alpha$ -fetoproteína (Campes y Tilghman (1989) Genes Dev. 3:537-546).

10 15 20 25 30 La invención proporciona además un vector de expresión recombinante que comprende una molécula de ADN clonada en el vector de expresión en una orientación antisentido. Es decir, la molécula de ADN está operativamente unida a una secuencia reguladora de una manera que permite la expresión (mediante transcripción de la molécula de ADN) de una molécula de AR que es antisentido para el polipéptido que codifica mARN. Las secuencias reguladoras operativamente unidas a un ácido nucleico clonado en la orientación antisentido pueden elegirse para que dirijan la expresión continua de la molécula de ARN antisentido en una variedad de tipos de célula. Por ejemplo, los promotores y/o potenciadores virales, o secuencias reguladoras pueden elegirse con expresión constitutiva directa, específica de tejido o específica de tipo célula de ARN antisentido. El vector de expresión antisentido pueden tener forma de un plásmido recombinante, fagérido, o virus atenuado en el que los ácidos nucleicos antisentido se producen bajo el control de una región reguladora muy eficiente, cuya actividad puede determinarse por el tipo de célula en el que el vector se introduce. Para un análisis de la regulación de expresión genética usando genes antisentido, véase Weintraub et al. (Reviews—Trends in Genetics, Vol. 1(1) 1986).

25 30 Otro aspecto de la invención pertenece a células huéspedes en las que un vector de expresión recombinante de la invención se ha introducido. Los términos "célula huésped" y "célula huésped recombinante" se usan intercambiablemente en el presente documento. Se entiende que tales términos no solamente se refieren a la célula sujeta particular, sino a la progenie o progenie potencial de tal célula. Debido a que pueden ocurrir ciertas modificaciones en las siguientes generaciones debido a mutación o influencias ambientales, tal progenie puede no ser, de hecho, idéntica a la célula original, pero sigue estando incluida en el alcance del término como se usa en el presente documento.

35 40 45 Una célula huésped puede ser cualquier célula procariótica (por ejemplo, *E. coli*) o eucariótica (por ejemplo, células de insecto, células de levadura o mamíferos). Un número de líneas de célula huésped de mamífero adecuadas capaces de expresar polipéptidos glicosilados intactos se han desarrollado en la técnica, e incluyen las líneas celulares COS-1 (por ejemplo, ATCC CRL 1650), COS-7 (por ejemplo, ATCC CRL-1651), HEK293, BHK21 (por ejemplo, ATCC CR-10), CHO (por ejemplo, ATCC CRL 1610) y BSC-1 (por ejemplo, ATCC CRL-26), células Cos-7, células COS, células hepG2, P3X63Ag8.653, SP2/0-Ag14, células 293, células HeLa y similares, que están fácilmente disponibles, por ejemplo, en la Colección Americana de Cultivos Tipo, Manassas, Va ([www.atcc.org](http://www.atcc.org)). En el presente documento, "célula huésped" no incluye células madre del embrión humano.

50 55 60 Los vectores de expresión para estas células pueden incluir uno o más de las siguientes secuencias de control de expresión, un promotor, un potenciador y/o procesar sitios de información, tales como sitios de enlace de ribosomas, sitios de segmentación de ARN, sitios de poliadenilación y secuencias terminadoras transcripcionales (véase, por ejemplo, Ausubel et al., supra; Sambrook, et al., supra).

55 60 65 El vector de ADN puede introducirse en células procarióticas y eucarióticas mediante técnicas convencionales de transformación o transfección. Como se usa en el presente documento, los términos "transformación" y "transfección" pretenden referirse a una variedad de técnicas reconocidas en la técnica para introducir ácido nucleico externo en una célula huésped, incluyendo co-precipitación de fosfato cálcico o cloruro cálcico, transfección mediada por DEAE-dextrano, lipofección o electroporación. Los métodos adecuados para transformar o transfectar células huéspedes pueden encontrarse en Sambrook, et al. (supra) y otros manuales de laboratorio.

55 60 65 Para transfección estable de células mamíferas, es bien conocido que, dependiendo del vector de expresión y la técnica de transfección usada, solamente una pequeña fracción de células puede integrar el ADN externo en su genoma. Con el fin de identificar y seleccionar estos integrantes, un gen que codifica un marcador seleccionable (por ejemplo, para resistencia a antibióticos) se introduce generalmente en las células huéspedes junto con el gen de interés. Los marcadores seleccionables preferentes incluyen aquellos que confieren resistencia a los fármacos, tales como, cloranfenicol, tetraciclinas, gentamicina, kanamicina, ampicilina, G418, higromicina, metotrexato, etc. Las células establemente transfectadas con el ácido nucleico introducido pueden identificarse mediante selección de fármacos (por ejemplo, células que han incorporado el gen marcador seleccionable sobrevirán, mientras que otras células morirán).

65 Una célula huésped de la invención, tal como una célula huésped procariótica o eucariótica, puede usarse para producir un polipéptido (por ejemplo, una proteína o anticuerpo). Por consiguiente, la divulgación proporciona además métodos para producir un polipéptido usando las células huéspedes de la invención. En una realización, el

método comprende cultivar la célula huésped de la invención (en la que se ha introducido un vector de expresión recombinante que codifica un polipéptido) en un medio adecuado de tal manera que se produzca el polipéptido. En otra realización, el método comprende además aislar el polipéptido del medio o la célula huésped.

5 Las células huéspedes de la invención también pueden usarse para producir animales transgénicos no humanos. Por ejemplo, en una realización, una célula huésped de la invención es un oocito fertilizado de una célula madre embrionaria en la que al menos se ha introducido una secuencia que codifica un polipéptido. Tales células huéspedes pueden después usarse para crear animales transgénicos no humanos en los que se han introducido secuencias exógenas que codifican un polipéptido en su genoma o animales recombinantes heterólogos en los que 10 se han alterado secuencias endógenas que codifican un polipéptido. Tales animales son útiles para estudiar la función y/o actividad del polipéptido y para identificar y/o evaluar moduladores de la actividad del polipéptido. Como se usa en el presente documento, un "animal transgénico" es un animal no humano, preferentemente, un mamífero, más preferentemente un roedor, tal como una rata o ratón, en el que una o más de las células del animal incluye un transgen. Otros ejemplos de animales transgénicos incluyen primates no humanos, ovejas, perros, vacas, cabras, 15 pollos, anfibios, etc. Un transgen es ADN exógeno que se integra en el genoma de una célula de la que se desarrolla un animal transgénico y que permanece en el genoma del animal maduro, dirigiendo de ese modo la expresión de un producto de gen codificado a uno o más tipos de célula o tejidos del animal transgénico. Como se usa en el presente documento, un "animal recombinante homólogo" es un animal no humano, preferentemente un mamífero, más preferentemente un ratón, en el que el gen endógeno se ha alterado mediante recombinación homóloga entre el 20 gen endógeno y una molécula de ADN exógeno introducida en una célula del animal, por ejemplo, una célula embrionaria del animal, antes del desarrollo del animal.

25 Un animal transgénico de la invención puede crearse introduciendo ácido nucleico que codifica un polipéptido en los pro-núcleos machos de un oocito fertilizado, por ejemplo, mediante microinyección, infección retroviral, y permitiendo que el oocito se desarrolle en un animal de acogida hembra pseudoembarazado. Las secuencias intrónicas y las señales de poliadenilación también pueden incluirse en el transgen para aumentar la eficiencia de expresión de transgen. Una secuencia o secuencias reguladoras específicas de tejido pueden estar operativamente unida al transgen para dirigir la expresión del polipéptido a células particulares. Métodos para 30 generar animales transgénicos mediante manipulación de embrión y microinyección, particularmente, animales tales como ratones, se han convertido en convencionales en la técnica y se describen, por ejemplo, en Patentes de Estados Unidos Números 4.736.866, 4.870.009 y 4.873.191 y Hogan, Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1986). Se usan métodos similares para la producción de otros animales transgénicos no humanos. Un animal fundador transgénico puede identificarse en base a la presencia del transgen en su genoma y/o expresión de mARN que codifica el transgen en tejidos o células de los animales. Un animal fundador transgénico puede después usarse para engendrar animales adicionales que lleven el transgen. 35 Además, los animales transgénicos no humanos que llevan el transgen pueden además engendrar otros animales transgénico no humano que lleven otros transgenes.

40 Para crear un animal recombinante homólogo, se prepara un vector que contiene al menos una parte de un gen que codifica un polipéptido en el que se ha introducido una eliminación, adición o sustitución para alterar de ese modo, por ejemplo, interrumpir funcionalmente, el gen. En una realización preferente, el vector está diseñado de tal manera que, después de la recombinación homóloga, el gen endógeno se interrumpe funcionalmente (es decir, deja de codificar una proteína funcional; también referido como un vector "noquear"). Alternativamente, el vector puede estar diseñado de tal manera que, después de la recombinación homóloga, el gen endógeno se mute o de otra 45 manera se altere pero que aún codifique la proteína funcional (por ejemplo, la región reguladora corriente arriba puede alterarse para alterar de ese modo la expresión de la proteína endógena). En el vector de recombinación homóloga, la parte alterada del gen está flanqueada en sus extremos 5' y 3' por ácido nucleico adicional del gen para permitir que ocurra la recombinación homóloga entre el gen exógeno transportado por el vector y un gen endógeno en una célula madre embrionaria. Las secuencias adicionales de ácido nucleico flanqueantes tienen la longitud suficiente para la recombinación homóloga buena con el gen endógeno. Típicamente, se incluyen varias kilobases de ADN flanqueante (ambas en los extremos 5' y 3') en el vector (véase, por ejemplo, Thoma y Capecchi 50 (1987) 51:503 para una descripción de vectores de recombinación homólogos). El vector se introduce en una línea de célula madre embrionaria (por ejemplo, mediante electroporación) y se seleccionan las células en las que el gen introducido se ha recombinado homólogamente con el gen endógeno (véase, por ejemplo, Li et al. (1992) Cell 69:915). Las células seleccionadas se injectan después en un blastocisto de un animal no humano (por ejemplo, un ratón) para formar quimeras de agregación (IRL, Oxford, 1987), págs. 113-152). Un embrión químérico no humano puede después implantarse en un animal de acogida hembra pseudoembarazado y el embrión se lleva a término. La progenie que alberga el ADN homólogamente recombinado en sus células germinales puede usarse para engendrar animales en los que todas las células del animal contengan el ADN homólogamente recombinado por la transmisión 55 de la línea celular del transgen. Los métodos para construir vectores de recombinación homólogos y animales recombinantes homólogos se describen con más detalle en Bradley (1991) Current Opinion in Bio/Technology 2:823-829 y en Publicaciones PCT Números WO 90/11354, WO 91/01140, WO 92/0968 y WO 93/04169.

60 En otra realización, pueden producirse animales transgénicos no humanos que contienen sistemas seleccionados que permite la expresión regulada del transgen. Un ejemplo de tal sistema es el sistema recombinasa cre/loxP de bacteriófago P1. Para una descripción del sistema recombinasa cre/loxP, véase, por ejemplo, Lakso et al.

la. (1992) Proc. Natl. Acad. Sci. USA 89:6232-6236. Otro ejemplo de un sistema de recombinasa es el sistema FLP recombinasa de *Saccharomyces cerevisiae* (O'Gorman et al. (1991) Science 251:1351-1355). Si un sistema recombinasa cre/loxP se usa para regular la expresión del transgen, se requieren animales que contengan transgenes que codifiquen Cre recombinasa y una proteína seleccionada. Tales animales pueden proporcionarse a través de la construcción de animales transgénicos "dobles", por ejemplo, emparejando dos animales transgénicos, uno que contenga un transgen que codifica una proteína seleccionada y el otro que contenga un transgen que codifique una recombinasa.

5 Los clones de los animales transgénicos no humanos descritos en el presente documento también pueden 10 producirse de acuerdo con los métodos descritos en Wilmut et al. (1997) Natura 385:810-813 y Publicaciones PCT Números WO 97/07668 y WO 97/07669.

#### Anticuerpos

15 La presente invención incluye además, aunque no se limita a, métodos de uso de ácidos nucleicos y polipéptidos codificados para hacer de este modo anticuerpos y anticuerpos anti-idiotípico, que incluyen 20 composiciones de diagnóstico y terapéuticas, métodos y dispositivos. Tales anticuerpos afectan además opcionalmente a un ligando específico, tal como pero sin limitar a, donde tal anticuerpo modula, disminuye, aumenta, 25 antagoniza, agoniza, mitiga, alivia, bloquea, inhibe, abroga y/o interfiere con al menos una actividad o enlace de proteína, o con actividad o enlace del receptor, in vitro, in situ y/o in vivo. Como un ejemplo no limitativo, un anticuerpo adecuado, parte especificada o variante puede enlazarse al menos con una proteína, o partes 30 especificadas, variantes o dominios de la misma. Un anticuerpo adecuado, parte especificada o variante puede también afectar opcionalmente al menos a una de la actividad o función de la proteína, tal como pero sin limitar a, síntesis de ARN, ADN o polipéptido, liberación de proteína, señalización del receptor, segmentación de membrana, actividad de proteína, producción y/o síntesis de proteína. Los anticuerpos útiles en los métodos y composiciones de 35 la presente invención pueden estar opcionalmente caracterizados por un enlace de alta afinidad con sus antígenos y, opcionalmente y preferentemente, por tener baja toxicidad.

30 Como se usa en el presente documento, un "anticuerpo", y similares incluyen cualquier polipéptido o péptido que contiene una molécula que comprende al menos una parte de una molécula de inmunoglobulina, tal como pero sin limitar a, al menos una región determinante de la complementariedad (RDC) de una cadena pesada o ligera o una parte de enlace de ligando de la misma, una región variable de cadena pesada o cadena ligera, una 35 región constante de cadena pesada o cadena ligera, una región marco, o cualquier parte, fragmento o variante de la misma, o al menos una parte de un receptor o polipéptido de enlace, que puede incorporarse en un anticuerpo.

35 Los anticuerpos pueden incluir uno o más de al menos un RDC, al menos una región variable, al menos una región constante, al menos una cadena pesada (por ejemplo, g1, g2, g3, g4, m, a1, a2, d, e), al menos una cadena ligera (por ejemplo, kappa y lambda), o cualquier parte o fragmento de la misma, y pueden además 40 comprender enlaces de disulfuro inter-cadena o intra-cadena, regiones bisagra, sitios de glicosilación que pueden separarse por una región bisagra, así como cadenas pesadas y cadenas ligeras. Los cadenas ligeras tienen típicamente un peso molecular de aproximadamente 25Kd y las cadenas pesadas oscilan entre 50K y 77 Kd. Las 45 cadenas ligeras existen en dos formas distintas de isotipos, kappa (k) y lambda (l), que pueden combinarse con cualquiera de los tipos de cadena pesada. Todas las cadenas ligeras tienen al menos una región variable y al menos una región constante. Se considera que el anticuerpo IgG es una estructura típica de anticuerpo y tiene dos enlaces de disulfuro intra-cadena en la cadena ligera (uno en la región variable y uno en la región constante), con cuatro en la cadena pesada, y tal enlace abarca un bucle de péptido de aproximadamente 60-70 aminoácidos que comprenden un "dominio" de aproximadamente 110 aminoácidos en la cadena. Los anticuerpos IgG pueden 50 caracterizarse en cuatro clases, IgG1, IgG2, IgG3 e IgG4. Cada clase de inmunoglobulina tiene un conjunto diferente de funciones. La siguiente tabla resume las propiedades fisicoquímicas de cada una de las clases y subclases de inmunoglobulina.

Propiedad	IgG1	IgG2	IgG3	IgG4	IgM	IgA1	IgA2	SigA	IgD	IgE
Cadena pesada	γ1	γ1	γ1	γ1	μ	α1	α2	α1 α2	δ	ε
Conc. Media de suero (mg/ml)	9	3	1	0,5	1,5	3,0	0,5	0,05	0,03	0,00005
Constante de sedimentación	7s	7s	7s	7s	19s	7s	7s	11s	7s	8s
Peso Mol. (X 10 <sup>3</sup> )	146	146	170	146	970	160	160	385	184	188
Vida media (días)	21	20	7	21	10	6	6	¿	3	2
% distribución intravascular	45	45	45	45	80	42	42	Rastro	75	50
Carbohidrato (%)	2-3	2-3	2-3	2-3	12	7-11	7-11	7-11	9-14	12

La siguiente tabla resume ejemplos no limitativos de funciones efectoras de anticuerpo para clases y subclases de anticuerpo humano.

5	Función efectora	IgG1	IgG2	IgG3	IgG4	IgM	IgA	IgD	IgE
	Fijación complemento	+	+/-	++	-	++	-	-	-
	Transferencia placentaria	+	+/-	+	+	-	-	-	-
10	Enlace con Staph A	+++	+++	-	+++	-	-	-	-
	Enlace con Strep G	+++	+++	+++	+++	-	-	-	-

+++ = muy alto; ++ = alto; + = moderado; +/- = mínimo; - = ninguno; ? = cuestionable

15 Como se describe más abajo, existen varios métodos para producir anticuerpos. Una vez que un anticuerpo se produce mediante cualquiera de estos métodos, su aminoácido y correspondientes secuencias genéticas pueden identificarse y, opcionalmente, modificarse, (por ejemplo, optimizarse, humanizarse, etc.) de tal manera que el anticuerpo pueda después producirse recombinantemente.

20 Por ejemplo, un polipéptido especificado, o un fragmento del mismo, puede usarse como un inmunógeno para generar anticuerpos usando técnicas estándares para la preparación de anticuerpos policlonales y monoclonales. Puede usarse el polipéptido de longitud completa o proteína o alternativamente, la divulgación proporciona fragmentos peptídicos antigenicos para su uso como inmunógenos. El péptido antigenico comprende al menos 8 (preferentemente 10, 15, 20 o 30 o más) residuos de aminoácido de una secuencia de proteína y abarca un epítopo de la proteína de tal manera que un anticuerpo se levanta contra el péptido forme un complejo inmune específico con la proteína.

25 30 Un inmunógeno típicamente se usa para preparar anticuerpos inmunizando un sujeto adecuado (es decir, inmunocompetente), tal como un conejo, cabra, ratón u otro mamífero o vertebrado. Una preparación inmunogénica apropiada puede contener, por ejemplo, polipéptido recombinantemente expresado o químicamente sintetizado. La preparación puede además incluir un adyuvante, tal como adyuvante completo o incompleto de Freund, o un agente inmunoestimulante similar.

35 40 Las células que producen anticuerpos pueden obtenerse a partir de sangre periférica o, preferentemente, los nodos del bazo o linfáticos de humanos u otros animales adecuados que se han inmunizado con el inmunógeno de interés. Cualquier otra célula huésped adecuada también puede usarse para expresar ácido nucleico heterólogico o endógeno que codifica un anticuerpo, fragmentos especificados o variantes del mismo. Las células fusionadas (hibridomas) o células recombinantes pueden aislarse usando condiciones selectivas de cultivo u otros métodos conocidos, y clonarse limitando la dilución y la clasificación celular, u otros métodos conocidos. Las células que producen anticuerpos con la especificidad deseada pueden seleccionarse mediante un ensayo adecuado (por ejemplo, ELISA).

45 55 En una técnica, un hibridoma se produce fusionando una línea celular inmortal (por ejemplo, una línea celular de mieloma, tal como, aunque sin limitar a, Sp2/0, Sp2/0-AG14, NS0, NS1, NS2, AE-1, L.5, <243, P3X63Ag8.653, Sp2 SA3, Sp2 MA1, Sp2 SS1, Sp2 SA5, U937, MLA 144, ACT IV, MOLT4, DA-1, JURKAT, WEHI, K-562, COS, RAJI, NIH 3T3, HL-60, MLA 144, NAMALWA, NEURO 2A, o similares), o heteromieloma, productos de fusión de los mismos, o célula o célula de fusión derivada de los mismos, o cualquier línea celular adecuada como las conocidas en la técnica (véase, por ejemplo, [www.atcc.org](http://www.atcc.org), [www.lifetech.com](http://www.lifetech.com), y similares), con células que producen anticuerpo, tales como, aunque sin limitarse a, bazo aislado o clonado, sangre periférica, linfa, amígdala, y otras células inmunes o que contienen célula B, o cualquier otras células que expresen secuencias constantes o variables o RDC de cadena pesada o ligera, ya sea ácido nucleico endógeno o heterólogico, recombinante o endógeno, viral, bacteriano, algal, procariótico, anfibio, insecto, reptil, pez, mamífero, roedor, equino, ovino, de cabra, oveja, primate, eucariótico, ADN genómico, cADN, rADN, ADN o ARN mitocondrial, ADN o ARN cloroplasto, hnARN, mARN, tARN, de único, doble o triple hebra, hibridizado, y similares y cualquier combinación de los mismos. Véase, por ejemplo, Ausubel, supra, y Colligan, Immunology, supra, capítulo 2.

60 65 Otros métodos adecuados de producir o aislar anticuerpos de la especificidad requerida que pueden usarse incluyen, aunque no se limitan a, métodos que seleccionan anticuerpo recombinante de un péptido o biblioteca de péptidos (por ejemplo, aunque sin limitar a, un bacteriófago, ribosoma, oligonucleótido, ARN, cADN, o similares, biblioteca de exposición; por ejemplo, como la disponible en Cambridge antibody Technologies, Cambridgeshire, Reino Unido; MorphoSys, Martinsried/Planegg, Alemania; Biovation, Aberdeen, Escocia, Reino Unido; BioInvent, Lund, Suecia; Dyax, Corp., Enzon, Affymax/Biosite; Xoma, Berkeley, CA; Ixys. Véase, por ejemplo, EP 368.684, PCT/GB91/01134; PCT/GB92/01755; PCT/GB92/00883; PCT/GB93/00605; US 08/350260(5/12/94); PCT/GB94/01422; PCT/GB94/02662; PCT/GB97/01835; (CAT/MRC); WO90/14443; WO90/14424; WO90/14430; PCT/US94/1234; WO92/18619; WO96/07754; (Scripps); EP 614 989 (MorphoSys); WO95/16027 (BioInvent); WO88/06630; WO90/3809 (Dyax); US 4.704.692 (Enzon); PCT/US91/02989 (Affymax); WO89/06283; EP 371 998;

EP 550400; (Xoma); EP 229 046; PCT/US91/07149 (Ixs); o péptidos o polipéptidos estocásticamente generados-US 5723323, 5763192, 5814476, 5817483, 5824514, 5976862, WO 86/05803, EP 590 689 (Ixs, ahora Applied Molecular Evolution (AME) o la que se basa en la inmunización de animales transgénicos (por ejemplo, ratones SCID, Nguyen et al., *Microbiol. Immunol.* 41:901-907 (1997); Sandhu et al., *Crit. Rev. Biotechnol.* 16:95-118 (1996); 5 Eren et al., *Immunol.* 93:154-161 (1998), así como patentes y solicitudes relacionadas) que son capaces de producir un repertorio de anticuerpos humanos, como los conocidos en la técnica y/o descritos en el presente documento. Tales técnicas incluyen, aunque no se limitan a, exposición de ribosoma (Hanes et al., *Proc. Natl. Acad. Sci. USA*, 10 94:4937-4942 (Mayo 1997); Hanes et al., *Proc. Natl. Acad. Sci. USA*, 95:14130-14135 (Nov. 1998)); tecnologías de producción de anticuerpo de única célula (por ejemplo, método de anticuerpo de linfocitos seleccionado ("SLAM") (Patente de Estados Unidos Nº 5.627.052), Wen et al., *J. Immunol.* 17:887-892 (1987); Babcock et al., *Proc. Natl. Acad. Sci. USA* 15 93:7843-7848 (1996); microgotita de gel y citometría de flujo (Powell et al., *Biotechnol.* 8:333-337 (1990); One Cell Systems, Cambridge, MA; Gray et al., *J. Imm. Meth.* 182:155-163 (1995); Kenny et al., *Bio/Technol.* 13:787-790 (1995)); Selección de célula B (Steenbakkers et al., *Molec. Biol. Reports* 19:125-134 (1994); Jonak et al., *Progress Biotech*, Vol, 5, *Inmunización In Vitro en Tecnología de Hibridoma*, Borrebaeck, ed., Elsevier Science Publishers B. V., Amsterdam, Holanda (1988)).

Los métodos para diseñar o humanizar anticuerpos no humano o humanos también pueden usarse y son bien conocidos en la técnica. Generalmente, un anticuerpo humanizado o diseñado tiene uno o más residuos de aminoácido de una fuente que no es humana, por ejemplo, pero sin limitarse a, ratón, rata, conejo, primate no humano y otro mamífero. Estos residuos de aminoácido humanos a menudo son referidos como residuos de "importación", que típicamente se toman de un dominio variable, constante y otro dominio de "importación" de una secuencia humana conocida. Las secuencias Ig humanas conocidas se desvelan, por ejemplo, en los siguientes sitios web:

20 [www.ncbi.nlm.nih.gov/entrez.queri.fcgi](http://www.ncbi.nlm.nih.gov/entrez.queri.fcgi); [www.atcc.org/phage/hadb.html](http://www.atcc.org/phage/hadb.html); [www.sciquest.com/](http://www.sciquest.com/); [www.abcam.com/](http://www.abcam.com/); [www.antibodyresource.com/onlinecomp.html](http://www.antibodyresource.com/onlinecomp.html); [www.public.iastate.edu/~pedro/research\\_tools.html](http://www.public.iastate.edu/~pedro/research_tools.html); [www.mgen.uni-heidelberg.de/SD/IT/IT.html](http://www.mgen.uni-heidelberg.de/SD/IT/IT.html); [www.whfreeman.com/Immunology/CH05/kuby05.html](http://www.whfreeman.com/Immunology/CH05/kuby05.html); [www.library.thinkquest.org/12429/Immune/Antibody.html](http://www.library.thinkquest.org/12429/Immune/Antibody.html); [www.hhmi.org/grnts/lectures/1996/vlab/](http://www.hhmi.org/grnts/lectures/1996/vlab/); [www.path.cam.ac.uk/~mrc7/mikeimages.html](http://www.path.cam.ac.uk/~mrc7/mikeimages.html); [www.antibodyresource.com/](http://www.antibodyresource.com/); [mcb.harvard.edu/BioLinks/Immonology.html](http://mcb.harvard.edu/BioLinks/Immonology.html); [www.immunologylink.com/](http://www.immunologylink.com/); [pathbox.wustl.edu/~hcenter/index.html](http://pathbox.wustl.edu/~hcenter/index.html); [www.biotech.ufl.edu/~hcl/](http://www.biotech.ufl.edu/~hcl/); [www.pebio.com/pa340913/340913.html](http://www.pebio.com/pa340913/340913.html); [www.nal.usda.gov/awic/pubs/antibody/](http://www.nal.usda.gov/awic/pubs/antibody/); [www.m.ehime-u.ac.jp/~yasuhito/Elisa.html](http://www.m.ehime-u.ac.jp/~yasuhito/Elisa.html); [www.biodesign.com/table.asp](http://www.biodesign.com/table.asp); [www.icnet.uk/axp/facs/davies/links.html](http://www.icnet.uk/axp/facs/davies/links.html); [www.biotech.ufl.edu/~fcc1/protocol.html](http://www.biotech.ufl.edu/~fcc1/protocol.html); [www.isac-net.org/sites\\_geo.html](http://www.isac-net.org/sites_geo.html); [aximt1.imt.uni-marburg.de/~rek/AEPStart.html](http://aximt1.imt.uni-marburg.de/~rek/AEPStart.html); [baserv.uci.kun.nl/~jraats/links.html](http://baserv.uci.kun.nl/~jraats/links.html); [www.recab.uni-hd.de/immuno.bme.nwu.ed/](http://www.recab.uni-hd.de/immuno.bme.nwu.ed/); [www.mrc-cpe.cam.ac.uk/imt-doc/public/INTRO.html](http://www.mrc-cpe.cam.ac.uk/imt-doc/public/INTRO.html); [www.ibt.unam.mx/vir/V\\_mice.html](http://www.ibt.unam.mx/vir/V_mice.html); [imgt.cnusc.fr:8104/](http://imgt.cnusc.fr:8104/); [www.biochem.ucl.ac.uk/~martin/abs/index.html](http://www.biochem.ucl.ac.uk/~martin/abs/index.html); [antibody.batch.ac.uk/](http://antibody.batch.ac.uk/); [abgen.cvm.tamu.edu/lab/wwwabgen.html](http://abgen.cvm.tamu.edu/lab/wwwabgen.html); [www.unizh.ch/~honegger/AHOseminar/Slide01.html](http://www.unizh.ch/~honegger/AHOseminar/Slide01.html); [www.cryst.bbk.ac.uk/~ubcg07s/](http://www.cryst.bbk.ac.uk/~ubcg07s/); [www.nimr.mrc.ac.uk/CC/caewg/caewg.html](http://www.nimr.mrc.ac.uk/CC/caewg/caewg.html); [www.path.cam.ac.uk/~mrc7/humanisation/TAHHP.html](http://www.path.cam.ac.uk/~mrc7/humanisation/TAHHP.html); [www.ibt.unam.mx/vir/structure/stat\\_aim.html](http://www.ibt.unam.mx/vir/structure/stat_aim.html); [www.biosci.missouri.edu/smithgp/index.html](http://www.biosci.missouri.edu/smithgp/index.html); [www.cryst.bioc.cam.ac.uk/~fmolina/Web-pages/Pept/spottech.html](http://www.cryst.bioc.cam.ac.uk/~fmolina/Web-pages/Pept/spottech.html); [www.jerini.de/fr\\_products.html](http://www.jerini.de/fr_products.html); [www.patents.ibm.com/ibm.html](http://www.patents.ibm.com/ibm.html). Kabat et al., *Sequences of Polypeptides of Immunological Interest*, US. Dept. Health (1983).

45 Tales secuencias importadas pueden usarse para reducir la inmunogenicidad o reducir, aumentar o modificar en enlace, afinidad, constante de asociación, constante de disasociación, avidez, especificidad, vida media, o cualquier otra característica adecuada, como las conocidas en la técnica. Generalmente, parte de todas las secuencias RDC no humana o humanas se mantienen mientras que las secuencias no humanas de las regiones variable y constante se sustituyen por humanos u otros aminoácidos. Los anticuerpos pueden opcionalmente humanizarse con la retención de una elevada afinidad para el antígeno y otras propiedades biológicas favorables. Para conseguir este objetivo, los anticuerpos humanizados pueden prepararse opcionalmente mediante un proceso 50 de análisis de las secuencias parentales y varios productos conceptuales humanizados que usan modelos tridimensionales de las secuencias parentales y humanizadas. Los modelos tridimensionales de inmunoglobulina están comúnmente disponibles y son familiares para aquellos expertos en la técnica. Hay disponibles programas de ordenador que ilustran y exponen estructuras conformacionales tridimensionales probables de secuencias de inmunoglobulina seleccionadas candidatas. La inspección de estas exposiciones permite el análisis del papel 55 probable de los residuos en el funcionamiento de la secuencia de inmunoglobulina candidata, es decir, el análisis de residuos que influyen en la habilidad de la inmunoglobulina candidata para enlazarse con su antígeno. De esta manera, los residuos FR pueden seleccionarse y combinarse de las secuencias de consenso e importación para que se consigan las características deseadas del anticuerpo, tales como la mayor afinidad para el antígeno o antígenos diana(s). En general, los residuos RDC están directamente y más sustancialmente implicados en la influencia del 60 enlace de antígeno. La humanización o diseño de anticuerpos de la presente invención puede realizarse usando cualquier método descrito, tales como aunque sin limitar, aquellos descritos en, Winter (Jones, et al. *Nature* 321:522 (1986); Riechmann et al., *Nature* 332:3213 (1988); Verhoeyen et al., *Science* 239:1534 (1988)), Sims et al. *Immunol.* 151:2296 (1993); Chothia y Lesk, *J. Mol. Biol.* 196:901 (1987), Carter et al., *Proc. Natl. Acad. Sci. U.S.A.* 89:4285 (1992); Presta et al., *J. Immunol.* 151:2623 (1993); Patentes de Estados Unidos Números: 5723323; 5976862; 5824514; 5817483; 5814476; 5763192; 5723323; 5766886; 5714352; 6204023; 6180370; 5693762; 5530101; 65 5585089; 5225539 y 4816567; PCT/: US98/16280; US96/18978; US91/09630; US94/01234; GB89/01334;

GB91/01134; GB92/01755; WO90/14443; WO90/14424 y WO90/14430; EP 229246.

5 Los anticuerpos también pueden generarse opcionalmente mediante inmunización de un animal transgénico (por ejemplo, ratón, rata, hámster, primate no humano, y similares) capaz de producir un repertorio de anticuerpos humano, como se describe en el presente documento y/o se conoce en la técnica. Las células que producen un anticuerpo pueden aislarse de tales animales e inmortalizarse usando métodos adecuados, tales como los métodos descritos en el presente documento.

10 Los ratones transgénicos que pueden producir un repertorio de anticuerpos humanos que se enlazan con antígenos humanos pueden producirse mediante métodos conocidos (por ejemplo, aunque sin limitar a, Patentes de Estados Unidos Números: 5.770.428, 5.569.825, 5.545.806, 5.625.126, 5.625.825, 5.633.425, 5.661.016 y 5.789.650 presentado por Lonberg et al; Jakobovits et al. WO 98/50433, Jakobovits et al. WO 98/24893, Lonberg et al. WO 98/24884, Lonberg et al. WO 97/13852, Lonberg et al. WO 94/25585, Kucherlapate et al. WO 96/34096, Kucherlapate et al. EP 0463 151 B1, Kucherlapate et al. EP 0710 719 A1, Surani et al. Patente de Estados Unidos Nº 5.545.807, Bruggemann et al. WO 90/04036, Bruggemann et al. EP 0438 474 B1, Lonberg et al. EP 0814 259 A2, Lonberg et al. GB 2 272 440 A, Lonberg et al. Nature 368:856-859 (1994), Taylor et al., Nature Genetics 15:146-156 (1997), Taylor et al., Nucleic Acids Reserch 20(23):6287-6295 (1992), Tuailon et al., Proc Ntal Acad Sci USA 90(8)3720-3724 (1993), Lonberg et al., Int Rev Immunol 13(1):65-93 (1995) y Fishwald et al., Nat Biotechnol 14(7):845-851 (1996). Generalmente, estos ratones comprenden al menos un transgen que comprende ADN de al menos un locus de inmunoglobulina humana que se reorganiza funcionalmente, o que puede sufrir recolocación funcional. Los lugares de inmunoglobulina endógena en tales ratones pueden interrumpirse o borrarse para eliminar la capacidad del animal para producir anticuerpos codificados por genes endógenos.

25 Los anticuerpos también pueden prepararse en leche administrando al menos un anticuerpo que codifica ácido nucleico a animales transgénicos o mamíferos, tales como cabras, vacas, caballos, ovejas, y similares que produzcan anticuerpos en su leche. Tales animales pueden proporcionarse usando métodos conocidos, Véase, por ejemplo, aunque sin limitar las Patentes de Estados Unidos Números 5.827.690; 5.849.992, 4.873.316, 5.849.992, 5.994.616, 5.565.962, 5.304.489, y similares. Los anticuerpos pueden además prepararse usando al menos un anticuerpo que codifica un ácido nucleico para proporcionar plantas transgénicas y células de plantas cultivadas (por ejemplo, aunque sin limitar a, tabaco y maíz) que producen tales anticuerpos, partes especificadas o variantes en la partes de la planta y en las células cultivadas de las mismas.

35 Los anticuerpos pueden enlazarse con antígenos con un amplio rango de afinidades ( $K_D$ ). En una realización preferente, al menos un anticuerpo monoclonal humano de la presente invención puede enlazarse opcionalmente con su antígeno con alta afinidad. Por ejemplo, un anticuerpo monoclonal humano puede enlazarse con un antígeno humano con un  $K_D$  igual o inferior a aproximadamente  $10^{-7}$  M, tal como pero sin limitar a 0,1-9,9 (o cualquier rango o valor en el mismo)  $\times 10^{-7}, 10^{-8}, 10^{-9}, 10^{-10}, 10^{-11}, 10^{-12}, 10^{-13}$  o cualquier rango o valor en el mismo.

40 La afinidad o avidez de un anticuerpo para un antígeno puede determinarse experimentalmente usando cualquier método adecuado. (Véase, por ejemplo, Berzofsky, et al., "Antibody-Antigen Interactions". En Fundamental Immunology, Paul, W. E., Ed., Raven Press: Nueva York, NY (1984); Kuby, Janis Immunology, W. H. Freeman & Company: Nueva York, NY (1992); y métodos descritos en los mismos). La afinidad medida de una interacción particular anticuerpo-antígeno puede variar si se mide bajo diferentes condiciones (por ejemplo, concentración de sal, pH). De este modo, las mediciones de afinidad y otros parámetros de enlace con antígeno (por ejemplo, KD, Da, Dd) se hacen preferentemente con soluciones estandarizadas de anticuerpo y antígeno, y un tampón estandarizado, tal como el tampón descrito en el presente documento.

50 Un anticuerpo dirigido contra un polipéptido (por ejemplo, anticuerpo monoclonal) puede usarse para aislar el polipéptido mediante técnicas estándares, tales como cromatografía por afinidad o inmunoprecipitación. Además, tal anticuerpo puede usarse para detectar la proteína (por ejemplo, en un lisado celular o sobreanadante celular) con el fin de evaluar la abundancia y patrón de expresión del polipéptido. Los anticuerpos pueden también usarse mediante diagnóstico para controlar los niveles de proteína en tejido como parte de un procedimiento clínico de análisis, por ejemplo, para determinar la eficacia de un régimen de tratamiento dado. La detección puede facilitarse uniendo el anticuerpo a una sustancia detectable. Ejemplos de sustancias detectables incluyen varias enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes y materiales radioactivos. Ejemplos de enzimas adecuadas incluyen peroxidasa de rábano picante, fosfatasa alcalina,  $\beta$ -galactosidasa o acetilcolinesterasa; ejemplos de complejos adecuados de grupo prostético incluyen estreptavidina/biotina y avidina/biotina; ejemplos de materiales fluorescentes adecuados incluyen umbelíferona, fluoresceína isotiocianato, rodamina, díclorotriacilamina fluoresceína, cloruro de dansilo o ficoeritrina; un ejemplo de material luminiscente incluye luminol; ejemplos de materiales bioluminiscentes incluyen luciferasa, luciferina y acuorina, y ejemplos de material radioactivo adecuado incluyen  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{135}\text{S}$  o  $^3\text{H}$ .

#### Características de Vectores

65 Se investigaron los promotores inherentes de líneas celulares transfectadas con una fuerte capacidad de producción de anticuerpo monoclonal. Se descubrió inesperadamente que se obtuvieron buenos niveles de

producción, es decir, aproximadamente 300 mg/L, o 25-30 pg por célula por día sobre una base de productividad específica, incluso con promotores de gen humano que se reconocieron por factores de transcripción de ratón en la célula huésped de mieloma murino.

5 Se observó que las líneas celulares transfectadas C379B y C381B fueron fuertes productores de un anticuerpo monoclonal humano con la citoquina IL-12 (anticuerpo anti-IL-12 o IL-12 mAb) codificado por los genes IgHC y LC completamente humanos. La invención proporciona las secuencias completas de ADN de los plásmidos de expresión HC y LC usados para crear las líneas celulares C379B y C381 B y que contienen el gen de anticuerpo anti-IL-12. Los vectores y plásmidos se han diseñado para permitir una inserción conveniente con una única etapa de varios genes de región variable de anticuerpo. En una realización preferente, el vector se usa en un proceso de una única etapa para sustituir la secuencia de región variable anti-IL-12 por una secuencia de región variable que codifique otro anticuerpo monoclonal de interés, por ejemplo, un anticuerpo para una interleuquina humana, un factor de crecimiento, etc. Además, los vectores y plásmidos permiten la sustitución de regiones constantes de anticuerpo.

10 15 Los mapas de dos plásmidos de expresión de acuerdo con la invención, p1560 y p1558 se representan esquemáticamente en las Figs. 1A y 1B. Sus secuencias se desvelan en las SEQ ID NOS: 11 y 12. El plásmido HC específicamente descrito en el presente documento, p1560, codifica regiones constantes del isótipo IgG1 humano. El plásmido LC específicamente descrito en el presente documento, p1558, codifica una región constante del isótipo kappa humano. La secuencia completa de ADN para cada uno de los dos plásmidos tiene alrededor de 20 kb en longitud y sus características seleccionadas, incluyendo, aunque sin limitación, el posicionamiento de las varias secuencias de polinucleótido de control de expresión, se enumeran en la Tabla 1.

20 25 30 Los genes de anticuerpo en estos dos plásmidos están presentes en la orientación inversa/complementaria con el fin de que los ejes centrales del vector pSV2gpt se presenten en su orientación convencional, como se representa en las Figuras 1A y 1B. Como consecuencia, los números de posición del nucleótido correspondientes al extremo de terminal carboxílico de la secuencia codificadora de anticuerpo (secuencia codificadora cH3) son inferiores y los números de posición del nucleótido correspondientes al extremo de terminal amino de la secuencia codificadora de anticuerpo (sitio de clonación de región V) más altos. Además, todas las secuencias que codifican anticuerpos corresponden a la "hebra menos" como se presente en las secuencias mostradas aquí. Esta orientación es razonable para la descripción de los sitios de enlace del factor de transcripción ya que tales sitios de enlace pueden aplicarse a la "hebra más" o "hebra menos".

35 **Tabla 1. Posiciones de nucleótido de características seleccionadas en p1560 y p1558**

Característica	p1560	p1558
Secuencia flanqueante 5'	13200-14950	12614-14871
Región promotora de gen Ig	11209-13199	10623-12613
Secuencia de señal	11060-11070 y 11154-11199	10422-10432 y 10559-10612
Secuencia codificadora de región V	10702-11059	10100-10421
Potenciador de intrón	8449-9030	6765-7209
Secuencia codificadora de CH1	5002-2595	N.A
Secuencia codificadora bisagra	4566-4610	N.A
Secuencia codificadora de CH2	4118-4447	N.A
Secuencia codificadora de CH3	3701-4020	N.A
Secuencia codificadora de kappa C	N.A	5943-6262
Marcador seleccionable de gpt	16200-16655	16122-16577
Gen de resistencia a Amp	18410-19267	18332-19189

40 45 50 55 Los vectores de la presente invención pueden usarse para expresar las secuencias de región constante para IgG1, IgG2, IgG3, IgG4, IgD, IgE o IgM humano, IgG1, IgG2a, IgG2b, IgG3 de ratón o IgG1, IgG2a, IgG2b, IgG2c de rata. Alternativamente, los vectores de la presente invención pueden usarse para expresar las versiones ΔCH1 de las secuencias de región constante anteriormente enumeradas.

60 65 Cualquier anticuerpo, proteína derivada de inmunoglobulina, proteína de fusión, otra proteína o parte de la misma puede sustituirse por la secuencia codificadora de región V en los vectores/plásmidos de la presente invención, por ejemplo, dominio extracelular de un receptor TNF. Para tal proteína de fusión que incluye el dominio CH1, sería necesario alguna forma de LC para la secreción de la célula.

5 El plásmido de expresión p1560 HC como el mostrado esquemáticamente en el presente documento contiene el promotor HC y un potenciador de intrón; sin embargo, el tipo de promotor y potenciador puede variar. El plásmido de expresión p1560 HC contiene una única secuencia que codifica la región constante. Es una variante alotípica que codifica un residuo Arg en la posición 214 (G1m(1,3)alotípico) en lugar del residuo Lys (G1m(1,17 alotípico)

### Análisis Secuencial del Promotor de Gen HC

10 Para definir los elementos en el promotor del gen HC de anticuerpo anti-IL-12 que podría impactar la transcripción del gen y ser parcialmente responsable de los altos niveles de expresión del anticuerpo anti-IL-12 observados en células transfectadas, se realizó un análisis bioinformático de la secuencia 2000-base corriente arriba del codón de inicio de translación HC. Este análisis identificó motivos de secuencia reconocidos por los factores de transcripción relevantes (TFs). La base de datos TF más actualizada, TRANSFAC 7,2 (Matys et al., 2003), se usó para la comparación. Después, se aplicaron varios algoritmos de búsqueda de matriz y patrón para identificar secuencias relevantes. Algunos de los resultados se validaron posteriormente mediante la bibliografía relevante que sostiene la conclusión de que la sinergia y la combinación de estos TFs pueden impulsar la alta producción de anticuerpos.

20 Se construyeron varios modelos de factor de transcripción de alta calidad y específicos de ratón con diferentes parámetros para una búsqueda de matriz (Goessling et al., 2001). Un modelo específico de factor de transcripción de linfocito de ratón también se construyó para una búsqueda de matriz. Además, se realizaron varias búsquedas de ajustes para las diferentes longitudes posteriores (6 bps y superiores).

25 En base a estos resultados, se identificaron 21 sitios potenciales de enlace para TFs de ratón (Tabla 2). La mayoría de estos TFs fueron específicos de linfocito B. algunos de los TFs pueden activarse durante diferentes fases del desarrollo de célula B. Su activación también es dependiente de la presencia de otros sitios de enlace y la interacción de otros factores.

30 Los códigos de accesión de la base de datos TRANSFAC para estos TFs de ratón son: T00032 para Ap1; T01575 para STATx; T00111 para c-Ets-1; T01852 para HMG\_IY; T00479 para Lyf-1; T00278 para YY1; T00613 para NF-Y; T00989 para CREB; T00402 para IRF; T02057 para IPF1; T00422 para IRF-1; T01432 para cMaf; T00814 para TFE3-S; T00017 para C/EBPBeta; T01114 para C/EBPdelta; T01864 para POU2F2/Oct-2; T00644 para POU2F2a/Oct-1; T00651 para POU5F1/Oct-3; T00702 para PU.1; T00169 para c-Rel; y T01159 para TFIID.

35 Se descubrió que seis TFs estaban localizados 500 bp corriente arriba del sitio de iniciación de la transcripción. Estos son POU2F2/Oct-2, POU2F2a/Oct-1, POU5F1/Oct-3, pU.1, c-Rel y TFIID de ratón. Se identificaron los siguientes TFs auxiliares o TFs en asociación con funciones potenciadoras: Ap1; STATx; c-Ets-1; HMG\_IY; Lyf-1; YY1 y TFIID.

40 **Tabla 2: Motivos secuenciales en promotor p1560 relevantes para transcripción**

TF	Posición en p1560	Sito de enlace ID	Secuencia reconocida por TF (SEQ ID NO)	Puntuación Central	Matriz/Parche	Fuentes/Método de Búsqueda	Anotación
AP-1	12292-	V\$API_Q4	agTGACTgacg (13)	1	0,984	HQ_mus/Inmuno/Parche	Interactivo con c-ETs-1 y Nf-AtP
STAT x	12643-	V\$STAT_01	TTCCCTgaa (14)	1	0,969	HQ_mus	Expresado en célula B, particularmente en centro germinal (Henderson, 1998)
C-Ets-1	12697-	V\$ETS1_B	gcAGGAAgtgaaagt (15)	1	0,964	HQ_mus/Inmuno/Parche	Desarrollo acelerado de célula B de propB a célula de plasma
HMG_IY	12381-	V\$HMG1_Y_06	GGAAAGt (16)	1	0,979	Inmunoespecífico	Factor auxiliar para otros factores de transcripción como NF-kB o ATF-2
LyF-1	12239-	V\$LYFI_01	TtTGGGAgg (17)	1	0,989	Inmunoespecífico	Expresado de probB a célula de plasma
YY1	11788-	V\$YY1_Q6	CaaaATGGC (18)	1	0,997	Inmunoespecífico	Represor ubicuo
NF-Y	12311	V\$NFY_Q6_01	AgttagCCAATGG (19)	1	0,974	HQ_mus	Se enlaza con elementos de caja Y de genes MHC clase II
CREB	12289-	V\$CRE_B_01	TGACGtag (20)	1	0,971	HQ_mus	Se enlaza en linfocitos; media la respuesta cAMP

	TF	Posición en p1560	Sito de enlace ID	Secuencia reconocida por TF (SEQ ID NO)	Puntuación Central	Matriz/Parche	Fuentes/Método de Búsqueda	Anotación
5	IRF/ICSBP	11955-; 12694-	V\$IRF_Q6	aaaaaTGAAA gaact (21); ggaagTGAAA gtaat (22)	1	0,967	HQ_mus	Expresado en bazo y timo. Inducido por IFN-gama
10	IPF1	12058-	V\$IPF1_Q4	gtgcTAATGaa a (23)	1	0,965	HQ_mus	Células beta de páncreas; promotor de insulina factor 1
15	IRF-1/IRF-2	11960+	RATONE_S_COX2_02	TTTCATTTTT (24)	0	100	Parche	Requerido para linfopoyesis B; interactivo con ICSBP
20	C-Maf	12833+; 11862	RATONE_S_IL4_02	TCAGCA (25)	0	100	Parche	Expresado en célula Th2, implicado en activación de gen específico de Th2
25	TFE3-S	13105+	RATONE_S_IGH_10	CATGTG (26)	0	100	Parche	Expresado de proB a células de plasma
30	C/EBP Beta	12868+	RATONE_S_INOS_06	TGATGTAAT (27)	0	100	Parche	Expresado de células GC B a células de plasma; sinergiza con Nf- $\kappa$ B/rel
35	C/EBP Delta	11886-	V\$CEBP_DELTA_Q6	ggtgcaGCAAT g (28)	1	0,965	HQ_mus/inmun o	Expresado ubicuamente; sinergiza con Nf- $\kappa$ B
40	PU.1	11506+; 12280-	RATONE_S_GSHPX1_01	CTTCTC (29)	0	100	Parche	Expresado de progenitor a GC B; sin células B si PU.1 es deficiente
45	POU2 F2/Oct -2	11304+	RATONE_S_IGH_44	ATTTGCAT (30)	0	100	Parche	Expresado de proB a células de plasma; defecto en secreción si es deficiente
50	TFIID	11601+	RATONE_S_MBPO_04	TTCAAA (31)	0	100	Parche	Interactivo con PU.1, REL
55	C-Rel	11580-	V\$CREL_01	TggccTTTCC (32)	1	0,968	HQ_mus	Prevalece en células B maduras, activador potente
60	POU5 F1/Oct -3	11304+	RATONE_S_IGH_45	ATTTGCAT (33)	0	100	Parche	Pueden tener función de represión de potenciador de IgH
65	POU2 F1/Oct -1	11303-; 12598-	V\$OCTI_B	tATGCAaatg (34); CCGAAatgcA ATTC (35)	1	1	HQ_mus/ Parche	El enlace de ADN se reducir por GR de una manera dependiente de ligando
	<ul style="list-style-type: none"> <li>- Posición en vector HC en p1560 corresponde al primer nucleótido; "+" indica más hebra"; "-" indica menos hebra.</li> <li>- Sitio de enlace en la identificación de sitio de enlace TF asignado por la base de datos TRANSFAC.</li> <li>- Puntuación central indica la puntuación de enlace calculada por TRANSFAC para los cinco nucleótidos más conservados, consecutivos usados en una matriz.</li> <li>- Puntuación coincidencia/parche indica la similitud de una subsecuencia a la matriz de TRANSFAC. La puntuación es de 0 a 1, siendo 1 la mejor coincidencia. Si la puntuación es 100, fue una búsqueda Parche.</li> <li>- Fuente/método de búsqueda indica si se usó el modelo de matriz específica de inmunocélula, el modelo de matriz de alta calidad y/o el método de búsqueda de parche.</li> <li>- La subsecuencia Oct-2 se publicó por Henderson et al., 1995.</li> </ul>							

#### Análisis Secuencial del Promotor de Gen LC

La región del promotor de gen LC se analizó mediante los mismos métodos, que produjeron una lista de 20 factores potenciales de transcripción de ratón y sus sitios de enlace (mostrados en la Tabla 3). Los códigos de acceso de la base de datos TRANSFAC para estos TFs son: T02057 para IPF12, T01786p para E12, T00169 para c-Rel, T00549 para NF-AT, T01675 para NKx2-5, T01575 para STATx, T01852 para HMGIY; T00111 para c-Ets-1, T00702 para Pu.1, T00278 para YY1, T00032 para Ap1, T01429 para Sox-5, T00017 para C/EBPBeta, T00814 para TF3-S, T01554 para Mitf, T01159 para TFIID, T00613 para NF-Y, T01114 para C/EBPdelta, T01864 para

POU2F2/Oct-2 y T00644 para POU2F2a/Oct-1. Entre ellos, se descubriró que NF-AT, C/EBPdelta, POU2F2a/Oct-1, Sox-5, E12, Statx, NF-Y, TFE3-S, Mitf, C/EBPBeta, TFIID y POU2F estaban localizados 500 bps corriente arriba del sitio de inicio de transcripción. Los TFs auxiliares de cadena ligar fueron Ap1, STATx, c-ETS-1, HMGIY, YY1 y TFIID.

5 **Tabla 3. Motivos secuenciales en promotor p1558 relevantes para transcripción**

TF	Posición en p1560	Sito de enlace ID	Secuencia reconocida por TF (SEQ ID NO)	Puntuación Central	Matriz/Parche	Fuentes/Método de Búsqueda	Anotación
HMGIY	11827-	V\$HMGIY_Q6	aaTTTCC (36)	1	1,00	InmunoEspec	Factor auxiliar para otro TF tal como NF-kB o ATF-2
NF-AT	12249-	V\$NFAT_06	gctTTTCCtaaa (37)	1	0,983	InmunoEspec	Células T activadas; compuestas de un componente AP-1 y de tipo Rel
PU.1	11676-	V\$PUI_Q6	CTTCCCt (38)	1	0,969	InmunoEspec	Expresado de progenitor a células GC B. Sin células B es deficiente.
C/EBP delta	10913-	V\$CEBP_DELTA_Q6	cATTGCtccatc (39)	1	0,966	HQ_mus/ InmunoEspec	Expresado de manera ubicua; actúa sinergisticamente con NF-kB
POU2F 1a/Oct-1	10696-	V\$OCT1_Q6	ctcaTTTGCatgttc (40)	1	0,994	HQ_mus/ Parche	Enlace de ADN se reduce por GR de una manera dependiente de ligando.
Sox-5	11090	V\$SOX5_Q1	ggaACAAATgt (41)	1	0,992	HQ_mus/ Parche	Células germinales post-meioticas; gen 5 de caja HMG relacionado con SRY
Nkx2-5	11937-	V\$NKX25_01	TtAAGTG (42)	0	0,986	HQ_mus	Expresado en estroma de nodo linfático; activador transcripcional TSL-1
C-Rel	12253-	V\$CREL_01	gggcTTTCC (43)	1	0,979	HQ_mus	Prevalece en células B maduras; activador potente
E12	10781-; 12363-	V\$E12_Q6	gaCAGGTgggg (44); ggCAGGTgggt (45)	1	0,976	HQ_mus/ Parche	Redundancia funcional de E2A y E2-2, excepto en células B maduras
STATx	10857-	V\$STAT_01	TTCCCAataa (46)	1	0,972	HQ_mus	Expresado en células B, particularmente en centros germinales (Henderson, 1998)
IPFI	12530-	V\$IPFI_Q4	tatCATTAGgc (47)	1	0,960	Parche	Células Beta de páncreas; factor 1 de promotor de insulina
NF-Y	10929-	V\$NFY_Q6	catacaCCAAATga (48)	1	0,966	HQ_mus	Se enlaza con elementos de caja Y de genes II de clase MHC
TFE3-S	11133-	RATONE_SIGH_10	CATGTG (49)	0	100	Parche	Expresados de proB a células de plasma
Mitf	11133-	RATONE_STBX2_01	CATGTG (50)	0	100	Parche	Célula de melanocito; forma heterodímeros con proteínas relacionadas con TFE3
Ap-1	111631-; 120251-; 120924-	RATONE_SIL_2_11	AGGTAAT (51)	0	100	Parche	Interactúa con c-Ets-1 o NF-AtP
Nkx2-5	11937-	V\$NKX25_01	TtAAGTG (42)	0	0,986	HQ_mus	Expresado en estroma de nodo linfático; activador transcripcional TSL-1
C-Rel	12253-	V\$CREL_01	gggcTTTCC (43)	1	0,979	HQ_mus	Prevalece en células B maduras; activador potente
C/EBP beta	11065-	V\$CEBP_B_02	aaatgtTGCAAgta (52)	1	0,917	InmunoEspec	Expresado de células B de centro germinal para células de plasma
YY1	11564-	RATONE_S_CMYC_05	ACATGG (53)	0	100	Parche	Represor ubicuo
c-Ets-1	11792-	RATONE_S_TIMP1_02	CAGGAAG (54)	0	100	Parche	Desarrollo acelerado de célula B de proB para células de plasma
TFIID	118911-; 110151+	RATONE_S_MBp_04	TTCAAA (55)	0	100	Parche	Interactivo con PU.1 y REL
POU2F 2/Oct-1	10700-	RATONE_SIGH_44	ATTTGCAT (56)	0	100	Parche	Expresado de proB a célula de plasma. Secreción defectuosa si es deficiente

Resumen de las Propiedades del Vector

55 Los únicos promotores HC y LC aquí descritos han demostrado ser capaces de impulsar altos niveles de expresión de gen, como en el caso de las líneas celulares transfectadas C379B y C381B. La identificación de motivos secuenciales que pueden determinar niveles de transcripción proporciona información para aumentar la transcripción de genes customizando las secuencias del promotor y potenciador y usando células huéspedes que expresan la mejor combinación de factores de transcripción para esos promotores. Esto podría incluir el uso de una célula huésped que se ha modificado para sobreexpresar los factores de transcripción que pueden aumentar la expresión y/o una célula huésped que se ha modificado para expresar de manera mínima los factores de transcripción que pueden impedir expresión, entre otros. Por ejemplo, las células huéspedes pueden co-transfектarse con un gen que codifica el factor de transcripción OBF-1 para obtener niveles más altos de OBF-1, o usan ARN antisentido, interferente (por ejemplo, siARN o shARN), o técnicas con gen noqueado para reducir la expresión de TFs que pueden regular de manera negativa la transcripción de gen Ig, por ejemplo, NF- $\mu$ NR.

La divulgación también comprende métodos para identificar la interacción entre las secuencias de las regiones promotoras y potenciadoras y factores de transcripción en vectores/plásmidos y huéspedes celulares, moderar la interacción alterando las secuencias para efectuar transcripción, traslación y niveles de expresión de genes, y determinar las modificaciones y ajustarlas con el fin de controlar los niveles. La alteración secuencial puede ser por mutaciones a las secuencias en la región o por una completa sustitución de las regiones compatibles con la línea celular usada.

## EJEMPLOS

10 Los siguientes ejemplos específicos pretenden ilustrar la invención y no deberían construirse como limitadoras del alcance de la invención.

### EJEMPLO 1: Clonación de genes de anticuerpo humano anti-humano interlequina-12 y preparación de líneas celulares muy productivas

#### Resumen

La línea celular de hibridoma C340A que secreta el mAb humano anti-humano IL-12 produce menos de 0,4 µg/ml de cultivo celular gastado. Este nivel de producción fue suficiente para generar las grandes cantidades de material necesario para futuros estudios. Para aumentar el nivel de producción de mAb IL-12 y para tener una línea celular bien caracterizada como la fuente, los genes de cadena pesada y ligera de mAb IL-12 se clonaron e introdujeron en célula de mieloma de ratón P3X63Ag8.653 (653) o SP2/0-Ag14 (SP2/0). Los genes de cadena pesada y ligera de mAb IL-12 se clonaron de una biblioteca de bacteriófago preparada a partir de ADN genómico de hibridoma C340A y transfirieron a vectores de plásmido. Los plásmidos de expresión resultantes de cadena pesada y ligera se secuenciaron después y las secuencias se compararon con secuencias de línea celular conocida. Estos plásmidos se introdujeron en células de mieloma de ratón y los clones transfectados se filtraron para expresión Ig humano mediante ELISA para identificar los productores más altos. Los clones parentales seleccionados se subclonaron después para identificar los clones con mayor producción y homogéneos. Los subclones de interés se caracterizaron después mediante análisis de curva de crecimiento y se analizaron para estabilidad de producción de mAb durante conforme avanzaba el tiempo. Un ensayo de enlace de antígeno realizado con sobrenadantes de estas líneas celulares mostró que el mAb recombinante se unió a interleuquina 12 humana inmovilizada (hIL-12) con una afinidad indistinguible del anticuerpo purificado del hibridoma original C340, confirmando que los genes que se clonaron y expresaron fueron los genes que codifican IL-12 mAb. Los subclones con mayor producción en cultivos de matraz T gastado fueron transfectante C379B 653 que produjo 135 µg/ml de IL-12 mAb y transfectante C381B SP2/0 que produjo 150 µg/ml de IL-12 mAb.

#### Materiales

Trisol se obtuvo de Gibco BRL, Grand Island, NY. Proteinasa K fue de Sigma Chemical, St. Lois, MO. Las enzimas de restricción se compraron en New England Biolabs, Beverly, MA. El Kit de Clonación Lambda EMBL3/Gigapack II Gold fue de Stratagene, La Jolla, CA. Las membranas Protran fueron suministradas por Schleicher & Schuell, Keen, NH. ARNasa y el Kit de Etiquetado Random Prime fueron de Boehringer Mannheim, Alemania. ADNasa fue de Pharmacia, Uppsala, Suecia.  $\alpha^{32}$  P-dCTP se compró en DuPont NEN, Boston, MA. IL-12 humano se compró en RDI Immunochemicals, Flanders, NJ. Los oligonucleótidos a la medida se compraron en Biosource International, Camarillo, CA. Los nombres, números de identificación de secuencia, y secuencias de los oligonucleótidos en este trabajo fueron los siguientes:

	Nombre	SEQ ID	Secuencia
5	5'-10	1	CCCAGGTGCAGCTGGTG
10	5'-46	2	CTCAGGTGCAGCTGGTGG
15	5'-63	3	CCCAGGTGCAGCTACAG
20	5'-73	4	CCGAGGTGCAGCTGGTG
25	hujH4	5	AACCTCGAGTTAACGGAGG
30	hujH6	6	GCAGGAAACCCACAGG
35	5'VK1	7	ATCCAGATGACCCAGTCT
40	5'VK2	8	ATCGTGTGACACAGTCTCCA
45	hujK3	9	AATATGCACAAAACTTGCAC
50	hujK5	10	ATTGAGCCTCTAAAGGTC

#### Clonación de los Genes IL-12 mAb de Cadena Pesada y Ligera

Para clonar y aislar los genes IL-12 mAb de cadena pesada y ligera de ADN de hibridoma C3430, se preparó primero una biblioteca genómica. El ADN genómico total se aisló de  $5 \times 10^7$  células mediante digestión con 0,1 mg/ml proteinasa K en un tampón que contenía 100 nM NaCl, 10 mM Tris-Cl (pH 8,0), 25 mM EDTA (pH 8,0) y 0,5% SDS, y una posterior extracción de fenol/cloroformo y precipitación de etanol. Despues se digirieron doscientos µgs de ADN genómico C340 con 8 unidades/ml enzima de restricción Sau3A y las alícuotas se tomaron cada 5 minutos. En base al análisis de tamaños de fragmento mediante electroforesis de gel de agarosa, se acumularon las

alícuotas de 5, 10 y 15 minutos que mostraron la digestión parcial deseada. Estas alícuotas se cargaron en un gradiante de sacarosa de 10% a 40% y giraron a 22.000 RPM durante 42 horas en un ultracentrifugadora Beckman (rotor SW41TI)O para fraccionar los fragmentos por tamaño. Se recogieron fracciones de quinientos  $\mu$ l y después se analizaron sobre un gel de agarosa para visualizar los fragmentos de ADN. El ADN se aisló de las fracciones que contenían fragmentos en el rango de tamaño 15-23 kba. Los fragmentos de ADN genómico purificado se ligaron al vector bacteriófago Lambda EMBL3 y se empaquetaron en partículas bacteriófagas usando el Kit de Clonación Lambda EMBL3/Gigapack III Gold de acuerdo con el protocolo del fabricante. La biblioteca genómica final consistió en aproximadamente 1 millón de clones.

5                   Para filtrar la biblioteca genómica para los genes IL-12 mAb de cadena pesada y ligera, se mezclaron 640.000 bacteriófagos de la biblioteca Sau3A con cepa E.coli Y1090r y se colocaron en dieciséis placas de agar de 150 mm en una densidad de 40,000 clones por placa. El ADN bacteriófago de las 16 placas se transfirió a dos conjuntos de membranas de nitrocelulosa Protran. En resumen, las membranas de nitrocelulosa se colocaron en las placas, se dejaron 2 minutos, se retiraron y posteriormente se trataron con solución de desnaturización, solución de neutralización, y tampón Tris como se describe en el manual de instrucciones del Kit de Clonación Lambda EMBL3/Gigapack II Gold. Se usó un entrecruzador Stragene UV para fijar ADN desnaturizado a las membranas.

10                  Para filtrar la biblioteca genómica para los genes IL-12 mAb de cadena pesada y ligera, se mezclaron 640.000 bacteriófagos de la biblioteca Sau3A con cepa E.coli Y1090r y se colocaron en dieciséis placas de agar de 150 mm en una densidad de 40,000 clones por placa. El ADN bacteriófago de las 16 placas se transfirió a dos conjuntos de membranas de nitrocelulosa Protran. En resumen, las membranas de nitrocelulosa se colocaron en las placas, se dejaron 2 minutos, se retiraron y posteriormente se trataron con solución de desnaturización, solución de neutralización, y tampón Tris como se describe en el manual de instrucciones del Kit de Clonación Lambda EMBL3/Gigapack II Gold. Se usó un entrecruzador Stragene UV para fijar ADN desnaturizado a las membranas.

15                  Las sondas para las cadenas pesada y ligera se prepararon usando fragmentos de ADN genómico que contenía secuencias de región constante de cadena pesada de IgG1 humano (2,8 kb fragmento EcoR1/HindIII de p747) o secuencias de región constante de cadena ligera kappa humana (fragmento 2,4 kb Ncol/Xba de p95). Ambas sondas se etiquetaron con  $\alpha$ -<sup>32</sup> P-dCTP usando un Kit de Etiquetado de ADN Random Primed.

20                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 25                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

25                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 30                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

35                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 40                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

45                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 50                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

55                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 60                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

65                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 70                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

75                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 80                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

85                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 90                  Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

95                  Las membranas de nitrocelulosa se prehíbridizaron durante una hora en un tampón de hibridización (50% formamida, 5X SSC, 3,75 mM Tris pH 7,8, 1X Solución Denhardts, 25  $\mu$ g/ml ADN cortado y 5% sulfato de dextrano). 100                 Las sondas con un CPM de aproximadamente 1 x 10<sup>8</sup> se desnaturizaron a 100 °C y después se añadieron al tampón de hibridización y se dejaron hibridizar a 42 °C durante la noche. Los filtros se lavaron dos veces en 2X SSC, 0,1% SDS durante 10 minutos a TA, después dos veces en 0,2X SSC, 0,1% SDS a 65 °C mientras se agitaba. Los filtros se secaron y se expusieron a película de rayos x.

#### 40                  Preparación de Plásmidos de Expresión IL-12 mAb

45                  Con el fin de caracterizar los insertos de ADN clonado, se purificó ADN de los clones bacteriófagos positivos usando un método de cultivo líquido. Doscientos  $\mu$ l de cultivo bacteriano Y1090r se combinaron con 100  $\mu$ l de bacteriófago en tampón de dilución lambda 1X y la mezcla se incubó a 37 °C durante 15 minutos para permitir que el bacteriófago se uniera a la bacteria. Después, se añadieron 10 ml de medio LB que contenía 10 mM MgSO<sub>4</sub> y los cultivos se incubaron durante la noche a 37 °C con agitación. El siguiente día se añadieron 100  $\mu$ l de cloroformo a cada tubo para lisar por completo la bacteria. Después de clarificación, se añadieron 20  $\mu$ g de ARNasa y 20  $\mu$ g de ADNasa y las muestras se incubaron a 37 °C durante 30 minutos. Para precipitar las partículas fago, después se añadieron 20 ml de una solución que contenía 20% PEG y 2,5 M NaCl y las muestras se incubaron en hielo durante 1,5 horas. Las partículas bacteriófagas se transformaron en bolas y después se lavaron una vez con tampón de dilución de lambda 1X. La bola se volvió a suspender después en una solución de 0,2% SDS y 20 mM EDTA y se incubó en 68EC durante 15 minutos. El ADN se extrajo de las partículas bacteriófagas usando fenol y cloroformo, y después el ADN se precipitó con isopropanol y se lavó con 70% etanol. El ADN purificado se volvió a suspender en 50  $\mu$ l de tampón TE.

55                  Se usó PCR para determinar cuál, si alguna, de los clones contenía regiones variables humanas además de la región constante conocida por que estuvieran presentes en base a la homología para las sondas radioetiquetadas. Debido a que la secuencia de las regiones variables en los extremo 50 era desconocida, los cebadores se acumularon y usaron para preparar las reacciones de secuencias. Para la reacción HC, se usó una mezcla de cuatro cebadores 5'-10, 5'-46, 5'-63 y 5'-73, que representan cada una de las posibles secuencias. En el extremo 3', se usaron los cebadores huJH4 y huJH6 que son homólogos a las regiones JH4 y JH6. Para LC, en el extremo 5', se usó una mezcla de dos cebadores de consenso 5'VK1 y 5'VK2. En el extremo 3', se usaron los cebadores huJK3 o huJK5 homólogos a las regiones JK3 o JK5. Las reacciones estándares PCR 100  $\mu$ l se realizaron usando 0,5 nanogramos de ADN fago como plantilla. Los cebadores se templaron a 48 °C y se realizaron 25 ciclos de PCR. Las reacciones se fraccionaron sobre un gel de agarosa para su visualización.

5 Para transferir los insertos del vector bacteriófago a un plásmido que contiene el gen marcador seleccionable *gpt*, ADN bacteriófago que contenía regiones variables y constantes se digirió con *Sa*I. Los fragmentos resultantes se purificaron y clonaron entre los sitios *Xba*I y *Sal*I de plásmido p1351. Debido a que había un sitio interno *Sal*I en la secuencia codificadora de la región variable de cadena pesada, dos fragmentos *Sal*I que tenían un tamaño de 13 kb y 4 kb tuvieron que transferirse al vector de expresión p1351 secuencialmente. Primero el fragmento *Sal*I de 13 kb se clonó en los sitios *Xba*I y *Sal*I de p1351 y el plásmido resultante designó p1557. Posteriormente, el fragmento de 4 kb se transfirió a p1557 y el plásmido de expresión final resultante se llamó p1560. El inserto *Sal*I de cadena ligera de 15 kb fue capaz de transferirse en una etapa y el plásmido de expresión final se designó p1558.

10 **Transfecciones celulares, Filtrado para Productores Altos y Expresión de los Genes Clonados**

15 El plásmido p1560 de cadena pesada se linealizó mediante digestión con enzima de restricción *Pvu*I y el plásmido p1558 de cadena ligera se linealizó usando enzima de restricción *Sal*I. Las células 653 y SP2/0 se transfectaron por separado con los plásmidos linealizados pre-mezclados mediante electroporación y las células cultivadas y transfectantes se seleccionaron usando ácido micofenólico como el descrito. Los sobrenadantes celulares de las colonias resistentes al ácido micofenólico se sometieron a ensayo aproximadamente dos semanas más tarde para IgG humano. Pare este experimento, los sobrenadantes celulares se incubaron en placas ELISA con 96 pozos que habían sido previamente cubiertas con anticuerpos específicos Fc de IgG de cabra anti-humano. IgG humano unido se detectó usando anticuerpo IgG (H+L) anti-humano de cabra conjugado con alcalina fosfatasa y sustratos de alcalina fosfatasa como se ha descrito. Las células de los clones más productores se transfirieron a platos de cultivo con 24 pozos en medio estándar, IMDM, 5% FBS, 2 mM glutamina, mezcla de selección de ácido micofenólico (5g/L Xantana, 250 mg/L Hipoxantana, 50 mg/L ácido micofenólico, 50 mM NaoH); más tarde los sobrenadantes recogidos de los cultivos gastados se cuantificaron cuidadosamente con ELISA haciendo diluciones dobles en serie de cada muestra y comparando los valores O.D con la curva estándar preparada usando mAb purificado de células C340. Los clones seleccionados se transfirieron después a matrices T75 y la producción de IgG humano se cuantificó de manera similar mediante ELISA. En base a estos valores, los transfectantes con mayor producción se subclonaron sembrando una media de una célula por pozo y realizando ensayos ELISA de sobrenadantes celulares a partir de colonias individuales de subclones.

30 Los subclones seleccionados para los ensayos finales fueron el transfectante C379B de 653 y los transfectantes C381B de SP2/0. Los niveles máximos de producción para C318B fueron 135 µg/ml y 150 µg/ml, respectivamente. Las densidades celulares pico para C379B y C381B fueron  $1 \times 10^6$  células/ml y  $1,45 \times 10^6$  células/ml, respectivamente.

35 **Ensayo para Enlace de IL-12**

40 Los sobrenadantes de las tres líneas parentales (clon 2 y clon 18 de transfectantes 653 y clon 1 de transfectante SP2/0) se usaron para analizar las características de enlace de anti-hull-12 mAb recombinante. Las concentraciones de mAg en las tres muestras de sobrenadante se determinaron primero mediante ELISA. Las cantidades de titulación de las muestras sobrenadantes, o mAb purificado de células C340 como un control positivo, se incubaron después en placas de 96 pozos que previamente se habían cubierto con 2 µg/ml de hull-12. El mAb unido se detectó después con anticuerpo IgG (H+L) anti-humano de cabra conjugado con alcalina fosfatasa y los apropiados sustratos de alcalina fosfatasa.

45 Las características de enlace de antígeno de sobrenadantes de tres líneas celulares parentales separadas se compararon con mAb purificado de células C340. Se ha demostrado que mAB de sobrenadantes de célula transfectada se unió específicamente a hull-12 inmovilizado de una manera indistinguible de mAb purificado de la línea celular original C340. Este resultado confirmó que se clonaron y expresaron los genes correctos.

50 **Secuenciación**

55 Las secuencias de los insertos de plásmido de expresión de cadena ligera y pesada se determinaron usando química terminadora BigDye y un secuenciador de ADN automatizado ABI377. Los cebadores se diseñaron para secuenciar corriente arriba o corriente debajo de secuencias conocidas en la región variable, región constante y eje central del vector. Cada reacción de secuenciación dio aproximadamente 500 bp de secuencia adicional que se alineó usando el programa AssemblyLIGN. Después de cada ronda de secuenciación, los cebadores se diseñaron para continuar las secuencias corriente arriba y corriente debajo de la secuencia conocida. Las secuencias de p1558 y p1560 se muestran en la SEQ ID NO: 11 y 12.

60 El inserto clonado que contiene las secuencias codificadoras de cadena ligera de IL-12 mAb tiene aproximadamente 15 kb de longitud. El codón de inicio está situado 4310 pares base del extremo 5' del inserto. Hay un intrón 3838 bp J-C entre las secuencias que codifican la región variable y constante y aproximadamente 6 kb de secuencia no codificadora corriente debajo de la región constante. La región constante de la cadena ligera es el alotipo Km(3), el mismo que cA2 y 7E3.

5 El inserto clonado que contiene las secuencias codificadoras de cadena pesada de IL-12 mAb tiene aproximadamente 16 kb de longitud. El codón de inicio está situado 4802 bp del extremo 5' del inserto. Hay un intrón 5406 y aproximadamente 4,8 kb de secuencia no codificadora corriente debajo de la región constante. El alotipo IgG es G1m(f), definido por la arginina (arg) en la posición 214. Éste es diferente que cA2 que tiene una lisina (lis) en la posición 214 y por lo tanto del alotipo G1m(z).

10 Aunque se ha ilustrado y descrito anteriormente con referencia a ciertas realizaciones específicas, la presente invención no se limita sin embargo a los detalles mostrados. Más bien, la presente invención está dirigida a un vector de expresión que tiene varias secuencias, células huéspedes y kits desvelados en el presente documento, 15 cuyo alcance se define en las reivindicaciones adjuntas.

Lista de Referencia

- 15 (1) V Matys, e Fricke, R Geffers, E Gößling, M Haubrock, R Hehl, K Hornischer, D Karas, A Kel, O Kel-Margoulis, D-U Klosos, S Land, B Lewicki-Potapov, H Michael, R Münch, I Reuter, S Rotert, H Saxel, M Scheer, S Thiel y E Wingender. TRANSFAC®: regulación transcripcional, de patrones y perfiles Nucleic Acids Res. 31, 374-378 (2003).
- 20 (2) E Goessling, O Kel-Margoulis, A Kel, E Wingender. MATCH™- una herramienta para buscar sitios de enlace del factor de transcripción de secuencias de ADN: Aplicación para el análisis de cromosomas humanos. Conferencia Alemana sobre Bioinformática 2001. <http://www.bioinfo.de/isb/gcb01/poster/index.html>
- 25 (3) A Henderson y K Calame, Regulación transcripcional durante el desarrollo de célula B. Annu. Rev. Immunol. 16:163-200 (1998).
- (4) A Henderson y K Calame, Lecciones en regulación transcripcional aprendidas de estudios sobre genes de inmunoglobulina. Critical Rev en Eucaryotic Gene Expr 5:255-280 (1995).

30 **LISTADO DE SECUENCIA**

<110> CENTOCOR, INC.

<120> ANTI-IL-12 ANTIBODY BASED VECTORS, HOST CELLS, AND METHODS OF PRODUCTION AND USES

35 <130> CEN5089 PCT

<140> TO BE ASSIGNED

40 <141> 2005-12-21

<150> 60/637,936

<151> 2004-12-21

45 <160> 56

<170> FastSEQ for Windows Version 4.0

<210> 1

50 <211> 17

<212> DNA

<213> Artificial Sequence

<220>

55 <223> Primer

<400> 1

cccaggtgca gctggtg 17

60 <210> 2

<211> 18

<212> DNA

<213> Artificial Sequence

65 <220>

<223> Primer

<400> 2  
 ctcaggtgca gctgggtgg 18

5 <210> 3  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence

10 <220>  
 <223> Primer

<400> 3  
 cccaggtgca gctacag 17

15 <210> 4  
 <211> 17  
 <212> DNA  
 <213> Artificial Sequence

20 <220>  
 <223> Primer

<400> 4  
 ccgaggtgca gctggtg 17

25 <210> 5  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

30 <220>  
 <223> Primer

<400> 5  
 35 aacctcgagt taacggagg 19

<210> 6  
 <211> 17  
 <212> DNA

40 <213> Artificial Sequence

<220>  
 <223> Primer

45 <400> 6  
 gcagggaaacc ccacagg 17

<210> 7  
 <211> 18

50 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer

55 <400> 7  
 atccagatga cccagtc 18

<210> 8  
 60 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

65 <220>  
 <223> Primer

<400> 8  
 atcgtgtga cacagtctcc a 21

5 <210> 9  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

10 <220>  
 <223> Primer

<400> 9  
 aatatgcaca aaacttgac 20

15 <210> 10  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

20 <220>  
 <223> Primer

<400> 10  
 atttgaggct ctaaagggtc 19

25 <210> 11  
 <211> 19490  
 <212> DNA  
 <213> Artificial Sequence

30 <220>  
 <221> misc\_feature  
 <222> (809)  
 <223> Expression plasmid where n can equal a, c, g, or t

35 <220>  
 <221> misc\_feature  
 <222> (995)  
 <223> Expression plasmid where n can equal a, c, g, or t

40 <220>  
 <221> misc\_feature  
 <222> (1046)  
 <223> Expression plasmid where n can equal a, c, g, or t,

45 <220>  
 <221> misc\_feature  
 <222> (8220)  
 <223> Expression plasmid where n can equal a, c, g, or t

50 <220>  
 <221> misc\_feature  
 <222> (8283)  
 <223> Expression plasmid where n can equal a, c, g, or t

55 <220>  
 <221> misc\_feature  
 <222> (13101)  
 <223> Expression plasmid where n can equal a, c, g, or t

60 <400> 11

gctcgacctg caggtcaacg gatccacttc acctgttaggc aaggcacagc acaggggtga 60  
 gcgaggccac agccctgccc ccgagcccca cccacccctc agggcactga ggccacctct 120  
 65 ctgccccaaa ggcccaccca cccgtcagtc cacgcaggcc acagccctgc ccctgaggcc 180

5	catccggccc	ctcatgcac	ccaggtgcac	tggcctcacc	actgcctgtc	ctgagggttg	240
	gtgatagaga	gcagagccag	ggcacccaaca	gcatgtggac	agcacagaag	acagcgtca	300
	ggacagggtgg	ggacagcgtg	ggggacagtg	tcagacacag	gtgaggacag	tgtggggac	360
	agtgtcaggg	acaggtggag	acagagtgtg	gaagagtgtt	ggggacagtt	gaggacagca	420
	tggaggagag	tgttggggag	agatggggac	agtgtcaggg	agaggtgggg	actgtgtga	480
10	ggacagcatc	agggacaggt	ggggacagca	tggggacacag	tgtgcatac	aggagggac	540
	ggtgtgggg	acagtgtcag	ggattataagg	ggacagagtg	ggagacagtg	tcagggaccg	600
	gtggagacca	tgtgggggac	agggtggggac	agcatgggg	acaggtgggg	acagcatggg	660
	ggacagtgtc	agggatagga	ggggacacgga	aacagtgggg	acattgtcag	ggacaggggaa	720
	gatagcatgg	gggcacagtgt	tggggacatg	ggggacacga	tggaggacag	tgttggggac	780
15	tgggtggggac	agcatgggg	acagtgtang	agacagggtt	ggacaggatg	gaggatagtg	840
	ttggggacag	gtggggacag	tgtgggggac	agtgtcaggg	acaggagggg	agatcgtga	900
	gaacagtgtc	cgggacaggt	ggggacagca	tgggggacag	tatcaggggac	agggtggggag	960
	agtgtgggg	acagtgttgg	actggtaggt	acagnccggg	ggacagcatc	ggggacaggt	1020
	ggggactgca	tgggggacaa	tatcanggac	aggtggggac	atggaggaga	gtgttgggaa	1080
20	caggtgggg	cagcatgggg	gactgtgtg	aggacaggtg	gggacagctt	ggggacaggt	1140
	ggtcatacgt	ggaggggacg	gcgtkggga	cagtgtcagg	gattataggg	gacagagtgg	1200
	gagacagtgt	cagggacagg	tggggacagc	atggcggaca	gtgtcaggg	taggaggaa	1260
	caggaggaaa	cagtggggac	attgtcgggg	acagggggga	taggtgggg	gacagtgtt	1320
	gggacaggtg	gggacagtgt	ggagggaaagt	tttggggact	ggtggggaca	gcatggggaa	1380
25	cagtgttaggg	gacaggtgtgg	gacaggaggg	gacagcatgg	aggatagtgt	tggggacagg	1440
	tggggacagt	gtgggggaca	gtgtcgggg	caggagggga	cagcgtgggg	gacagtgtca	1500
	gggacaggtg	gggacacgat	gggggacagt	gttgtactgg	tagtacagc	ctgggggaca	1560
	gcattgggg	caggtggaga	ctgcatgggg	gacaatatca	gggacaggtg	gggacagcat	1620
	ggaggagagt	gttggggaca	ggtggggaca	gcatggggga	cagtgtttag	gacaggtggg	1680
30	gacaaacgtg	gggtacagtg	tcggagatgg	gtggggacag	catggaggaa	agtgtcgggt	1740
	ttaggtaaaga	acaacgtgga	ggagagtgtc	ggggacaggt	ggggacagtg	tggggatata	1800
	tgtcacggac	agggtggggac	agcatgtggg	acagggttcc	atacaggagg	ggacagcatg	1860
	ggggacagtg	ccaggtactg	taggggacag	cgtggggac	agggccagga	actataggg	1920
	acagagtggg	ggatgggtgc	agggacaggt	ggggaaagca	tgggggacag	tgtcagggac	1980
35	agggtggaaac	tgtgggggac	agtgttgggg	acaaaggggg	acagggtggg	ggacagtgtt	2040
	ggggacagat	ggggacagca	tgggggacag	tgtcggggac	atgtggggag	agcctgggg	2100
	acactgttgg	acaggtgggg	gcagcattgg	ggacaatgtc	aggacagtt	tgtgagagca	2160
	tgggggacag	cgtcagggac	agggtggggac	agcctgtggg	acagtgtcag	agacagttt	2220
	tgacacgtg	ggggacaatg	tcaaggacag	ctggggacaaa	cgtgcggccg	accttgaaga	2280
40	agggtacggt	ggcactgtag	cacacgctt	acaggaagag	tgtatgaag	atggtgatgg	2340
	tcttccacag	cccgtccagc	tccccgtct	gcccctccg	acagctctcc	tccagtttca	2400
	gctctggaca	ggaagggggt	ggtcagtgtc	gtgtccccct	ggcttggggc	ctctgggggt	2460
	gattccctct	gtgggggggc	ccaggatgt	gggccccggcc	gggatggggcc	aacaatgtcc	2520
	tgaggtcagc	tccccacagc	tgcccgcct	gggcaccaggc	tttggggcccg	ggactcagcc	2580
45	agacacccgg	ccctagatag	cgacctggcc	ctcagcagga	cccgtcccc	gtctccctgt	2640
	tccctccctg	agggccagag	ggcaggagga	tggtaagcc	cacacccat	gtgacccca	2700
	ctgcagggaa	gggcggcatt	gggaagtggg	ccagtgcag	ggacggcagc	tggcgtgt	2760
	tccctctgt	gtggggggct	gtgtgtgtgt	ggcggctca	ggggcacctt	gtgagaggag	2820
	ggctgggttt	gtctgagctg	gtcagcatgt	ggagaagctg	ccgagcgggt	cgtggccctt	2880
50	gaggtgcgc	gtggggctcg	tggggccctg	tgtccgagga	gtttcacgt	gtgcgaggac	2940
	cttgcctctgg	tctgggtgt	gtgcgggtcg	cccggtttag	gtccctgtgt	tgaggcgtgc	3000
	acgtgtgtgt	gtggtgcccg	tgtggccggc	caacctca	gcggggtttg	ttgaacgggt	3060
	ctgggtctag	tgtgtgtgt	ggcatctgc	ccagtctctc	cacaggcccc	gagagtgtcat	3120
	gtccccagga	gtcgggtgt	tccccatgt	ggtgcgaggc	tggcaggccc	tgccagggggt	3180
55	tagtgcctg	ggggtagatg	ggtgaggggag	ggcctgtccc	tacgcgcatt	gactaggcat	3240
	gcccccgagt	gggcacatcg	gtcgaggac	agggcgtca	cagcacagga	cagtctctta	3300
	cagaggcagg	ggctgtgtgt	ctgtccccag	gggtctctag	ggcttctcg	ggcccagccc	3360
	agggcagctg	ctgtcggagg	gagggccacg	ctggcaaaatc	ccccacccctg	ccgagggcag	3420
	cccctggctg	agccccaccc	tagggcccc	aggcacacct	gcacagcctg	ggccagttgt	3480
60	gggcacagtgg	gacccgctt	gcctccctca	tgccacttc	gcctcagact	ccgcctgacc	3540
	cgtggaaaga	accatcacag	tctcgcqqqq	ccccaggqca	qcqctqqqtq	ctttatattcc	3600

5	atgctggcg cccgggaagt atgtacacgg ggtacgtgcc aagcatcctc gcgcgacccc 3660
	gagagcccg ggagcagggg cttgccggcc ctggcaactca ttacccgga gacaggaga 3720
	ggctttctg cgttagtgg ttgtcagag cctcatgcat cacggagcat gagaagacgt 3780
	tccctctgtc ccacctgtc ttgtccacgg tgagcttgcgt gtagaggaag aaggagccgt 3840
	cgaggtccag cacgggaggg gtggtcttgc agttgttctc cggctgccc ttgctctccc 3900
10	actccacggc gatgtcgtc ggatagaagc ctttgaccag gcaaggtcagg ctgacctgg 3960
	tcttggtcag ctcatcccg gatggggca ctctcaggc agaggggtggg ccgagccggc 4080
	ctgtagggac agaggttgc acagcggtca ctctgtccat gtggccctcg cacccacgg ggggttacac ctgtggttct cggggctgcc 4020
	tcgatggggc ctggggagggc ttttggag accttgcact tgcactcctt gccattcagc 4200
15	cagtctggc gcaggacggt gaggacgtc accacacggt acgtgttgcgtt gtactgtcc 4260
	tcccgccgt ttgtcttggc attatgcacc tccacgcgt ccacgtacca gttgaacttg 4320
	acctcagggt ctctgtggc cacgtccacc accacgcgt tgacctcagg ggtccgggag 4380
	atcatgaggg tgccttggg ttttgggggg aagaggaaga ctgacgggtcc ccccaggagt 4440
20	tcagggtctg aggaagagat ggaggtggac gtgtcagcac cggctggggc cctgtccctg 4500
	gatgcaggct actcttagggc acctgtcccg ctttaggtc gaggggcagg cctgggtctgg 4560
	cttacctggg cacgggtggc atgtgtgagt ttgtcaca gatttgggct ctgcagagag 4620
	aagattggg gttactggaa tctgggagga gagaagggtg ccgagctgag ggagtggaga 4680
	gtttggcctt tgggggtggc tttagtgcagg ggcagggtcc tccggatat ggctcttggc 4740
25	aggtctgagc ccagcacctg cccctttgtg tgcagggtc gggtagggg cacctagct 4800
	gtgcctgccc agagcctggg gaaaaagcca gaagaccctc tccctgagca tgagtggggc 4860
	gggcagaggc ctccgggtga agaggcagac ggggcctgc ttgtgtccct ggactggggc 4920
	tgcatacgcc ggtatgcgtcc aggcaggagc gctgagcctg gctccagca gacaccctcc 4980
	ctccctgtgc tggcctctca ccaactctt tgcacccctt ggtttgtctg ggcttgtat 5040
30	tcacgttgca gatgttaggtc tgggtggca agtgcgttgc gggcacggtc accacgtgc 5100
	tgagggagta gatgtctgag gactgttagga cagccggaa ggtgtgcacg ccgctggta 5160
	gggcgcctga gttccacgac accgtcacgg gttccccggaa gtagtcttgc accaggcagc 5220
	ccaggccgc tgcgtccccca gaggtgtct tggaggaggg tgccaggggg aagaccgtat 5280
	ggcccttggg ggagggtcga agagagggtgg tgccatgtga ccgggtgtg ggacagagct 5340
35	gggcccaggg cgcagagggc cctgggtct taactgtccg cgagggtcag cgtccagtgt 5400
	ctgggctcat gggcattggg tgcacactg gtcggccatca cttagcccc 5460
	tccctgcccc aaagccaagg tcaggcctgg cctgccccag aaagcttcag ctgctcgacc 5520
	tgcagggcatg caagcttgcg ggacccggc cctgtgtg ccctctgca ggcaccctcg 5580
	cagcttagga ggcgggggtc ggcagccagg tcaagcgtct gtgcctgccc ggagtca 5640
40	cagtccagtg tctctagtt ggcctcagct ctggccatcg gtgccaccc agggacggct 5700
	catgcccatt ggcggccatc cagccttta tgggtgcctg gttgaccag tggacactgt 5760
	tctcagatgg ctcttgggtt gtcggcccttgc cccctgtaa gccctgaccc tggcgctcca 5820
	gcatggccct cccctagtgc gtggccctga cttgcccagg gcctgttca tagcctgccc 5880
	tctgccttcc aaggccctt tcttctgtgc agcagagggg ccagacactg catagggtcg 5940
45	gccccttca gccccagggc cccgaaaccc ctcgccttg aatatcgccc cgggagcc 6000
	ctcctcagcc ttccttccctt ttccttctta gcccaggatgt gcagcagccc aggtcaggc 6060
	cctgagtgcc tggatgcggc ctgcctccca gtgcctgtca ttacttctgg aggtcagtc 6120
	accacaacctt caccctccca gcctggctt ggccttcgt gccaccagcc caccctcc 6180
	ctctctccag agttttcccc ggcaagggtcc ctgctgggtt caaccctggc ccccccagcac 6240
50	aggttaggagc ttgcacactg cccttggccc tccccaccc tgcattggcc aggacccca 6300
	ggccacaggg aggccccatt tctctgtcc gtcggcccttgc tggcccttgg gtcggactgc 6360
	aggtgggggtg tgccttgcg ctctgaggag gtaagtgcg ctgcctctag ccaggccatc 6420
	ccctctgtc agccccaggg ccccgctcac cacccttcc ctcacactgc accacaggct 6480
	ctggctgact ctgcctcaggc cctgaatggg cccctcttgc tgcctctgc tgctacactg 6540
55	tcctgcacca ctccactca gtttattgt gtcgggtggcc ctgcctctgc gcaagccatc 6600
	ttgctcttcc tggggcacca gcctcagagg ctttctgtcc cagggtccgc tgggaccagc 6660
	cgtgggaccc tccttggctc aagcacacgt tccccctgc gccacacctg cccctgcctg 6720
	agagcccagc cccgaggccct ggaacgcctt cccttcttca tccagctcg cccttgc 6780
60	ctgctcagtg gatggactc acactccctt cccggacca ggaggctgc ctgcactt 6840
	accagccctc agctgtctgt tgcctcagcaac taccctgc tgcctaaatgctt ctaggagctg 6900
	agtatgcct cccaccaggcc ctgctcactt gtggctgcct tgccttgcac tcttagtgcct 6960
	gtcccttggg gcaacttagc ccagctcagc tcaaccctgt tcaactcagc 7020

5	ccagttcagc tcagtcagc ccagttcagc cttgtttagt ctaggtcagt ttaggtcagt 7080 tttgcacccatc tgagtccatt tctgaaagct ggatggagtt gtcacggcca gaaatggca 7140 gcccaccaga cctgtttgtc tcagctaaag ccatctcatt gccaggttcc tgcacagcca 7200 ggctggcttc catctttgt ctcctctac ttgatacccc agtccctgc agtcctgccc 7260 cagcgccacc tgggttttg ttccaaagca ttaccaatca ttaccacccct ccactacctg 7320 ggtggaatat ttctttgtc cttttaaagtc attaaaacat cttgagaatg agacccaagaa 7380
10	tttaggagcc tgtgctgtg taaaaatgag caggcccct tgctctagaa gtggcagcat 7440 atcttctgca ccaagaggag ggtattgaga tgctcagac ctcacccctt ccggagacatc 7500 ccctcccttc tgagtctgca gtaaaccctt gcctttaaat tccctctaga taacagtcat 7560 cattggaaac aaccaagaaa tgcatttat ctgaatttgc cacttaaat tctgcccattt 7620 accataaaatc gcttttggaaag gcatgggcta cttcaaggg tgcatgtatg acctacatgc 7680 aatgacttag acaagggcga tgccagtggg gcttggatg ttctcaagca tcattaccca 7740
15	tgccatcccc attcagaggt tttggaaacag agggaaagtt aaatgggagt gaccagaca atggtcactc aaaagactca cataaaatgag 7860 tctctgtctc ttcatcaagc aattaagacc agtcccccctt ctatgtggaaa taagacgtta 7920 aatacaaagt ttaagagaaa gcaaattgcag cagccggccgc tgctgtctc ttaccatgtc 7980 gggcgcctgg tcaactgcgag ctttgcggaaat ctttggcatg gaatcattcc tccaagtcca 8040 ttaacaaggg ctggggccctg agcagccagt cggccggca gcagaagcca cgcattccag 8100 ctctgggttag tccggggaga cccaaagccc aggccggggcc tggcagccac cctcccaagag 8160 cctccgctag gccagtcctg ctgacgcccgc atcggtgatt cggaaacagaa tctgtccctn 8220
20	taagggtgtct ccacagtctt gtcttcagca ctatctgatt gagtttctc ttatgcccacc 8280 aantaacatg ctttaactgaa ataattcagg ataatgtatgc acattttacc taaaacttat 8340 cctaaagtga gtagttgaaa agggcttga aaaatactaa aatgaaggcc actctatcag 8400 aatatcaaag tttttctctt taatcacaat gaaaaacgaa gttaaacctaa aaagattgtg 8460 aacacagtca ttatgaaaat aatgtcttga ggtatcgaaa aatgttttga gatttagttat 8520
25	30 cacatgaagg gataacaagc taattttaaa aactttttaa atacagtcat aaactctccc 8580 taagactgtt taatttctta aacatcttac tttaaaaatg aatgcagttt agaagttgat 8640 atgctgtttt cacaacttag cagttgataa gctaagattt gaaatgaaat tcaagatgtt 8700 aaaaaaaaagcc ttttcagttt cggtcagccct cgccttattt tagaaacgca aattgtccag 8760 35 gtgttggttt gtcagtaga gcacttttag atctggccct gggaaaaacc acctcttcac 8820 aaccagaagt gataaaattta ccaattgtgt tttttgtttt ccttaaaatag actctcgccg 8880 tgacctgctt cctgccaccc gtcgtgggtt ccggagaccc ccatgcagcc atcttgactc 8940 taattcatca tctgtttcca gtttcgttca attaattttaaa aaaataaaact tgatttatga 9000 tggtaaaaac gcagttccgc atcggggccg acagcactgt gctgttattt cttagctgag 9060 40 cttgtttgg cctcaatttttcc agacacatat cactcatggg ttttaatcaa atgataagaa 9120 tttcaaaatac ttggacaggta aaaaaattta atatacttga aaatctctca catttttaag 9180 tcataattttt ctttaaccatt ttttcagaa gccacttcaaa acatatcttgc tcttttaaca 9240 gtaagcatgc ctccataagat aaacaatctt ttttcatgg aaaccagttt caaggcactg 9300 aggctcttggaa gcctccctaa gcccgttca ggacggcagc cactgtttt gggctacccc 9360 45 tgcccccaac cctgttctca tcaagacccgg ggctacgcgt ccctcctggc tggatttcacc 9420 cactccgaca gttcttttc cagccaaaaaa agaatctaag atgcaggttgc acacacagcg 9480 cacctcataa ttctaaagaa aatatttccat gattcgttgc tttgtcagcgatc tcttgcagtc 9540 ctacagacac cgttcttgcg acacatttcc tttttttttt cttttttttt gttttttttt 9600 ggcatttcgc ccaaattgtgg cggcccaagc ccccaaggctc agtacttcca tcagacgcac 9660 50 ccaacctttag tcccattttcc caaaggcatc ggaaaatcca cttttttttt cttttttttt 9720 aggcacccca gtggccgttcc cttctggcc agtccgcacca ggtttttttt tttttttttt 9780 cgaggaccaa ctttgcataatga tcagaaaaacc ccacaggccg tagcagaaaaaa cttttttttt 9840 agaatggcaa ctgactgtcc gtggccctgt tttttttttt tttttttttt 9900 55 ctccccacgtt ccccacattt tggcttgcacc cttttttttt tttttttttt 9960 aagcccaact gcaagcagac ggcttcttaag tttttttttt tttttttttt 10020 gcacccgttcc cttttttttt tttttttttt 10080 cgacagcccccc tttttttttt tttttttttt 10140 gagccccccca ctttgcgttcc tttttttttt tttttttttt 10200 60 cttttttttt tttttttttt 10260 ccttgatggcc tggcaagttt tttttttttt 10320 caccttggat tttttttttt 10380 gtgacaacaa agccaggacc ccaggcaaga tttttttttt 10440



5	agccaaagca gatgtttcac ggagagtgtg ggcagggca tgctgcagg tcgagtgtt 13920
	tctacataag ttatgggacc atagcagagt ttcccaactc ttctgagtca ctttctgatc 13980
	ctttagtgtc ttgcccacct ctggtcagta actccttcga tctccctta attaattaaa 14040
	ctctgttgc agcccaattg ccaccagcag ctagtttcc taaaagggtg tagtgtgtt 14100
	tggtatggc agagcattcc actagagcca tgagaaaatc ggagttctac aaggacacag 14160
10	ccatggataa agacaattcc tactcaagct agctgaagaa ctgtaaattt gaaagccagt 14220
	gtctgagcgc cctctactgg aagcaatggc acattactcc cagcatcatt ttccagcaca 14280
	aaagacctaa attgttggaa tgagacattt gcttagcatt ttcttgcactc tcactttcag 14340
	tttctccag aactcctgat gttgacacaa gacatgacat atgaagggtt aagtttaatta 14400
	aaagacgctt tccagcatcc aagatgaggg aaaggtgaga aatactgaga ggctgtgtaa 14460
15	aatggagtcg tttcttgcgt catgtaagg ataaatagca tccttgcga ttcttgcag 14520
	cctgcccaga aaaatgtgag ttctagcagc agtttctcag ttatcactga cagaacctga 14580
	ttgcttaatg atttataatgtt ttattacata gtggccaaga gcaagccctt ctggtaagtc 14640
	acttcttcaa tttgggtgaca agaaaccagc catcaggaat gatgtatttc ttgttaagtt 14700
	ccttgagaaa tacgagattt ttccttgcgt tgcttgcgtg tgggtttat gcccgtgtt 14760
20	gtgtgtgtgt gtgtgtgtgt gtatgtgtgt atgtgtatgt gtgtgtgtgt ttgtatgtt 14820
	gtgtgtatgt atgtgtgtgt gtgtatctct agtccctgtct ttttctttct actgactttg 14880
	cagttgtctt gattatgtt tggtttacta cagtaacttgg gatccgttga cctgcagggtc 14940
	gaccggatcc agacatgata agatacattt atgagtttgg acaaaccaca actagaatgc 15000
	agtaaaaaaa atgttttattt tgtgaaattt gtgtatgtat tgctttatattt gtaaccattt 15060
25	taagctgcaa taaaacaagtt aacaacaaca attgcattca ttttatgttt caggttcagg 15120
	gggaggtgtg ggaggttttt taaagcaagt aaaacctcta caaatgtgtt atggctgatt 15180
	atgatctcta gtcaaggcac tatacatcaa atattccctt ttaaaccctt tacaattaa 15240
	aaagctaaag gtacacaattt ttgagcata gtttataata gcagacactc tatgcctgt 15300
	tggagtaaga aaaaacagta tggatattttt ataaactgtt tgcctactt taaaggttac 15360
30	agaatattttt tccataattt tcttgcgtat cagtcagct ttttccttgcgt tgggttaat 15420
	agcaaaagcaa gcaagagttc tattactaaa cacagcatga ctcaaaaaac ttagcaattt 15480
	tgaaggaaag tccttgggtt ctctacctt tctcttcttt tttggaggag tagaatgtt 15540
	agagtctgca gtagcctcat catcaactaga tggcatttct tctgagcaaa acaggtttt 15600
35	ctcattaaag gcattccacc actgctccca ttcatcagtt ccataggttga gaatctaaaa 15660
	tacacaaaca attagaatca gtagtttac acattataca cttaaaaattt ttatattttac 15720
	cttagagctt taaatctctg tagtagttt gtccaaattat gtcacaccac agaagtaagg 15780
	ttccttcaca aagatccggg gcccactcat aaatccagtt gcccacgg tagccaatca 15840
	ccgtatcgta taaatcatcg tcggtacgtt cggcatcgat catcacaata cgtgcctgga 15900
40	cgtcgaggat ttgcgtggg tcaatgcgcg gccagatcca catcagacgg ttaatcatgc 15960
	gataccagtg agggatgggtt ttaccatcaa gggccgactg cacaggcggt tgcgcgcgt 16020
	gattaaagcg gcggactagc gtcgagggtt caggatgttt aaagcggggtt ttgaacaggg 16080
	tttcgtctcg gtttgcctgt gtcatggatg cagccctccag aataacttact ggaaactatt 16140
	gtaaccggcc tgaagttaaa aagaacaacg cccggcagtg ccaggcgttga aaaagattag 16200
45	cgaccggaga ttggcgggac gaatacgacg cccatatccc acggctgttc aatccaggta 16260
	tcttgcgggaa tatcaacaac atagtcata accagcggac gaccagccgg ttttgcgaag 16320
	atgggtacaa agtgcgtttt tggatataattt tcacgaatcg caaccgcagt accaccggta 16380
	tccaccaggat catcaataac gatgaaggct tcgcccattcg cttctgcgtg tttcagcact 16440
	ttaagctcgc gctgggttgcgt tgcgtatgttgcgt ctggaaatatac aaacggatc gacatgacg 16500
50	ataccaggat cacgcgcaccatcga taacgcaccc ggtaccagac cgccacggct tacggcaata 16560
	atgccttcc attgttgcata aggcatacgat cggcttgcga gtttacgttgc atggatctgc 16620
	aacatgtccc aggtgacgat gtattttcg ctatgtgaa gtgtcccagg ctgtttatct 16680
	acggcttaaa aagtgttgcgaa gggggaaaata ggttgcgcga gattatagag atctggcgca 16740
	ctaaaaacca gtatttccata tgagtcccg tcttttacg cactgcctct ccctgacgcg 16800
55	ggataaaatgt gtattctcaa acatatctcg caagcctgtc ttgtgtccaa gtagctttt 16860
	tgcaaaagcc taggcctcata aaaaagcctc ctcaactt ctggaaatago tcagaggccg 16920
	aggcgccctc ggcctctgca taaataaaaaa aaatttagtca gccatggggc ggagaatggg 16980
	cggaactggg cggagtttagg ggcgggatgg gcggagtttag gggggggact atggttgtct 17040
60	actaatttgcgat atgcatacttc tgcctgttgcgtt ggagcctgttgcgt tgcgtactaa 17100
	ttgagatgca tgctttgcata acttctgcgtt gctggggagc ctggggactt tccacaccct 17160
	aactgacaca cattccacacg ctgcctcgcg cgtttcggttgcgtt atgacgggtga aaacctctga 17220
	cacatgcagc tcccggagac ggtcacagct tgcgttgcata cgatgcccgg gagcagacaa 17280

5                    gcccgtcagg gcgcgtcagc ggggtttggc ggggttcggg gcgcgccat gacccagtca 17340  
 ctagcgata gcggaggtgt tactggctta actatgcggc atcagagcg attgtactga 17400  
 gagtgcacca tatgcgggtg gaaataccgc acagatgcgt aaggagaaaa taccgcata 17460  
 10                ggcgccttc cgcttcctcg ctcactgact cgctgcgtc ggtcgttcgg ctgcggcgag 17520  
 cggtatcagc tcaactcaaag gggtaatac ggttatccac agaatcaggg gataacgcag 17580  
 gaaagaacat gtgagcaaaa ggccagcaaa aggccagggaa ccgtaaaaag gccgcgttgc 17640  
 tggcgtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 17700  
 agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 17760  
 15                tcgtgcgtc tcctgttccg accctgcgc ttaccggata cctgtccgoc ttctccctt 17820  
 cgggaagcgt ggcgcgttct catagctcac gctgttaggt ttcgttagttcg gtgttaggtcg 17880  
 ttcgcctcaa gctgggtgt gtgcacgaac ccccggtca gcccgcaccgc tgccgccttat 17940  
 ccggtaacta tcgtctttag tccaaccggg taagacacga cttatcgcca ctggcagcag 18000  
 ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 18060  
 20                ggtggctaa ctacggctac actagaagga cagtatttgg tatctgcgt ctgctgaagc 18120  
 cagttacctt cggaaaaaaga gttgttagct ttgtatccgg caaacaaaacc accgctggta 18180  
 gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaaaagga tctcaagaag 18240  
 atcctttagt ctttctacg gggctgtacg ctcagtgaa cgaaaaactca cgttaaggga 18300  
 25                ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaat taaaatgaa 18360  
 gttttaaatc aatctaaatg atatatgagt aaacttggtc tgacagttac caatgcttaa 18420  
 tcagtggggc acctatctca ggcgcgtc tatttcgttc atccatagtt gcctgactcc 18480  
 ccgtcggtta gataactacg atacgggagg gcttaccatc tggcccccagt gctgcaatga 18540  
 taccgcgaga cccacgcgtca cccgcgtccag atttatcagc aataaaaccag ccagccggaa 18600  
 30                gggccgagcg cagaagtggc cctgcgtactt tatccgcctc catccagtct attaattgtt 18660  
 gccgggaagc tagagtaagt agttcgccag ttaatagttt ggcgcacgtt gttgccattg 18720  
 ctgcaggcat cgtgggtgtca cgctcggtgt ttggatggc ttcatcgac tccgggttccc 18780  
 aacgatcaag gcgagttaca tgatccccca tggatggc aaaaagcggtt agctccttcg 18840  
 gtcctccgat cgttgcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 18900  
 35                cactgcataa ttctcttact gtcatgccc ccgtaaatgtt ctttctgtg actggtgagt 18960  
 actcaaccaa gtcattctga gaatagtgtt tgccggcgacc gagttgtct tgcccgccgt 19020  
 caacacggga taataccgcg ccacatagca gaactttaaa agtgcgtatc attggaaaac 19080  
 gttctccggg gcgaaaactc tcaaggatct taccgcgttt gagatccagt tcgatgtaac 19140  
 ccactcggtc acccaactga tcttcagcat ctttacttt caccagcggt tctgggtgag 19200  
 40                caaaaacagg aaggaaaaat gcccggaaaa aggaaataag ggcgcacacgg aatgttgaa 19260  
 tactcatact ttccctttt caatattttt gaagcattt tcagggttat tgtctcatga 19320  
 gcggtacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 19380  
 cccgaaaaagt gccacactgac gtcataagaaa ccattattat catgacatca acctataaaa 19440  
 ataggcgtat cacgaggccc ttctgtttc aagaagaatt ccaagctatt 19490

45                <210> 12  
 <211> 19412  
 <212> DNA  
 <213> Artificial Sequence

50                <220>  
 <221> misc\_feature  
 <222> (790)  
 <223> Expression plasmid where n can equal a, c, g, or t

55                <220>  
 <221> misc\_feature  
 <222> (13735)  
 <223> Expression plasmid where n can equal a, c, g, or t

60                <400> 12

ttccaaagcta tgctcgacct gcaggtcaac ggatccacat atataattta gaagtttca 60  
 ctagccatgt tataaaaagta aaaagaaaatc tgtgcaatca attttataaa tattttgtt 120

5	tgcacacaag atatacaggg taatctcaac acgtatttt tgtacaaaat gttaaaaatt agtatggact ggccttatct ccagagctca ggatatcaag tctagaacat acgtgaatta gagttatgac tgaagaacaa gttcaaggag ctactccttc cattactaat gataacaaat tgatgcttag aactttatata gaaactgtttc gaccccccacc cccttcaga tgagaaaact tgacttgtga gcatattttc ccaggcaatt ccaataattt attcactata ttctccatg ggtcacccac ttgtgctccc tcagtgaggg agtactcan gtttgggatc tccagcagct gttagatcct accctgataa ctgttataaa acccagccca aggggcactt tctatggaaa actgggggtg caggttcaa gaaagaacaa gaagagttaa gtcttggaaat tgtaatacc aaacacacaa acatcaatct gtcctcctct atcagtcaga ctttgccttc tcttctct tttcttctta tttccattt attgttcttc ctttctcttt tctccaaaat aatatgttca gcaatgtat tagcacccaa ccttagaaac aaaaaaagagt taaaaataga aagtcacagt tgatgtatag acatttagat cagaaaaaga cacagacagt gtgagggagg gacaagggga tagtggtctg ataagagtgt cagtgaggt gtgacactaa acatttattt ttatactaac tttcttactg tgccttagtc tccatggcat atcatttcctc tcccacagtt gatgtctggc tctctcagaa tgactgctgt ttcatgctt tctgacctgg agaaaaggata cccatcctac cctggtttc cttggcagac acattttct tttccctccc tcttgattt cttcttttt tagatctacc atgggaccca gggagcttg tttactcatg ctatttctt taatcttcac atctgcaccc tacaggttag aaaggaggc acacagctg tacatgtcag agtaaggca aaagcatacc acgtctcagg agaaggcaggc atgctgaccc gaactttctg gatttgatg gagttggattt tgccagccac ctggggaggt tagaaggggag ttcccttcaga ggtacgtgt ttgggtctgc cctaacaaaag taccacaaac tgtcataggt ctggagacag aaatctgaat caggttctcg gggagaaaacc atcccacacc ttgggtctct ttggctctca gtcgtatccc ttcttcttat gtgtcactgc atctccaatc gatttagggc ccactcgat ccagcataaa attatagatt cacagaaaatg agcaagaca tgcacccagt ttctcctgtt ggttacttct gaatttgaca tgataaaaat aagtgtgtac tgcctgcaac taccactgca atcaagatac tttatagttt cgcctactct tcttctcca tctctatccc tataattttg tcacttcaac tgaccctttc agattggctt ttcttactca tctgtatcaa cacttcgttc cttttattt acaatttggtt aggcttagttg cctattgagg aacaataaaa gtcgtgtga acaatcatgt tcctgggggtt aattcccaa ggagttacaat	atgttaatcat ataaaatatt aatattttac 180 ggatgtgtt ttacacgtat agcccatctt 240 ataactctgt gtgcttagtgc cagggatatt 300 attcaagtt caagcacatt aaggctgaaa 360 gaaatcaaac aatagccaaa tttcaactgag 420 atttatacat cccttagtat atactaggcc 480 atthaatctt cacagcaatc ttacaagtaa 540 gaagcacaaa agtgcagtaa ttgttattt 600 atatgcacag gatgtatgtgc gtaaaggcagc 660 ctcatggtaa ggactgactt gatccctgg 720 aagggtgc aagggccctgg cccgtgtcct 780 gctatgggtg acaattttat ttagtaaaaat 840 aaatccctct tccccttccc cagcacaggt 900 gatgaaaaga actgacatgg acttgagctc 960 agagagctgt tagttgaag agacgcagtt 1020 tggacaagg cacgggtgtga aaagacaaag 1080 tttctctcc ttttcttattt cttatccctcc 1140 ccttcctccaa ctttcttcccttcc 1200 caaaattcgt atactcatcc tcttcttctcc 1260 acttactgac ttgttcaggg atggcgaattt cttatccat 1320 gatcactatg 1380 ccagtttga agccacagaa gacccactc 1440 cttggatgtcc 1500 attgtgtcc caagcgtact ggtgttagc atttggatgtg 1560 cccactgcca actaccaggaa agtttttattt 1620 tttcccccata gatgtccaggc 1680 atcatttcctc tggatgttcc 1740 tctgtccatg ggggttggc ctgcctcttc 1800 tgcgttccatg gtagccacg cagttatgcca agtgcgttcc 1860 tctgacccatgg ccatatcatg ggtggccctt 1920 cacttttttt ttctctctt cccttcttcc 1980 ccttttctt ctcttgcctc acttctgccc 2040 gagccatgca acccacagtg tgctaagctt 2100 aattcttacat cttatcaggca ttaccatcag 2160 ccacacggat taagaaatgt gtcaggggcc 2220 ttctcatgca aagtttattt ttcagagcca 2280 tagatctttt gttctgttgc cttttaactc 2340 actctcagac ctgtggagcc cgatttctgt 2400 gaacccttgc tgggttatgg ctcatttagt 2460 ttctctcgcc atccccattt tagtagttc 2520 taagagccgt aaaacaaacag aagtttattt 2580 tcaaggccaca ggcagagtt gtgtcttctg 2640 tctccagctt ctgggtggctc ccagcaatcc 2700 tctggctctt ctgggtctct gcctccttgg tcatgtggct 2760 ttcttcttat gtgtcactgc atctccaatc ttcttcttctt tataaggaca ccagtcatgt 2820 aaaaaaagttt aaaaaaaatata ttggaaata 2880 ggatataaaag cacagaggc tgcgtgtcc 2940 tacgttaatta cagtagata taaaacccaa 3000 agttctatgt cattttatca catgtctaga 3060 agattttatcc catgatcaca aagatctccc 3120 ccacccttaac ccctggcaac cagaatctgt 3180 aatgttacgt aaaaggaatc acatggatag 3240 gtagaaaagtc cttcagaatc actccagttc 3300 cctcatagca ttctatggca tggatataattt 3360 tcctqqqatca tatqqtaqqt tttttttttt 3420 ttttttttttt tttttttttt tttttttttt 3480 ttttttttttt tttttttttt tttttttttt 3540
10		
15		
20		
25		
30		
35		
40		
45		
50		
55		
60		

5 tagttttca agaaaactgac aactatttc tagagtagct ataccctttt atattccac 3600  
 cagcaatata tgaatgtatcc attttcacccg catcctcaca aacatttcag tatgacctaa 3660  
 ttttgacttg attacctctg caaagaccct atttccaaat aaagtcacat ttatgtgacc 3720  
 catttatgg taccaagagc aaggagttga gcataccctt ttgaaggaca caattccacc 3780  
 tgcagcacct gtcttcactt gcttaggtta tgggtgggtt tcacttcaga gtgttatctt 3840  
 ctgagaagaa gtttgaatt cataactgga ttatgcctgc ctttgattgt tcaatttgcc 3900  
 10 ccccaggat gagtgagcaa gttcatgtca ttggaaggaa cattactgag tctcaccctt 3960  
 actcaaaccct cccaaagtttga aaagatttga tactcatcaa catttcagaag aggccagaagg 4020  
 gcaagggtgtc tccttatgtta gtttctccat gactgtccctt ccccaacttcc ccagactgaa 4080  
 ctggaaagaa aaacaaagca ggacaaagggtt attgtcagtc atttggggaa agggatgtc 4140  
 ttttagttac tcctgaaccc cttcagatgt taggcagaga gaggtactta gaaagggtgt 4200  
 15 ttggatggat gtagtagctga gaggggcagg gccagagtgc cccatagaat gaggaagagt 4260  
 cctcatttac atgtgcctt cccccccttccaa aacatgtga gtgttcatac atgggtttcc 4320  
 agtgtctggg caaggtaat tctgtctctt gtcccccttgc tgaaggcttg tttcatttaa 4380  
 tttctatgtta tttgtgtctcc tagctagcca aacagggcac agatgccagg aaaatagcag 4440  
 20 gcctgtgtt ttgttgagca ctttgcggcc ttggaggccag actcatttcag ggctcaagag 4500  
 ggcaggatttta ggcaggtagc tatttgcattt caggcccccc cacttaccag ctgcgtggc 4560  
 ttgacatttacttatttctt cttctatcta ttaactgggtt ataaagatgtt ctttatgtga 4620  
 agttattttgc aacgataatgt ttttgcatac acagtgacca gatcataact gcacaacccc 4680  
 acacaggaa cagaacatttca ccaagacccc agaagccctt ttttgcatttcc tcccaggac 4740  
 taactcgcccc accccccagct aagagaatttccatccaaa ttttgcatttcc ttttttacc 4800  
 25 tggtttgaa ctttatatgg atgaaaacat acagaacata tttttgcatttcc ttttttcc 4860  
 ttatgttaaca ttatttttat gagattcatc ttttgcatttcc ttttttcc 4920  
 attctcctca ttgggttatttca taatgtataa acataccaca gtttataaact tcatttctt 4980  
 gttggacaat cgggcttatttccaa ccaatttttgcatttcc ttttttgcatttcc 5040  
 30 cttgcatagt tttttgggtt gttggggacac ttttgcatttcc ttttttgcatttcc 5100  
 atagaatata ttttgcatttcc ttttgcatttcc ttttttgcatttcc 5160  
 aatctcacag ttttgcatttcc ttttgcatttcc ttttgcatttcc 5220  
 cagacaggag ttttgcatttcc ttttgcatttcc ttttgcatttcc 5280  
 tctaagtacc atgatgaact ttttgcatttcc ttttgcatttcc ttttgcatttcc 5340  
 35 ttttgcatttcc ttttgcatttcc ttttgcatttcc ttttgcatttcc 5400  
 ttgaaaaattt ttttgcatttcc ttttgcatttcc ttttgcatttcc ttttgcatttcc 5460  
 acctgtcacc cttaaaaagg actatgaggg ttttgcatttcc ttttgcatttcc 5520  
 gtctcttccaa ccatggccca ggggactgtt ttttgcatttcc ttttgcatttcc 5580  
 actgtcttgc agaggatgtt ttttgcatttcc ttttgcatttcc ttttgcatttcc 5640  
 40 ttttgcatttcc ttttgcatttcc ttttgcatttcc ttttgcatttcc 5700  
 caacagataa taatttattaa atgagaaatggg ttttgcatttcc ttttgcatttcc 5760  
 atttatttcat ttttgcatttcc ttttgcatttcc ttttgcatttcc 5820  
 gtgaaagatgtt agctggagga ttttgcatttcc ttttgcatttcc ttttgcatttcc 5880  
 gccaaaggat ttttgcatttcc ttttgcatttcc ttttgcatttcc 5940  
 45 taacacttcc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6000  
 ttttgcatttcc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6060  
 aggctgttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6120  
 tggagggttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6180  
 aggcacacaa ctttgcatttcc ttttgcatttcc ttttgcatttcc 6240  
 50 acagatgggttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6300  
 aagttcttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6360  
 aaaacacat gtttgcatttcc ttttgcatttcc ttttgcatttcc 6420  
 ttgtatgtttt gtttgcatttcc ttttgcatttcc ttttgcatttcc 6480  
 55 ttttgcatttcc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6540  
 ctcaatgtt ttttgcatttcc ttttgcatttcc ttttgcatttcc 6600  
 aacaggtcttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6660  
 ttccctgttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6720  
 tgggttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6780  
 60 gtgtcagcccttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6840  
 aaggcctgttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6900  
 ggcacacttgc ttttgcatttcc ttttgcatttcc ttttgcatttcc 6960

5	gcctcttgg aatccccctc tggggattc cacccttgg gtgggagac agtagttaa 7020 atgtaaaata agaatctttt gctggagaa gtcaacagat agggagaagt cagctgataa 7080 cagaaatagt tttttttttt ttaaaacta acttcactgt taaccaagca gttcaacatg 7140 aaagactgaa tctcttatgt ttaatatttt cttctttt aatcttcata actaatttt 7200 ttcagataat tgataaaaat aaccatggta gcaaataat gtatcactg gaaaataagc 7260 agggaaaaac atgctatgaa gatactccta tctgggtgaa ttcttgatag ctttacattt 7320 10 ttcatctggc atttaaacat taaacagtta atgtatggta catggaaattt atttcaagtt 7380 atcttattag ttttaataga gttaaaaag tgttaaaag agtttcaaa aggctctaaa 7440 atcatttga aatagttaa aacagtttg aatcgttga agttatgtt aatagagctt 7500 taaaaaggcc ctaaaatagt cctatcaagt ttttgcagac caaaataatc tccttaata 7560 tcactttga gtcagctgg gttaaacgac agcaacacaa tgacaatca ttaaactatt 7620 15 ttagagatta tgaaattaaa atactcagat taaaattttc ctatcacaga attaaggta 7680 tggaaaatat gtttaagttt ttattatcc attgctataag gtttagatat tttgtacaac 7740 tgaataaaaa tcacacactg gcagctacat ttttgcagat taaaacatg gtcacgaata 7800 tatcttattt taaaatcagt taatatacct taatggtatt taatgcacaa ttcaagtg 7860 20 attgatcaag ccctcagtgg ccaggcatg ggtgtgat ttactctgaa agaattacat 7920 atttcttctt ttttgggtg gctttgtt tttaaataca tttgatgaga ggatattgaa 7980 ataataaaat agcaactgaaa aaaaaagct taaaattatt tacaatcccc taatggaaat 8040 tttcaataat gagatatcat aatgaatgtg aattttattt ctgaaatctc taataaata 8100 gtcttotccc tggtttccc agctcagcgc ccattacgtt tctgtcttct ttcccttagt 8160 25 ggcattattt gtagtactgt gcatcaggaa agctggctac ggcagcatca atcgggcaga 8220 cacagggtgg ccacggccac tagcggcaag gcggctgccc caagagcgcg gtggcatggc 8280 caccaaagcc actcaatcga gaaagaccgc ggctctgtt acagctcgcg gtgccacggc 8340 tttcttggca gaataaaaaat gtagacaagt aataacagag gataatgaaa gaacataact 8400 ttttaaataat ttcttatttt tttcacagac ccacggcat taaaatgc aattatttac 8460 30 ttttttcat ttaaacacat ttcttgaga ttgagctttt gggataaacc accttccac 8520 cattacaata agagataatt tcacgtttag tctaattgtac aaattggatt tttaaaaat 8580 gagctctat ttttgcggcc ttattcttata agaatgtgtc tttttagttt tattactt 8640 tacagactct aaaaacaaca ttgctgtga ttttcaagta agctgccttct tctacatagc 8700 aaataggtac acttcacttt tccctgat ttcttagggc gtgttattgt tttttattgt 8760 35 ttttctgtat catgcataat cagcaacaac caatacaata tttggcaaga gtgacaaaa 8880 ataaattttac ttttgcctt tagaaataca agggttccctt ttttagttaca tttttttttt 8940 ttactttgtg tcattcagtt tagagcaatt taatctttt ttctccaaat ccatttttga 9000 agctgagttt aacttttgc acccatggca aatcttaat gccccttaccatctt 9060 40 accaaactcc ttttaagcc tctaaaagtc aatactggcc atcagacccaa aatttcagaa 9120 gacaatagtg aaaaattact tacgttaat ctccagtcgt gtcccttggc cgaaggatc 9180 cacagtgtt acttaattac tttccctta aaaaaatctt cttttcgtt ttaatatc 9240 taacctgacc gatgcagaga aatcttgcattt gtcacttac tggctagcgc 9300 ttggctgttc ctaagatga actaattttc tatcccttac tcatctgact ttttggaaa 9360 45 atctggact ctttggattt gacctgagct aatatctcaa acacaaaaac gtccttggc 9420 taaaaccttta taagaaaaag cattaggaa gtgcacttac gtttgcatttcc caccttggc 9480 cctccggcga aagtggcca cagtggggat ttcacccctt tccctcaac aaaaacctt 9540 cttgaagccca atcatttgcattt ataggctgtc ttttgcggaaa 9600 50 ccattttttt catgaatccctt attctcttcaat ctttttttcaat gattcgttca ttttgcctt 9660 tttttaagtt cattttcttgc ttttgcatttcaat gtttgcatttcc gtttgcatttcaac 9720 caaaatatca ctttttttgc acaaaatgttcaat ctttttttcaat gtttgcatttcaac 9780 gacagatgtt ctttttttgc acaaaatgttcaat ctttttttcaat gtttgcatttcaac 9840 attcgttcaat ctttttttgc acaaaatgttcaat ctttttttcaat gtttgcatttcaac 9900 55 ccaatgttcaat ctttttttgc acaaaatgttcaat ctttttttcaat gtttgcatttcaac 9960 aacaattttcc ctttttttgc acaaaatgttcaat ctttttttcaat gtttgcatttcaac 10020 gactaatgtt ctttttttgc acaaaatgttcaat ctttttttcaat gtttgcatttcaac 10080 agtggaaaaa agtacttacg ttttgcatttcaat ctttttttcaat gtttgcatttcaac 10140 taaatattat actgttggca gtaataatgtt ctttttttcaat gtttgcatttcaac 10200 60 gtgagagtga aatctgtccc agatccactg ccccttgcatttcaat ctttttttcaat gtttgcatttcaac 10260 aaactggatg cagcatagat ctttttttgcatttcaat ctttttttcaat gtttgcatttcaac 10320 gctaaccaggc ttttttttgcatttcaat ctttttttcaat gtttgcatttcaac 10380
---	--

5	gatgcagaca gtgaggatgg agactgggtc atctggatgt cacatctggc acctgagatt 10440 ggaaacataa aaacaaatgt ccacacaatt aatcatgtt taagagaatt tccctgaata 10500 gtaaagcagt actgagcacg ctgggcttag taaaactgcta gtgttctcca tccttacctg 10560 ggaaacagag cagcaggagc cccagggagct gagcggagac cctcatgtcc atgctgtgtc 10620 ctgactgggt ctgattcctt cacaaggatctt gaccagccta ttaataaggc ttcaaggccag 10680 gaggttgtgc tctggaaaca tgcaaatttag gaggggatgg ggcaggctgg gcacagctgc 10740 agagctggcg catctgagta actcagcacc agctcagttt ccccacctgt ccaggtaag 10800 atcaaggtag ctc当地atggatgtt acttggggact atttttttttt 10860 ggggaaacatt ttatggttt tttttgacaa ttttggaaatat tccttggag tcgatggagc 10920 aatgtatttc attgggtgtat ggggatttt taggagaata ttctttttt taggaaacac 10980 atagtaaaat tttagaccctt acaatttca ggtcttcaaa agactctcat gtgatttctg 11040 ttagggaaagg tggtaactttt catataacttg caacattttt gtgagtttaa cattgttctt 11100 ttctaaaaaa aaaattaaaaa ataaaattta ttcacatgt gctacatata tttgtaaatg 11160 ttaggttaatg gtgttatgccc attgttcttta ccactgttaag atcaagcaat ttacttcaga 11220 tacactaagt tgataccgtg ttccctcaat gcatgcagca attacagatc caccattatc 11280 aagagctcta ggtcttta ataccagag actaaatggg ctgcacccctt ttccctgttt 11340 gggcacccctc atagtctacc ttctttctg ccattaagta ttatttccca acattcatct 11400 ctcttagtga ggggtatcat cgcattggagc atgtccctgc cacgcaccat aggtgacact 11460 ttcttctttt acttttatac agggacatca ttctgaccca gacaccaaga ttccctgtttaa 11520 catctttgg aaagaggccac tcaatttt atcaagcaac tgcccatgtta catggagaaa 11580 tcaagttgtt tccagatgaa actggacagg gatttgcact cgttatatct catatctcta 11640 atgtgcccua aaatgtcccg gcctggttcc gtggggagggg aagtggatcc aactacatca 11700 gcatcagtgg gctgcagccctt aggactccat aaaattttac tgatcctga cttaggagagc 11760 caaattcacag tgctgttagcc tctgcacaaa ctttccctgtt gcttataag ctgcctgaat 11820 tttaaggggaa attgcttata ttggaaagaaa ggaagaaaagc tccatttgcctt ctctaaatgt 11880 ttgctgaaaaa taaaactgaca aaagggaaagat taataagaga aaaggcaaaac aaaatttact 11940 taaagtgcag caggatatac tagcaaggtt attaccaga taactcaatg ggatccagg 12000 gtttatgttt cctttctagg gaagaggttta ttggggaaatg taggcaaccc ggagagaata 12060 gatgcaataa gaaaatgcattt ctcaaaagaa caggtatata cctcccttagg taaagtatca 12120 acttttagtc tcttccattt ttgatttgcctt ttttgcgtt atcttccctg acataaaaat 12180 tctcagggaaag agtttttcttta aaaattttttt ttccctctgaa gaatttgcctt tttaggcaggt 12240 aggggatgtt taggaaaagc cccctgtca ttccctgtt tctaaatgccc tttgggtttt 12300 tataatcatc atacaaatgc agcacatgtt gagatgtttt tttctggatt cctttacttg 12360 caacccaccc gccaagatcc tggccagag agatgtggct acagactgaa agagcagtt 12420 tcccctcaac aacgtggagg ttggggactt gaccactgtt gcagatgaaa acctgtgtag 12480 aacttttgc tctccaaactt atgtatgtt agattactt tgacttggaaag ccttaatgtat 12540 aaaataaaata gttgatatac ccttttttattt gttatacata ttatataactc tatttcttcca 12600 ataaaagtatg ctaaaagaaaa aatgttattt agaaaaacctt aaaaatgaga aaatataattt 12660 agtacttattt aacgcgttgc tcacaggttga cacacagaaag aaaaatataag tggatctgca 12720 aatttcaaaac ccaagttattt caagggtttaa ctgttccatgt atgaatgttag cagtccttcat 12780 ctatagtcta gggcttttcc ctttgcgtt ttttgcgttca ttcccaatgg ctatataat 12840 gtctttagtgc cttcatgtatc ttggggagag aggttgcctt gcatgccttgc ttggccagat 12900 ggcctgcaat gtttatctt taggaaatgtt ctatggtttgc actgttgcctt ccagaatoca 12960 tcgttataaa gttaaatctc agtggaaatgg tatttgccttgc ttggggccat tgagatgtgt 13020 tagatcatga ggggtggagcc ctcttagtggaa atacccat gccaactataa acagggttta 13080 tggggcttggaa atcttcttgc ttttgcgtt ggttgcgttcat tttaagacat ggccttcgtt 13140 cctttgaagg actgaaagctt ccaggcgtca tcttggaaagc agagaaaagag gagcttaacc 13200 tgcccatgccc ttatgtttaa atttcccttattt ctccagaaat gtgagaaaat aaatttctgt 13260 tctttatgaa ttacacagttt tcaaggaaacc taacctgttta tagcagcttgc aaagagaaca 13320 agagagacag ctcacaatca gtgaggacag gatgagggtt atacataacctt cagtttcctc 13380 gtctctcagg tggaaatagcc cagagaattt taagtccatg ttccacatg tgggtgatct 13440 tcagttatcc tgagtccatgtt ggggttgc ttttgcgtt accattcatc ttctgttccc 13500 tccctcactt tcttcccttc ctcccaatgtt aaatttgcgtt cctaaacagg aatccttattt 13560 gctgggtggac cccaaactaag atagtaaaataa aaatcattaa tcttttgcgtt gagggttattt 13620 tctcatctga attcagatcg atgtgttacc tggatcttgc acctgcaggc atgcaagctt 13680 aggaagaaaa atctaaaaaa atgaaatgtca cagaaatagc aagttagaaatg ctgggttacca 13740 gagactgcattt tgntaaaggg tacaatataat cagttctqata qqqqactaa qttttqqtqa 13800
---	--

5 tctgttgcata gcatggcga cattattata tagatgtatt tctggcaggc ttgattgcct 13860  
 attacatggta aagtcaacat acactaagac actgggggct gctgcagaga aagagattta 13920  
 atcccaaggc aattaaatga gaagacagga ggaagcctca tatctacctc cccaaagcact 13980  
 ttggggttaa agacttaag tggttttgga tggagaggca gattgggtga agagtgaagg 14040  
 atgaagtcat gggactgggaa ggtgaagaaa ctgctttctt atgtgactc ggttcttgg 14100  
 gaggggggtg gtcttttagac aagttgggtg cagctattct actggaattc agatctgg 14160  
 10 aaatatctca aagatttaggt tttatgctca taatggtgaa ggtgttatat ttggaaacaa 14220  
 tggggacgtt aatggctgtt atttattgca acttgacttt tattagttag aagctaagg 14280  
 aagtggctca gaatgttgc tgattaatgc ttaactatat ttcagtcgg aacctgaaat 14340  
 actgttcttgc ttaaccttat gacagtgggtt ttgttagttaa taataactata ttctgttatt 14400  
 caaaataaca aagtatattc aatgttctca gtacaaaaaa tgataaaatat tggaggtgag 14460  
 15 gactatgttta attaacctgta tttgattatt cgacaatgtt tacatgtatg catctcattt 14520  
 taccatcatat gtatatacat tactattgtt caattaaaaat caaaataaaaa cttttaaaaat 14580  
 ttcatgttagc cattgcctgg aataaataaa gaaaaatatg tgtgcacatg cttactata 14640  
 cccctgtaaa aattgaagct actaagtccaa gtaaaataact gaatcacaca tatttctgtc 14700  
 20 ttttagtttta gtggaaatag tcaaaaataact gggtcccaac aataacttata tttagtcttt 14760  
 cttcccatat ccaccaccc ttttattttt aatttcatttta aatacattt agtttcattt 14820  
 actcttgcattt gtaaccacatt tggatccgtt gacctgcagg tcgaccggat ccagacatga 14880  
 taagatacat tggatgatttt ggacaaaacca caactagaat gcagtggaaaaaaatgctt 14940  
 tttgtgaaat ttgtgatgtt attgttttata ttgttaaccat tataagctgc aataaaacaaag 15000  
 25 ttaacaacaa caattgcattt cattttatgtt ttcaggttca gggggagggtg tgggagggtt 15060  
 tttaaagcaa gtaaaaaccc tacaatgtt gtagggctga ttatgatctc tagtcaaggc 15120  
 actatacatc aaatattccattt tattaaacccca tttacaaattt aaaaagctaa aggtacacaa 15180  
 ttttgagca tagtttattaa tagcagacac tctatgcctg tggatggatggaaaaaaacaaag 15240  
 tatgttatga ttataactgtt ttttttttttataacatgtt ttttttttttataacatgtt 15300  
 30 ttttttttttataacatgtt ttttttttttataacatgtt ttttttttttataacatgtt 15360  
 tctattacta aacacagcat gactaaaaaa acttagcaat tctgaaggaa aagcaagagt 15420  
 gtcttcatttccattt tttttggagg agtagaaatgt tgaggtcag cagtagcctc 15480  
 atcatcaacta gatggcattt ttttttttttataacatgtt ttttttttttataacatgtt 15540  
 ccactgtcc cattcatcattt ttttttttttataacatgtt ttttttttttataacatgtt 15600  
 35 cagtagtttta acacattata cactaaaaaa ttttttttttataacatgtt ttttttttttataacatgtt 15660  
 ttttttttttataacatgtt ttttttttttataacatgtt ttttttttttataacatgtt 15720  
 gggcccaactc ataaatccat ttttttttttataacatgtt ttttttttttataacatgtt 15780  
 cgtcggtacg ttcggcatcg cttcatcattttataacatgtt ttttttttttataacatgtt 15840  
 ggtcaatgcc ggcggcagatc cacatcagac gtttaatcat gcgataccag tgagggatgg 15900  
 40 ttttttttttataacatgtt aaggccgac tgcacaggcg gttgtgcggc gtgattaaag cggccggacta 15960  
 gcgtcgaggt ttcaggatgt ttttttttttataacatgtt ttttttttttataacatgtt 16020  
 gtgtcatgga tgcaggctcc agaataacttta ctggaaacta ttttttttttataacatgtt 16080  
 aaaagaaacaa cggccggcag ttttttttttataacatgtt ttttttttttataacatgtt 16140  
 acgaatacga cggccatatac ctttttttttataacatgtt ttttttttttataacatgtt 16200  
 45 acatagtcat caaccagcg acgaccagcc gtttttgcga agatgggtgac aaagtgcgt 16260  
 ttttttttttataacatgtt ttttttttttataacatgtt ttttttttttataacatgtt 16320  
 acgatgaagc cttcgccatc gtttttttttataacatgtt ttttttttttataacatgtt 16380  
 ttttttttttataacatgtt ttttttttttataacatgtt ttttttttttataacatgtt 16440  
 50 agtaacgcac cccgttaccat accgcacacgg ttttttttttataacatgtt ttttttttttataacatgtt 16500  
 gaaggcatca gtcggcttgc gtttttttttataacatgtt ttttttttttataacatgtt 16560  
 atgttattttt ttttttttttataacatgtt ttttttttttataacatgtt 16620  
 gagggggaaaaa taggttgcgc gagattatag agatctggcg cactaaaaac cagttttca 16680  
 catgagtcgg cgttttttttataacatgtt ttttttttttataacatgtt 16740  
 55 aaacatatact ctttttttttataacatgtt ttttttttttataacatgtt 16800  
 ctttttttttataacatgtt ttttttttttataacatgtt 16860  
 ctttttttttataacatgtt ttttttttttataacatgtt 16920  
 gggccggat gggccggat ttttttttttataacatgtt ttttttttttataacatgtt 16980  
 ttttttttttataacatgtt ttttttttttataacatgtt 17040  
 60 atacttctgc ctgttggggaa gtttttttttataacatgtt ttttttttttataacatgtt 17100  
 agtcgcctcg cggccgttgcg ttttttttttataacatgtt ttttttttttataacatgtt 17160  
 acggtcacag ctttttttttataacatgtt ttttttttttataacatgtt 17220

5	gcgggtgttg gcggtgtcg gggcgaccc atgacccagt cacgtagcga tagcggagtg 17280 tatactggct taactatgcg gcatcagagc agattgtact gagagtgcac catatgcggt 17340 gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgctct tccgcttcct 17400 cgctcaactga ctcgctgcgc tcggctgttc ggctgcggcg agcggtatca gtcactcaa 17460 aggcggtaat acggttatcc acagaatcag gggataaacgc agggaaaagaaac atgtgagcaa 17520 aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcggtt ttccataggc 17580 tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga 17640 caggactata aagataccag gcgttcccc ctggaagctc cctcgctgcgc ttcctgttc 17700 cgaccctgcc gettaccgga tacctgtccg ctttctccc ttcggaagc gtggcgctt 17760 ctctatagtc acgctgttagg tatctcagtt cggtaggtt cgttcgctcc aagctgggt 17820 gtgtcacgaa acccccccgtt cagccgacc gctgcgcctt atccgtaac tatcgcttt 17880 agtccaaaccc ggtaaagacac gacttatcgc cactggcagc agccactggt aacaggatta 17940 gcagagcgag gtatgttaggc ggtgctacag agttcttggaa gtggggcct aactacggct 18000 acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccggaaaa 18060 gagttggtag ctcttgatcc ggcaaaacaaa ccaccgctgg tagcggtggg tttttgttt 18120 gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatccttgc atcttttcta 18180 cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat 18240 caaaaaggat cttcacctag atcctttaa attaaaaatg aagttttaaa tcaatctaaa 18300 gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcgtgag gcacctatct 18360 cagcgatctg tctatttcgt tcattccatag ttgcctgact ccccgctgt tagataacta 18420 cgatacggga gggcttacca tctggcccca gtgctcaat gataccgcga gaccacgct 18480 caccggctcc agatttataca gcaataaaacc agccagccgg aaggggccgag cgcagaagt 18540 gtcctgcaac ttatccgccc tccatccagt ctattaaatg ttgcccggaa gctagagtaa 18600 gtagttcgcc agttaatagt ttgcgcacg ttgttgcctt tgctgcaggc atcgtgggt 18660 cacgctcgcc gtttggtagt gtttgcattca gtcgggttc ccaacgatca aggcgagta 18720 catgatcccc catgttgc aaaaaagcgg ttagctcctt cggcctccg atcgttgta 18780 gaagtaagtt gcccgcagtg ttatcaactca tggttatggc agcactgcat aattctctta 18840 ctgtcatgcc atccgtaaaga tgctttctg tgactggta gtactcaacc aagtctttct 18900 gagaatagt tatgcggcga ccgagggtctt cttggccggc gtcaacacgg gataataccg 18960 cgccacatag cagaacttta aaagtgccta tcattggaaa acgttctcg gggcgaaaaac 19020 tctcaaggat ctaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaaact 19080 gatcttcagc atctttact ttccacccgg ttctgggtt agcaaaaaaca ggaaggcaaa 19140 atgccgaaaa aaagggaaata agggcgacac ggaaatgttg aataactcata ctcttcctt 19200 ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atattgaat 19260 gtattttagaa aaataaacaa ataggggttc cgccgcacatt tccccggaaaa gtgccacctg 19320 acgtctaaaga aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc 19380 ccttcgtct tcaagaagaa ttccaagcta tt 19412
40	<210> 13 <211> 11 <212> DNA <213> Mus musculus
45	<400> 13 agtgtactgac g 11
50	<210> 14 <211> 9 <212> DNA <213> Mus musculus
55	<400> 14 ttccctgaa 9
60	<210> 15 <211> 15 <212> DNA <213> Mus musculus
65	<400> 15 gcaggaagtgc aagt 15

<212> DNA  
 <213> Mus musculus

5 <400> 16  
 gaaaaagt 7

<210> 17  
 <211> 9  
 <212> DNA  
 10 <213> Mus musculus

<400> 17  
 tttgggagg 9

15 <210> 18  
 <211> 9  
 <212> DNA  
 <213> Mus musculus

20 <400> 18  
 caaaaatggc 9

<210> 19  
 <211> 13  
 25 <212> DNA  
 <213> Mus musculus

<400> 19  
 agtttagccaa tgg 13

30 <210> 20  
 <211> 8  
 <212> DNA  
 <213> Mus musculus

35 <400> 20  
 tgacgttag 8

<210> 21  
 40 <211> 15  
 <212> DNA  
 <213> Mus musculus

<400> 21  
 45 aaaaaatgaaa gaact 15

<210> 22  
 <211> 15  
 <212> DNA  
 50 <213> Mus musculus

<400> 22  
 ggaagtgaaa gtaat 15

55 <210> 23  
 <211> 12  
 <212> DNA  
 <213> Mus musculus

60 <400> 23  
 gtgctaatga aa 12

<210> 24  
 <211> 10  
 65 <212> DNA  
 <213> Mus musculus

<400> 24  
 tttcattttt 10

5 <210> 25  
 <211> 6  
 <212> DNA  
 <213> Mus musculus

10 <400> 25  
 tcagca 6

<210> 26  
 <211> 6  
 <212> DNA  
 15 <213> Mus musculus

<400> 26  
 catgtg 6

20 <210> 27  
 <211> 9  
 <212> DNA  
 <213> Mus musculus

25 <400> 27  
 tgatgtaat 9

<210> 28  
 <211> 12  
 30 <212> DNA  
 <213> Mus musculus

<400> 28  
 ggtgcagcaa tg 12

35 <210> 29  
 <211> 6  
 <212> DNA  
 <213> Mus musculus

40 <400> 29  
 cttctc 6

<210> 30  
 45 <211> 8  
 <212> DNA  
 <213> Mus musculus .

<400> 30  
 50 atttgcat 8

<210> 31  
 <211> 6  
 <212> DNA  
 55 <213> Mus musculus

<400> 31  
 ttcaaa 6

60 <210> 32  
 <211> 10  
 <212> DNA  
 <213> Mus musculus

65 <400> 32  
 tggccttcc 10

5                   <210> 33  
 <211> 8  
 <212> DNA  
 <213> Mus musculus  
 10                  <210> 33  
 <400> 33  
 atttgcat 8  
 15                  <210> 34  
 <211> 10  
 <212> DNA  
 <213> Mus musculus  
 20                  <400> 34  
 tatgcaaatg 10  
 <210> 35  
 <211> 15  
 <212> DNA  
 <213> Mus musculus  
 25                  <210> 35  
 <211> 7  
 <212> DNA  
 <213> Mus musculus  
 30                  <400> 35  
 ccgaatatgc aattc 15  
 35                  <210> 36  
 <211> 7  
 <212> DNA  
 <213> Mus musculus  
 40                  <400> 36  
 aatttcc 7  
 45                  <210> 37  
 <211> 12  
 <212> DNA  
 <213> Mus musculus  
 50                  <400> 37  
 gctttccta aa 12  
 55                  <210> 38  
 <211> 8  
 <212> DNA  
 <213> Mus musculus  
 60                  <400> 38  
 cttccct 8  
 65                  <210> 39  
 <211> 12  
 <212> DNA  
 <213> Mus musculus  
 70                  <400> 39  
 cattgctcca tc 12  
 75                  <210> 40  
 <211> 15  
 <212> DNA  
 <213> Mus, musculus  
 80                  <400> 40  
 ctcatttgca tgttc 15  
 85                  <210> 41  
 <211> 10  
 <212> DNA

<213> Mus musculus  
 <400> 41  
 ggaacaatgt 10  
 5  
 <210> 42  
 <211> 7  
 <212> DNA  
 <213> Mus musculus  
 10  
 <400> 42  
 ttaagt 7  
 <210> 43  
 <211> 10  
 15 <212> DNA  
 <213> Mus musculus  
 <400> 43  
 gggctttcc 10  
 20  
 <210> 44  
 <211> 11  
 <212> DNA  
 <213> Mus musculus  
 25  
 <400> 44  
 gacaggtggg g 11  
 <210> 45  
 30 <211> 11  
 <212> DNA  
 <213> Mus musculus  
 <400> 45  
 35 ggcaggtggg t 11  
 <210> 46  
 <211> 9  
 <212> DNA  
 40 <213> Mus musculus  
 <400> 46  
 ttccataaa 9  
 45 <210> 47  
 <211> 12  
 <212> DNA  
 <213> Mus musculus  
 50 <400> 47  
 tatcattaag gc 12  
 <210> 48  
 55 <211> 13  
 <212> DNA  
 <213> Mus musculus  
 <400> 48  
 60 catacaccaa tga 13  
 <210> 49  
 <211> 6  
 <212> DNA  
 <213> Mus musculus  
 65 <400> 49

5 catgtg 6  
 <210> 50  
 <211> 6  
 <212> DNA  
 <213> Mus musculus  
 10 <400> 50  
 catgtg 6  
 <210> 51  
 <211> 7  
 <212> DNA  
 <213> Mus musculus  
 15 <400> 51  
 aggtaat 7  
 <210> 52  
 20 <211> 14  
 <212> DNA  
 <213> Mus musculus  
 <400> 52  
 25 aaatgttgca agta 14  
 <210> 53  
 <211> 6  
 <212> DNA  
 <213> Mus musculus  
 <400> 53  
 30 acatgg 6  
 <210> 54  
 <211> 7  
 <212> DNA  
 <213> Mus musculus  
 <400> 54  
 35 caggaag 7  
 <210> 55  
 <211> 6  
 <212> DNA  
 <213> Mus musculus  
 <400> 55  
 40 ttcaaa 6  
 <210> 56  
 <211> 8  
 <212> DNA  
 <213> Mus musculus  
 <400> 56  
 45 atttgcat 8

**REIVINDICACIONES**

1. Un vector de expresión que comprende la secuencia de ácido nucleico de una cualquiera de SEQ ID NOs: 11-12.
- 5 2. Una célula huésped que comprende los vectores de expresión de acuerdo con la reivindicación 1.
3. La célula huésped de la reivindicación 2, donde la célula huésped es una célula huésped de mamífero.
- 10 4. La célula huésped de la reivindicación 3 donde la célula huésped es una célula huésped de mieloma murino.
5. Un kit que comprende el vector de la reivindicación 1.

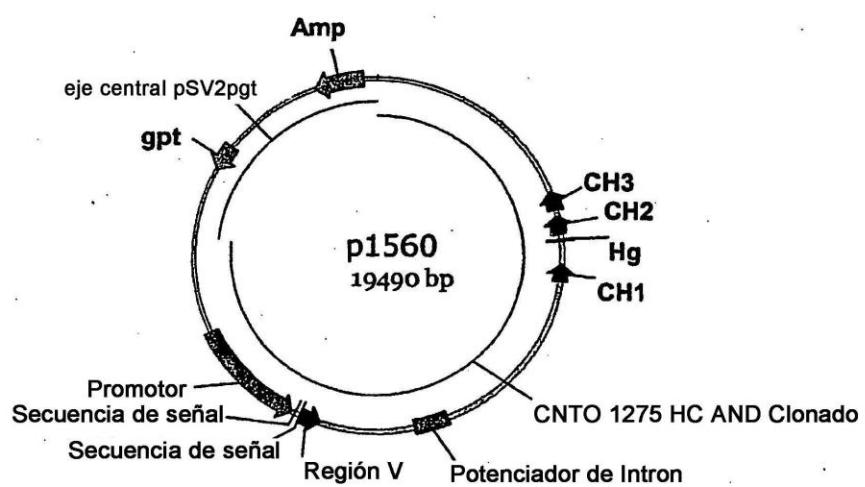


FIG. 1A

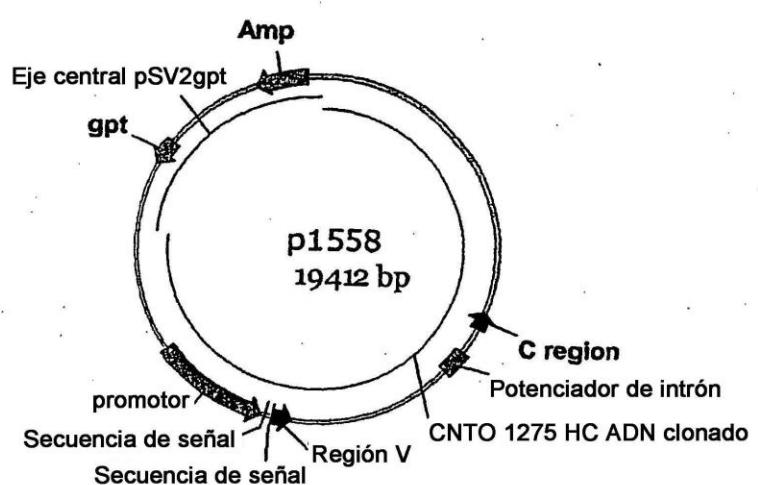


FIG. 1B