

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 678 221**

51 Int. Cl.:

**A01K 67/027** (2006.01)

**C07K 16/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.02.2014 PCT/US2014/017427**

87 Fecha y número de publicación internacional: **28.08.2014 WO14130690**

96 Fecha de presentación y número de la solicitud europea: **20.02.2014 E 14754019 (9)**

97 Fecha y número de publicación de la concesión europea: **18.04.2018 EP 2840892**

54 Título: **Animales no humanos con secuencias de cadena pesada de inmunoglobulina modificadas**

30 Prioridad:

**20.02.2013 US 201361766765 P**  
**18.09.2013 US 201361879338 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**09.08.2018**

73 Titular/es:

**REGENERON PHARMACEUTICALS, INC.**  
**(100.0%)**  
**777 Old Saw Mill River Road**  
**Tarrytown, NY 10591, US**

72 Inventor/es:

**MCWHIRTER, JOHN;**  
**GURER, CAGAN;**  
**MEAGHER, KAROLINA A.;**  
**MACDONALD, LYNN y**  
**MURPHY, ANDREW J.**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

Observaciones:

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

**ES 2 678 221 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

Animales no humanos con secuencias de cadena pesada de inmunoglobulina modificadas

## 5 Referencia cruzada con solicitudes relacionadas

## Listado de secuencias

## Campo de la invención

10

Se describen ratones y ratas modificados genéticamente, que comprenden una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada (es decir, una secuencia VDJ de cadena pesada reordenada) unida operativamente a una secuencia de ácido nucleico de región constante endógena. En algunos aspectos, los ratones y las ratas se modifican genéticamente para que tengan un locus de inmunoglobulina que comprenda una

15 secuencia de ácido nucleico (una secuencia VDJ) de región variable de cadena pesada reordenada unida operativamente a una secuencia génica de región constante de inmunoglobulina, en la que la secuencia VDJ es una secuencia VDJ humana, y la secuencia génica de región constante es endógena. En algunos aspectos, los ratones o las ratas comprenden un locus de cadena pesada de inmunoglobulina modificado genéticamente que comprende una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada, en

20 la que el dominio variable de la cadena pesada reordenada comprende una secuencia del segmento del gen V de la cadena pesada ( $V_H$ ) unida operativamente, mediante un espaciador, a una secuencia del segmento del gen J de la cadena pesada ( $J_H$ ), y en la que el espaciador comprende al menos un resto de aminoácido. Se describen ratones y ratas modificados genéticamente, que comprenden en sus genomas: (i) una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada, unida operativamente a una secuencia de ácido nucleico de región

25 constante; y (ii) un locus de cadena ligera de inmunoglobulina que comprende uno o más segmentos génicos de región variable de cadena ligera humana, pero menos que el número de segmentos de tipo silvestre. Se describen ratones y ratas modificados genéticamente que comprenden en sus genomas: (i) una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante; y (ii) un locus de cadena ligera de inmunoglobulina que comprende uno o más segmentos

30 génicos de región variable de cadena ligera humana, pero menos que el número de segmentos de tipo silvestre. En algunos aspectos, al menos uno de los segmentos génicos de región variable codifica uno o más restos de histidina que no están codificados por un segmento génico de región variable de cadena ligera de la línea germinal humana correspondiente. En el presente documento se describen métodos para crear ratones o ratas modificados genéticamente. Se describen métodos para la producción de secuencias de región variable de cadena ligera (por

35 ejemplo, una cadena ligera  $\kappa$  o  $\lambda$ ) de inmunoglobulina que pueden unirse a un antígeno en ausencia de una cadena pesada, y/o que pueden asociarse con un dominio variable de cadena pesada reordenada y/o que presentan características de unión a antígeno dependiente de pH, que son útiles para la producción de anticuerpos biespecíficos.

## 40 Antecedentes

Los anticuerpos biespecíficos son anticuerpos multifuncionales que comprenden sitios de unión a antígeno que pueden unirse a dos determinantes antigénicos distintos y han aparecido como uno de los principales productos

45 bioterapéuticos para el tratamiento de muchas enfermedades, incluyendo el cáncer. Aunque recientemente se han desarrollado diversos anticuerpos biespecíficos con dobles propiedades de unión a antígeno, en los anticuerpos biespecíficos convencionales, la especificidad y afinidad de los dominios variables de cadena ligera o de cadena pesada de inmunoglobulina, han tenido que sacrificarse hasta cierto grado porque, en los anticuerpos biespecíficos convencionales, cualquiera de solo un dominio variable de cadena ligera o de cadena pesada contribuye a la unión

50 de cada uno de los determinantes antigénicos, mientras que en los anticuerpos regulares las regiones variables tanto de cadena pesada como de cadena ligera pueden contribuir a la unión con el mismo determinante antigénico. Además, a la hora de conseguir un nivel de eficacia deseable, los anticuerpos terapéuticos, por ejemplo, anticuerpos terapéuticos biespecíficos, a menudo requieren dosis altas o múltiples de anticuerpos debido a su reciclabilidad limitada *in vivo*.

55 La mayoría de las proteínas de unión a antígeno desarrolladas hasta ahora que se dirigen a dos antígenos o epítomos, comprenden dos brazos de unión a antígeno: (i) un primer brazo de unión a antígeno que comprende un par de dominios variables de cadena ligera y pesada de inmunoglobulina que contribuye a la unión con un primer antígeno o epítomo; y (ii) un segundo brazo de unión a antígeno que comprende un segundo par de dominios variables de cadena ligera y pesada que contribuye a la unión con un segundo antígeno o epítomo. Aunque estas

60 proteínas de unión a antígeno son biespecíficas en todo el contexto de proteínas de unión a antígeno, no son necesariamente biespecíficas dentro de cada brazo de unión a antígeno, limitando el uso de las proteínas de unión a antígeno en formatos multiespecíficos, por ejemplo, proteínas de unión a antígeno triespecíficas. Como se desvela en el presente documento, un ratón o una rata que expresa un dominio variable de cadena pesada universal puede emplearse como una herramienta general para crear proteínas de unión a antígeno para su uso en muchos formatos

65 diferentes de proteínas de unión a antígeno.

Se hace referencia a los siguientes documentos:

- WO2013/022782 que tiene que ver con ratones de cadena ligera universal humanizada; y
- Osborn *et al.* (2013) *Journal of Immunology*, 190 (4): 1481-1490 que se refiere a anticuerpos de IgG de alta afinidad que se desarrollan de manera natural en ratas con Ig desactivada (*knockout*), que llevan locus de Ig IgH/Ig humana de la línea germinal, portadores de la región C<sub>H</sub> de rata.

### Resumen

En la técnica existe una necesidad de generar secuencias de dominio variable de cadena ligera de inmunoglobulina en las que la afinidad y especificidad antigénica es el resultado, exclusiva o principalmente, de la diversidad de dominios variables de cadena ligera de inmunoglobulina, y/o reside exclusiva o principalmente en esta diversidad. Dichas secuencias serían extremadamente útiles a la hora de diseñar proteínas de unión a antígeno, por ejemplo, anticuerpos biespecíficos, en los que cada dominio variable es independientemente responsable de la distinta unión específica de antígeno. Diversos aspectos descritos en el presente documento se basan en parte en el sorprendente descubrimiento de que las ratas y los ratones modificados genéticamente que comprenden dominios variables de cadena pesada de inmunoglobulina codificados por una secuencia génica variable de cadena pesada reordenada (por ejemplo, una secuencia VDJ de cadena pesada reordenada) pueden reunir esta necesidad. Las ratas y los ratones que codifican un dominio variable de cadena pesada de inmunoglobulina reordenada (es decir, un dominio variable de cadena pesada universal) enfocan los mecanismos de diversificación de anticuerpos en uno o más dominios variables de cadena ligera de anticuerpos no reordenada (es decir diversificable).

Se describen ratones y las ratas modificados genéticamente que, después de la estimulación con un antígeno de interés, producen anticuerpos con especificidad de unión a antígeno que reside exclusiva o principalmente en los dominios variables de cadena ligera de anticuerpo. Los aminoácidos de los dominios variables de anticuerpo de la cadena ligera y las correspondientes secuencias de ácido nucleico, pueden identificarse de anticuerpos producidos por dichas ratas y ratones modificados genéticamente, y las secuencias pueden utilizarse en anticuerpos recombinantes u otras proteínas de unión a antígeno para desarrollar dominios variables de cadena ligera que se unen a un determinante antigénico independientemente (y con suficiente especificidad y afinidad) de los dominios variables de cadena pesada. Las secuencias génicas de dominio variable de cadena pesada reordenada se dirigen a un locus de cadena pesada de tal manera que las secuencias de dominio variable de cadena pesada reordenada pueden unirse operativamente a una secuencia constante pesada endógena. Asimismo, las ratas y los ratones que comprenden una secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada pueden combinarse con modificaciones genéticas adicionales de locus de inmunoglobulina (por ejemplo, cruzarse con ratas y ratones que comprenden modificaciones genéticas adicionales de locus de inmunoglobulina). Además, utilizando ratones que tienen un repertorio restringido (limitado) de segmentos génicos de la región variable de cadena ligera (por ejemplo, un número restringido de secuencias génicas variables de cadena ligera que comprenden uno o más segmentos génicos V<sub>L</sub> humanos, pero menos que el número de segmentos de tipo silvestre, en combinación con la única secuencia de cadena pesada reordenada descrita anteriormente), puede producirse un dominio variable de cadena ligera de inmunoglobulina que puede emparejarse de un modo más eficaz con un dominio variable de cadena pesada de inmunoglobulina.

Por tanto, la presente invención proporciona una rata o un ratón que comprende en su genoma de línea germinal, en un locus de cadena pesada de inmunoglobulina endógena, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada bajo el control de un promotor y unida operativamente a una secuencia génica de la región constante de la cadena pesada endógena, en la que la secuencia de nucleótidos de región variable de cadena pesada reordenada codifica la secuencia de V<sub>H</sub>3-23/X<sub>1</sub>X<sub>2</sub>/J<sub>H</sub>, en la que X<sub>1</sub> y X<sub>2</sub> es cualquier aminoácido, y en la que todos los segmentos génicos V<sub>H</sub>, D<sub>H</sub> y J<sub>H</sub> funcionales endógenos están deletados de la cadena pesada de inmunoglobulina endógena o se han hecho no funcionales.

La presente invención proporciona además un método para crear una rata o un ratón, comprendiendo el método:

(a) modificar un genoma de una rata o un ratón para deletar o hacer no funcionales los segmentos génicos V<sub>H</sub>, D<sub>H</sub> y J<sub>H</sub> de cadena pesada de inmunoglobulina funcional endógena; y

(b) colocar en el genoma, en el locus de la cadena pesada de inmunoglobulina endógena, una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada, en la que la secuencia de nucleótidos de la región variable de la cadena pesada reordenada codifica la secuencia de V<sub>H</sub>3-23/X<sub>1</sub>X<sub>2</sub>/J<sub>H</sub> unida operativamente a una secuencia de la región constante de la cadena pesada endógena, en la que X<sub>1</sub> y X<sub>2</sub> es cualquier aminoácido.

La presente invención proporciona adicionalmente un método para obtener una secuencia de ácido nucleico que codifica un dominio variable de la cadena ligera (VL) de inmunoglobulina, capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

(a) inmunizar a una rata o a un ratón de la invención, con un antígeno de interés o con un inmunógeno del mismo;

(b) permitir que la rata o el ratón genere una respuesta inmunitaria;

(c) aislar, de la rata o ratón inmunizado, una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

(d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio VL) que pueda unirse al antígeno.

La presente invención también proporciona un método para crear una proteína de unión a antígeno que comprende un dominio variable de cadena ligera de inmunoglobulina que puede unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

(a) inmunizar a una rata o a un ratón modificado genéticamente de la invención, con un primer antígeno que comprende un primer epítipo o una parte inmunógena del mismo, en el que la rata o el ratón comprende en su genoma dos o más segmentos génicos de la región variable de la cadena ligera (VL y JL) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de la región constante de la cadena ligera de inmunoglobulina;

(b) permitir que la rata o el ratón genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

(c) aislar, de la rata o del ratón, una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de la cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(d) obtener de la célula de (c), la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une específicamente al primer epítipo o parte inmunógena del mismo;

(e) emplear la secuencia de ácido nucleico de (d), en una construcción de expresión, fusionada a una secuencia de ácido nucleico de la región constante de inmunoglobulina humana; y

(f) expresar la secuencia de ácido nucleico de (d), en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo para formar una proteína de unión a antígeno cuya cadena ligera esté codificada por el ácido nucleico de (d) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

La invención también proporciona una célula procedente de una rata o ratón de la invención, en la que la célula comprende en su genoma un locus de cadena pesada de inmunoglobulina que comprende una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada, en el que la secuencia de nucleótidos de la región variable de la cadena pesada reordenada codifica la secuencia de  $V_H3-23/X_1X_2/J_H$ , en la que  $X_1$  y  $X_2$  es cualquier aminoácido y en la que todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  funcionales endógenos están delecionados de la cadena pesada de inmunoglobulina endógena o la hacen no funcional.

En diversos aspectos, la única secuencia genómica de ácido nucleico que codifica el dominio variable de cadena pesada expresada por los ratones o las ratas modificados genéticamente, es el dominio variable de cadena pesada reordenada. Por consiguiente, la diversidad de los dominios variables de cadena pesada del anticuerpo producida por los ratones o ratas modificados genéticamente es extremadamente restringida.

En algunos aspectos, se describen ratas y ratones modificados genéticamente que tienen, en su genoma, un locus de inmunoglobulina que se ha modificado genéticamente de tal manera que sus secuencias de región variable consisten esencialmente en una región variable de cadena pesada humana reordenada sencilla. Se entiende que, en dichos ratones y ratas modificados genéticamente, puede que las diferentes células no siempre tengan secuencias completamente idénticas en la región variable de cadena pesada reordenada sencilla (por ejemplo debido a errores de replicación, hipermutación somática u otros mecanismos), pero independientemente, dichos ratones y ratas modificados genéticamente muestran una diversidad drásticamente restringida de los dominios variables de cadena pesada de anticuerpo en comparación con ratas y ratones que tienen secuencias variables de cadena pesada no reordenadas y/o con ratas y ratones cuyos genomas incluyen múltiples segmentos génicos de la región variable de la cadena pesada (por ejemplo, múltiples segmentos V, D y/o J, particularmente si no están reordenados).

En diversos aspectos, se describe un locus de cadena pesada de inmunoglobulina modificado genéticamente que comprende una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada (es decir, que comprende una secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada). En diversos aspectos, la secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada y el dominio variable de la cadena pesada reordenada que lo codifica, procede de un segmento génico V, D y J humano. En diversos aspectos, la secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada y el dominio variable de la cadena pesada reordenada que lo codifica, proceden de un segmento génico J<sub>H</sub> humano y V<sub>H</sub> humano. En diversos aspectos, la secuencia de nucleótidos de la región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de la región constante de la cadena pesada. En diversos aspectos, la secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de la región constante de la cadena ligera. En diversos aspectos, el locus de inmunoglobulina modificado genéticamente, está presente en la línea germinal de una rata o ratón. En diversos aspectos, las ratas o los ratones modificados genéticamente, comprenden todo el complemento de los segmentos génicos variables de la cadena ligera no reordenada, capaces de reordenarse para formar un gen de cadena ligera en unión operativa con una secuencia génica de la región constante de cadena ligera. En otros aspectos, las ratas y los ratones modificados genéticamente, comprenden una pluralidad de segmentos génicos variables de cadena ligera no reordenada, pero menos que un complemento completo (es decir, menos que un número de segmentos de tipo silvestre). En diversos aspectos, los segmentos génicos variables de cadena ligera no reordenada, están unidos operativamente a una secuencia génica de la región constante de cadena pesada. En otro aspecto, se describe una construcción de ácido nucleico que comprende una región variable de la cadena pesada de inmunoglobulina humana reordenada (es decir, que comprende una secuencia de nucleótidos que codifica un dominio variable de la cadena pesada reordenada; es decir, una secuencia VDJ de cadena pesada previamente reordenada) como se describe en el presente documento.

En el presente documento se describen numerosas variaciones de ratas y ratones modificados genéticamente con un locus de inmunoglobulina que comprende una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada (es decir, con un locus de inmunoglobulina que comprende una secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada). Cada variación tiene la capacidad de enfocar los mecanismos de diversificación de anticuerpos sobre secuencias de nucleótidos de la región variable de la cadena ligera de la inmunoglobulina.

En diversos aspectos, una secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada (es decir, un dominio variable de cadena pesada codificado por una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada) está unida operativamente a una secuencia génica de la región constante de la cadena pesada endógena (por ejemplo, una secuencia génica de la región constante de cadena pesada que codifica un isotipo de inmunoglobulina seleccionado de IgM, IgD, IgA, IgE, IgG y sus combinaciones). Por ejemplo, se describen ratones y ratas no humanos modificados genéticamente, que comprenden locus de inmunoglobulina en los que: (a) una primera secuencia de nucleótidos codifica un dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada), en el que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de la región constante de cadena pesada endógena; y (b) una segunda secuencia de nucleótidos codifica un dominio variable de cadena ligera (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de la cadena ligera de inmunoglobulina humana no reordenada), en el que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de la región constante de la cadena ligera humana o no humana.

En otro aspecto, se describen ratas y ratones no humanos modificados, en los que las ratas y los ratones comprenden una secuencia de nucleótidos reordenada que codifica un dominio variable de cadena pesada, en el que el dominio variable de cadena pesada comprende una secuencia variable de cadena pesada (V<sub>H</sub>) que está unido operativamente, mediante un espaciador, a una secuencia de segmento J de la cadena pesada (J<sub>H</sub>), en la que el espaciador codifica al menos un resto de aminoácido.

En otro aspecto, se describe una rata o un ratón que comprende un locus de inmunoglobulina modificado genéticamente que comprende una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada (es decir, comprende una secuencia de ácido nucleico que codifica un dominio variable de cadena pesada reordenada; es decir, una secuencia VDJ de cadena pesada reordenada), en el que el locus de inmunoglobulina modificado genéticamente está presente en la línea germinal de la rata o ratón. En algunos aspectos, el locus de inmunoglobulina modificado genéticamente es un locus de cadena pesada. En algunos aspectos, el locus de inmunoglobulina modificado genéticamente es un locus de cadena ligera.

Se describen ratones y ratas modificados genéticamente que tienen un locus genómico de inmunoglobulina modificado genéticamente que comprende una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada, en el que la secuencia de nucleótidos está unida operativamente a una secuencia génica de la región constante de la cadena ligera humana o no humana (por ejemplo, cadena ligera κ o λ).

Se describe una rata o un ratón que comprende un locus de inmunoglobulina modificado genéticamente que comprende: (a) una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia génica de la región constante de la cadena ligera; y (b) una secuencia de nucleótidos de la región variable de la cadena ligera humana o no humana (por ejemplo, cadena ligera  $\kappa$  o  $\lambda$ ) reordenada, unida operativamente a una secuencia génica de la región constante de la cadena pesada humana o no humana (por ejemplo, una secuencia génica de la región constante de la cadena pesada que codifica un isotipo de inmunoglobulina seleccionado de IgM, IgD, IgA, IgE, IgG y sus combinaciones.

En otro aspecto, se describe una rata o un ratón modificado genéticamente con un locus de inmunoglobulina que comprende:

(a) un primer alelo que comprende:

(i) una primera secuencia de nucleótidos que codifica un dominio variable de la cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada) unida operativamente a una secuencia génica de la región constante de cadena pesada, y

(ii) una segunda secuencia de nucleótidos que codifica un dominio variable de la cadena ligera (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de la cadena ligera de inmunoglobulina humana no reordenada) unida operativamente a una secuencia génica de la región constante de la cadena ligera; y

(b) un segundo alelo que comprende

(i) una tercera secuencia de nucleótidos que codifica un dominio variable de la cadena ligera (es decir, en el que la tercera secuencia de nucleótidos es una secuencia de nucleótidos variable de la cadena ligera de inmunoglobulina humana no reordenada) unida operativamente a una secuencia génica de la región constante de la cadena pesada, y

(ii) una cuarta secuencia de nucleótidos que codifica el dominio variable de la cadena pesada reordenada (es decir, en el que la cuarta secuencia de nucleótidos es una secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada) unida operativamente a una secuencia génica de la región constante de la cadena ligera.

En diversos aspectos, se describen ratas y ratones modificados genéticamente con locus o secuencias génicas de la región variable de la cadena ligera no reordenadas. En algunos aspectos, los ratones y ratas modificados genéticamente comprenden un número (es decir todos o sustancialmente todos) de segmentos génicos (es decir, secuencias) de la región variable de la cadena ligera de inmunoglobulina humana de tipo silvestre. En otros aspectos, las ratas y los ratones descritos en el presente documento, comprenden un repertorio limitado de segmentos génicos variables de cadena ligera, por ejemplo, (i) uno, dos o más segmentos génicos  $V_L$  humanos, pero menos que el número de segmentos de tipo silvestre; y (ii) uno o más segmentos génicos  $J_L$  humanos, unidos operativamente a una secuencia de ácido nucleico de la región constante de la cadena ligera no humana. La secuencia de ácido nucleico de la cadena pesada está presente en un locus de la cadena pesada de inmunoglobulina endógena y/o puede haber segmentos de la cadena ligera, por ejemplo, en un transgén o en un locus de inmunoglobulina endógeno.

En diversos aspectos, se describen ratones y ratas modificados genéticamente, en los que todos los dominios variables de la cadena pesada de inmunoglobulina del animal proceden de la misma secuencia génica de la cadena pesada variable reordenada, y en los que dichos dominios variables se expresan relacionados con un dominio variable de cadena ligera procedente de uno de al menos uno, dos o tres o más segmentos génicos  $V_L$  y al menos uno, dos o tres o más segmentos génicos  $J_L$ . Adicionalmente se describen ratones y ratas modificados genéticamente que comprenden en sus genomas: (i) un locus de cadena pesada de inmunoglobulina que comprende una secuencia de ácido nucleico de la región variable de la cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de la región constante de la cadena pesada humana o no humana y (ii) un locus de cadena ligera de inmunoglobulina que comprende uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre (por ejemplo, dos segmentos génicos  $V_k$  humanos y uno o más segmentos génicos  $J_k$  humanos), unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera no humana o humana. En algunos aspectos, la región constante de cadena ligera es una región constante de rata o de ratón, por ejemplo, una región constante  $C_k$  de rata o de ratón.

En otro aspecto, los ratones y las ratas modificados genéticamente del presente documento, expresan, después de la estimulación con un antígeno de interés, una proteína de unión a antígeno que comprende una secuencia de aminoácidos de cadena ligera y de cadena pesada de inmunoglobulina, en el que la secuencia de aminoácidos de cadena pesada procede de un locus de cadena pesada modificada genéticamente que comprende una secuencia de

ácido nucleico de región variable de cadena pesada humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada. En determinados aspectos, la secuencia de aminoácidos de cadena ligera procede de un locus de cadena ligera de inmunoglobulina modificada genéticamente que comprende uno o más segmentos génicos  $V_L$  humanos, pero menos que el número de segmentos de tipo silvestre, y (ii) dos o más segmentos génicos  $J_L$  humanos, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera no humana.

Se describen ratones y ratas modificados genéticamente que comprenden en sus genomas una secuencia de ácido nucleico de la región variable de la cadena pesada de inmunoglobulina humana reordenada que comprende un segmento génico  $V$  de la cadena pesada ( $V_H$ ) que está unido operativamente, mediante un espaciador, a una secuencia del segmento génico  $J$  de la cadena pesada ( $J_H$ ), en el que el espaciador codifica al menos un aminoácido (por ejemplo, 2 aminoácidos, 3 aminoácidos o 4 aminoácidos) y/o un segmento génico  $D$  ( $D_H$ ) modificado. En diversos aspectos, la secuencia de ácido nucleico de la región variable de la cadena pesada reordenada está unida operativamente a una secuencia de ácido nucleico de la región constante de la cadena pesada humana o no humana. En diversos aspectos, los ratones y las ratas comprenden adicionalmente, en sus genomas, un locus de cadena ligera de inmunoglobulina modificada genéticamente que comprende uno o más segmentos génicos de la región variable de la cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, por ejemplo, dos segmentos génicos  $V_k$  humanos y uno o más segmentos génicos  $J_k$  humanos, unidos operativamente a una secuencia de ácido nucleico de la región constante de la cadena ligera humana o no humana.

También se proporcionan métodos para crear y utilizar las ratas y los ratones modificados genéticamente descritos en el presente documento. Se describen métodos para colocar en el genoma de una rata o un ratón no humano, una secuencia de ácido nucleico de la región variable de la cadena pesada humana reordenada en unión operativa con una secuencia de ácido nucleico de la región constante de la cadena ligera o pesada de inmunoglobulina.

En otro aspecto, se describen métodos para obtener una secuencia de aminoácidos de la región variable de la cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de una secuencia de aminoácidos de la región variable de la cadena pesada.

En otro aspecto, mediante cualquiera de los métodos descritos en el presente documento, puede obtenerse un locus de inmunoglobulina modificado genéticamente

En diversos aspectos, en el presente documento se describen proteínas de unión a antígeno (por ejemplo, anticuerpos) producidas por, o procedentes de, los ratones y ratas no humanos modificados genéticamente. También se describen métodos para crear proteínas de unión a antígeno, incluyendo proteínas multiespecíficas (por ejemplo, biespecíficas o trispecíficas) de unión a antígeno. También se describen métodos para crear dominios efectores variables de cadena ligera de inmunoglobulina.

En otro aspecto, se describe una célula pluripotente, células madre pluripotentes o totipotentes inducidas, procedentes de una rata o un ratón, que comprenden las diversas modificaciones genómicas descritas en el presente documento. También se describen células que comprenden un núcleo que contiene una modificación genética como se describe en el presente documento, por ejemplo, una modificación introducida en una célula mediante inyección pronuclear.

En diversos aspectos, se describe un embrión de ratón o de rata que comprende una célula cuyo genoma comprende un locus de cadena pesada de inmunoglobulina que comprende una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante. En determinados aspectos, el embrión de rata o de ratón comprende adicionalmente un locus de cadena ligera de inmunoglobulina que comprende dos o más segmentos génicos de región variable de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera.

También se describen métodos para crear secuencias de ácido nucleico que codifican una secuencia de aminoácidos de región variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaces de unirse a un antígeno o a un epítipo del mismo, independientemente de un dominio variable de cadena pesada, que comprenden: (a) inmunizar a un ratón o a una rata con un antígeno de interés o con un inmunógeno del mismo, en el que el ratón o la rata comprende en su genoma (i) una secuencia de ácido nucleico de la región variable de la cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de la región constante de la cadena pesada, y (ii) una secuencia de ácido nucleico de la región variable de cadena ligera de inmunoglobulina humana no reordenada unida operativamente a una secuencia de ácido nucleico de la región constante de cadena ligera; (b) permitir que la rata o el ratón genere una respuesta inmunitaria; (c) aislar de la rata o del ratón inmunizado una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y (d) obtener de la célula, una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno. En diversos aspectos, la célula es un linfocito, incluyendo, pero sin limitación, linfocitos citolíticos naturales, linfocitos T y linfocitos B. En

diversos aspectos, el método comprende adicionalmente (c)' fusionar el linfocito con células cancerosas, por ejemplo una célula de mieloma.

5 También se describen métodos para crear secuencias de ácido nucleico que codifican una secuencia de aminoácidos de región variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaces de unirse a un antígeno o a un epítipo del mismo independientemente de un dominio variable de cadena pesada, que comprende: (a) inmunizar a una rata o a un ratón con un antígeno de interés o con un inmunógeno del mismo, en el que la rata o el ratón comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de la región constante de cadena pesada, y (ii) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera; (c) aislar de la rata o del ratón inmunizado una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y (d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ) que puede unirse al antígeno. En diversos aspectos, la célula es un linfocito, incluyendo, pero sin limitación, linfocitos citolíticos naturales, linfocitos T y linfocitos B. En diversos aspectos, el método comprende adicionalmente (c)' fusionar el linfocito con una célula cancerosa, por ejemplo, una célula de mieloma.

20 También se describen métodos para crear secuencias de ácido nucleico que codifican una secuencia de aminoácidos de región variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaces de unirse a un antígeno o a un epítipo del mismo independientemente de un dominio variable de cadena pesada, que comprenden: (a) inmunizar a una rata o a un ratón con un antígeno de interés o con un inmunógeno del mismo, en el que la rata o el ratón comprende en su genoma: (i) una secuencia de ácido nucleico de la región variable de la cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia de ácido nucleico de la región constante de cadena ligera, y (ii) segmentos génicos de la región variable de la cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana unidos operativamente a una secuencia de ácido nucleico de la región constante de la cadena pesada; (c) aislar del ratón o rata inmunizado una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y (d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de la cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno. En diversos aspectos, la célula es un linfocito, incluyendo pero sin limitación, linfocitos citolíticos naturales, células T y células B. En diversos aspectos, el método comprende adicionalmente (c)' fusionar el linfocito con una célula cancerosa, por ejemplo, una célula de mieloma.

35 También se describen métodos para crear proteínas de unión a antígeno, que comprenden:

(a) inmunizar a una rata o a un ratón con un primer antígeno que comprende un primer epítipo o parte inmunógena del mismo, en el que la rata o el ratón comprende en su genoma: (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada, y (ii) una secuencia de ácido nucleico de región variable de cadena ligera de inmunoglobulina humana no reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera;

45 (b) permitir que la rata o el ratón genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

(c) aislar de la rata o del ratón una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

50 (d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(e) emplear la secuencia de ácido nucleico de (c) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

55 (f) expresar el ácido nucleico de (c) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina que se una específicamente a un segundo antígeno o epítipo del mismo para formar una proteína de unión a antígeno cuya cadena ligera esté codificada por el ácido nucleico de (c) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

60 También se describen métodos para crear proteínas de unión a antígeno, que comprenden:

(a) inmunizar a una rata o a un ratón con un primer antígeno que comprende un primer epítipo o parte inmunógena del mismo, en el que la rata o el ratón comprende en su genoma: (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada, y (ii) dos o más segmentos génicos de

región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera;

5 (b) permitir que la rata o el ratón genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

(c) aislar de la rata o del ratón una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

10 (d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(e) emplear la secuencia de ácido nucleico de (c) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

15 (f) expresar la secuencia de ácido nucleico de (c) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo del mismo para formar una proteína de unión a antígeno cuya cadena ligera esté codificada por el ácido nucleico de (c) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

También se describen métodos para crear proteínas de unión a antígeno, que comprenden:

25 (a) inmunizar a una rata o a un ratón con un primer antígeno que comprenda un primer epítipo o parte inmunógena del mismo, en el que la rata o el ratón comprende en su genoma: (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera, y (ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena pesada;

30 (b) permitir que la rata o el ratón genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

35 (c) aislar de la rata o del ratón una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

40 (e) emplear la secuencia de ácido nucleico de (c) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

45 (f) expresar la secuencia de ácido nucleico de (c) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo del mismo para formar una proteína de unión a antígeno cuya cadena ligera está codificada por el ácido nucleico de (c) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

50 En diversos aspectos, se describe una rata o un ratón que comprende, en su genoma de línea germinal, un locus de cadena pesada de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. En algunos aspectos, la rata o el ratón son homocigotos para la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada no humana. En determinados aspectos, la secuencia génica de la región constante de cadena pesada no humana codifica una región Fc. En aspectos particulares, la secuencia génica de región constante de cadena pesada no humana es una secuencia génica de región constante de cadena pesada de rata o ratón. En aspectos particulares, la secuencia génica de región constante de cadena pesada se selecciona de un  $C_H1$ , una bisagra, un  $C_H2$ , un  $C_H3$  y de una combinación de los mismos. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada procede de un segmento génico  $V_H$  de cadena pesada humana, un segmento génico D de cadena pesada humana y un segmento génico  $J_H$  de cadena pesada humana. En determinados aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada procede de un segmento  $V_H$  de cadena pesada de la línea germinal humana, un segmento D de cadena pesada de la línea germinal humana y un segmento  $J_H$  de cadena pesada de la línea germinal humana.

65 En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada codifica la secuencia de  $V_{H3-23}/G_Y/J_{H4-4}$  humana. Todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$

funcionales endógenos están delecionados del locus de cadena pesada de inmunoglobulina del ratón o rata no humano o se han hecho no funcionales. En algunos aspectos, un dominio variable de cadena pesada de inmunoglobulina codificada por la secuencia de nucleótidos de región variable de cadena pesada reordenada, no es inmunógeno para la rata o el ratón. En algunos aspectos, la rata o el ratón comprende un gen Adam6a, un gen Adam6b o ambos. En algunos aspectos, la rata o el ratón comprende adicionalmente una secuencia de nucleótidos que codifica un segmento génico de cadena ligera ( $V_L$ ) de inmunoglobulina humana no reordenada y un segmento génico J de cadena ligera de inmunoglobulina humana no reordenada. En determinados aspectos, la secuencia de nucleótidos que codifica el segmento génico V de cadena ligera ( $V_L$ ) no reordenada y el segmento génico de cadena ligera ( $J_L$ ) no reordenada están unidos operativamente a una secuencia génica de región constante de cadena ligera de inmunoglobulina. En aspectos particulares, la secuencia génica de región constante de cadena ligera se selecciona de una secuencia génica de región constante de roedor o de ser humano. En determinados aspectos, el segmento génico de cadena ligera ( $V_L$ ) de inmunoglobulina humana no reordenada y el segmento génico de la cadena ligera ( $J_L$ ) de inmunoglobulina humana no reordenada, están unidos operativamente, en un locus endógeno de rata o de ratón, a una secuencia génica de región constante de inmunoglobulina de rata o ratón. En algunos aspectos, el locus de cadena pesada de inmunoglobulina comprende una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada.

En aspectos adicionales se describe un locus de cadena pesada de inmunoglobulina de rata o ratón en un genoma de una célula germinativa de rata o de ratón, que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia génica de región constante de cadena pesada, en el que la secuencia génica de región constante comprende una secuencia no humana, una secuencia humana o una combinación de las mismas. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de inmunoglobulina humana reordenada. En determinados aspectos, la secuencia génica de región constante de inmunoglobulina no humana endógena es una secuencia génica de región constante de cadena pesada de rata o ratón.

En aspectos adicionales, se describen métodos para crear una rata o un ratón, comprendiendo dichos métodos:

- (a) modificar un genoma de una rata o un ratón para delecionar o hacer no funcionales, segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  de cadena pesada de inmunoglobulina funcional endógena; y
- (b) colocar en el genoma una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada.

La secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada de inmunoglobulina no humana. La secuencia génica de región constante de cadena pesada de inmunoglobulina no humana es una secuencia génica de región constante de cadena pesada de inmunoglobulina de rata o ratón. La secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está colocada en el genoma, en un locus endógeno de cadena pesada de inmunoglobulina. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está presente en un genoma de la línea germinal de la rata o ratón. En algunos aspectos, la rata o el ratón comprenden una pluralidad de copias de la secuencia de nucleótidos de la región variable de la cadena pesada de inmunoglobulina humana reordenada. En algunos aspectos, la rata o el ratón comprende un gen Adam6a, un gen Adam6b o ambos. En algunos aspectos se describe un ratón o una rata que es heterocigoto(a) para el locus de cadena pesada de inmunoglobulina, como se describe en el presente documento, en el que el ratón o la rata expresa la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, predominantemente a partir del locus de cadena pesada de inmunoglobulina.

También se describen ratones y ratas que comprenden un locus de inmunoglobulina modificado genéticamente que comprende:

- (a) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que está unida operativamente a una secuencia génica de región constante de cadena ligera; y
- (b) una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada que está unida operativamente a una secuencia génica de región constante de cadena pesada.

En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$ . En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de la cadena ligera  $\lambda$ . En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera de ratón o de rata. En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera humana. En algunos aspectos, la secuencia génica de

región constante de cadena pesada es una secuencia génica de región constante de cadena pesada de ratón o de rata. En algunos aspectos, la secuencia génica de región constante de cadena pesada codifica una secuencia seleccionada un C<sub>H</sub>1, una bisagra, un C<sub>H</sub>2, un C<sub>H</sub>3 y de una combinación de los mismos. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada comprende una secuencia génica de dominio variable de cadena ligera κ humana. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada comprende una secuencia génica de dominio variable de cadena ligera humana. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana ordenada procede de un segmento génico V<sub>H</sub> de cadena pesada humana, un segmento génico D<sub>H</sub> de cadena pesada humana y un segmento génico J<sub>H</sub> de cadena pesada humana. En determinados aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada procede de un segmento V<sub>H</sub> de cadena pesada de la línea germinal humana, un segmento D<sub>H</sub> de cadena pesada de la línea germinal humana y un segmento J<sub>H</sub> de cadena pesada de línea germinal humana. En determinados aspectos, el segmento génico V<sub>H</sub> humano se selecciona del grupo que consiste en V<sub>H</sub>1-2, V<sub>H</sub>1-3, V<sub>H</sub>1-8, V<sub>H</sub>1-18, V<sub>H</sub>1-24, V<sub>H</sub>1-45, V<sub>H</sub>1-46, V<sub>H</sub>1-58, V<sub>H</sub>1-69, V<sub>H</sub>2-5, V<sub>H</sub>2-26, V<sub>H</sub>2-70, V<sub>H</sub>3-7, V<sub>H</sub>3-9, V<sub>H</sub>3-11, V<sub>H</sub>3-13, V<sub>H</sub>3-15, V<sub>H</sub>3-16, V<sub>H</sub>3-20, V<sub>H</sub>3-21, V<sub>H</sub>3-23, V<sub>H</sub>3-30, V<sub>H</sub>3-30-3, V<sub>H</sub>3-30-5, V<sub>H</sub>3-33, V<sub>H</sub>3-35, V<sub>H</sub>3-38, V<sub>H</sub>3-43, V<sub>H</sub>3-48, V<sub>H</sub>3-49, V<sub>H</sub>3-53, V<sub>H</sub>3-64, V<sub>H</sub>3-66, V<sub>H</sub>3-72, V<sub>H</sub>3-73, V<sub>H</sub>3-74, V<sub>H</sub>4-4, V<sub>H</sub>4-28, V<sub>H</sub>4-30-1, V<sub>H</sub>4-30-2, V<sub>H</sub>4-30-4, V<sub>H</sub>4-31, V<sub>H</sub>4-34, V<sub>H</sub>4-39, V<sub>H</sub>4-59, V<sub>H</sub>4-61, V<sub>H</sub>5-51, V<sub>H</sub>6-1, V<sub>H</sub>7-4-1, V<sub>H</sub>7-81 y una variante polimórfica de los mismos. En determinados aspectos, el segmento génico D humano se selecciona del grupo que consiste en D1-1, D1-7, D1-14, D1-20, D1-26, D2-2, D2-8, D2-15, D2-21, D3-3, D3-9, D3-10, D3-16, D3-22, D4-4, D4-11, D4-17, D4-23, D5-12, D5-5, D5-18, D5-24, D6-6, D6-13, D6-19, D6-25, D7-27, y una variante polimórfica de los mismos. En determinados aspectos, el segmento génico J<sub>H</sub> humano se selecciona del grupo que consiste en J<sub>H</sub>1, J<sub>H</sub>2, J<sub>H</sub>3, J<sub>H</sub>4, J<sub>H</sub>5, J<sub>H</sub>6 y una variante polimórfica de los mismos. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada codifica la secuencia de V<sub>H</sub>3-23/GY/J<sub>H</sub>4-4 (SEQ ID NO: 137) humana. En algunos aspectos, el locus de inmunoglobulina modificado genéticamente está presente en la línea germinal del ratón o de la rata. Todos los segmentos génicos V<sub>H</sub>, D<sub>H</sub> y J<sub>H</sub> funcionales endógenos están deletados del locus de cadena pesada de inmunoglobulina del ratón o de la rata o se han hecho no funcionales. En algunos aspectos, el ratón o la rata comprende un gen Adam6a, un gen Adam6b o ambos. En algunos aspectos, un dominio variable de cadena pesada codificado por la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada no es inmunógeno para el ratón o la rata. En algunos aspectos, el locus de inmunoglobulina modificado genéticamente comprende una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que están unidas operativamente a la secuencia génica de región constante de cadena ligera.

Aspectos adicionales describen un locus de inmunoglobulina en un genoma de la línea germinal de un ratón o de una rata que comprende:

- (1) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que está unida operativamente a una secuencia génica de región constante de cadena ligera, y
- (2) una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada que está unida operativamente a una secuencia génica de región constante de la cadena pesada.

En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera κ. En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera λ. En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera de ratón o rata.

Aspectos adicionales describen métodos de creación de un ratón o una rata que comprende un locus de inmunoglobulina modificado, comprendiendo los métodos:

- (a) modificar un genoma de un ratón o una rata para deletar o hacer no funcionales:
  - (i) segmentos génicos V, D y J de cadena pesada de inmunoglobulina funcionales endógenos, y
  - (ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcionales endógenos; y
- (b) y colocar en el genoma:
  - (i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que está unida operativamente a una secuencia génica de región constante de cadena ligera, y
  - (ii) una secuencia de nucleótidos de región variable de la cadena ligera de inmunoglobulina humana no reordenada que está unida operativamente a una secuencia génica de región constante de cadena pesada.

En algunos aspectos, la secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada codifica un dominio variable de cadena ligera κ. En algunos aspectos, la secuencia de nucleótidos de

región variable de cadena ligera de inmunoglobulina humana no reordenada se codifica como un dominio variable de cadena ligera. En algunos aspectos, la secuencia génica de región constante de cadena pesada es una secuencia génica de región constante de cadena pesada de inmunoglobulina no humana. En determinados aspectos, la secuencia génica de región constante de cadena pesada de inmunoglobulina no humana es una secuencia génica de región constante de cadena pesada de ratón o de rata. En aspectos particulares, la secuencia génica de región constante de cadena pesada codifica una secuencia seleccionada de un CH1, una bisagra, un CH2, un CH3 y una combinación de los mismos. En algunos aspectos, el locus de inmunoglobulina modificado está presente en un genoma de línea germinal de la rata o ratón. En algunos aspectos, el ratón o la rata comprende un gen Adam6a, un gen Adam6b o ambos.

También se describen ratones y ratas que comprenden un locus de cadena pesada de inmunoglobulina modificado que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que comprende una secuencia del segmento V de cadena pesada (V<sub>H</sub>) que está unida operativamente, mediante un espaciador, a una secuencia del segmento J de cadena pesada (J<sub>H</sub>), en la que el espaciador comprende al menos un resto de aminoácido. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada no humana. En algunos aspectos, la secuencia génica de región constante de cadena pesada no humana es una secuencia génica de región constante de un ratón o una rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada humana. En determinados aspectos, la secuencia génica de región constante de cadena pesada codifica una secuencia seleccionada de un C<sub>H</sub>1, una bisagra, un C<sub>H</sub>2, un C<sub>H</sub>3 y una combinación de los mismos. En algunos aspectos, la secuencia V<sub>H</sub> y la secuencia J<sub>H</sub> proceden de un segmento génico V<sub>H</sub> humano y de un segmento génico J<sub>H</sub> humano. En determinados aspectos, el segmento génico J<sub>H</sub> humano se selecciona del grupo que consiste en J<sub>H</sub>1, J<sub>H</sub>2, J<sub>H</sub>3, J<sub>H</sub>4, J<sub>H</sub>5, J<sub>H</sub>6 y una variante polimórfica del mismo. En determinados aspectos, el segmento génico J<sub>H</sub> humano es J<sub>H</sub>4-4 o una variante polimórfica del mismo. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina reordenada codifica la secuencia de V<sub>H</sub>3-23/GY/J<sub>H</sub>4-4 (SEQ ID NO: 137) humana. Todos los segmentos génicos V<sub>H</sub>, D y J<sub>H</sub> funcionales endógenos están delecionados del locus variable de cadena pesada de inmunoglobulina de la rata o ratón o lo hacen no funcional. En algunos aspectos, un dominio variable de cadena pesada codificado por la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada no es inmunógeno para el ratón o la rata. En algunos aspectos, el ratón o la rata comprende un gen Adam6a, un gen Adam6b o ambos.

Aspectos adicionales describen un locus de cadena pesada de inmunoglobulina en un genoma de línea germinal de una rata o un ratón, que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que comprende un segmento génico variable de cadena pesada (V<sub>H</sub>) que está unido operativamente, mediante un espaciador, a un segmento génico J de cadena pesada (J<sub>H</sub>), en el que el espaciador codifica al menos un resto de aminoácido. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada no humana. En determinados aspectos, la secuencia génica de región constante de cadena pesada no humana es una secuencia génica de región constante de cadena pesada de ratón o rata. En algunos aspectos, el locus de inmunoglobulina comprende una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada reordenada.

Se describen aspectos adicionales que son métodos para crear una rata o un ratón que comprende un locus de cadena pesada de inmunoglobulina modificado, comprendiendo los métodos:

(a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcionales, segmentos génicos V<sub>H</sub>, D y J<sub>H</sub> de cadena pesada de inmunoglobulina funcionales endógenos; y

(b) colocar en el genoma una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que comprende un segmento génico variable de cadena pesada (V<sub>H</sub>) que está unido operativamente, mediante un espaciador, a un segmento génico J de cadena pesada (J<sub>H</sub>), en el que el espaciador comprende al menos un resto de aminoácido.

En algunos aspectos, la secuencia de nucleótidos de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada de inmunoglobulina no humana. La secuencia génica de región constante de cadena pesada de inmunoglobulina no humana es una secuencia génica de región constante de cadena pesada de ratón o de rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada se coloca un locus de cadena pesada de inmunoglobulina endógena en el genoma. En algunos aspectos, el locus de cadena pesada modificada comprende una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. En algunos aspectos, el ratón o la rata comprende un gen Adam6a, un gen Adam6b o ambos. En algunos aspectos, el locus de cadena pesada de inmunoglobulina está presente en un genoma de la línea germinal del ratón o de la rata.

Aspectos adicionales describen una rata o un ratón modificado que comprende en su genoma:

(a) un locus de cadena pesada de inmunoglobulina que comprende una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y

(b) un locus de cadena ligera de inmunoglobulina que comprende uno o más segmentos génicos de cadena ligera  $V_L$  y  $J_L$  de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera.

En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia de la región constante de ratón o de rata que codifica un isotipo de inmunoglobulina seleccionado de IgM, IgD, IgG, IgE, IgA y de una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia de ácido nucleico de región constante  $\kappa$  o  $\lambda$  de roedor. En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena pesada humana reordenada se selecciona de un segmento  $V_H$  de la línea germinal humana, un segmento D de la línea germinal humana y un segmento  $J_H$  de la línea germinal humana. En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena pesada humana reordenada está unida operativamente a la secuencia de ácido nucleico de región constante en un locus endógeno. En algunos aspectos, el uno o más segmentos génicos de cadena ligera  $V_L$  y  $J_L$  de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, están unidos operativamente a la secuencia de ácido nucleico de región constante de cadena ligera en un locus endógeno. En algunos aspectos, el ratón o la rata comprenden no más de dos segmentos génicos  $V_L$  no reordenados de inmunoglobulina humana, y uno, dos o tres o más segmentos génicos  $J_L$  no reordenados humanos. En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena pesada reordenada codifica la secuencia de  $V_H3-23/GY/J_H4-4$  (SEQ ID NO: 137) humana. La secuencia de ácido nucleico de región variable de cadena pesada reordenada codifica la secuencia de  $V_H3-23/X_1X_2/J$ , en la que  $X_1$  y  $X_2$  es cualquier aminoácido. En determinadas realizaciones  $X_1$  es Gly y  $X_2$  es Tyr. En algunos aspectos, el locus de cadena pesada de inmunoglobulina comprende un gen Adam6a funcional, un gen Adam6b, o ambos. En determinados aspectos, el gen Adam6a, el gen Adam6b o ambos son genes Adam6 endógenos. En algunos aspectos, el ratón o la rata modificados genéticamente comprende un gen Adam6a, un gen Adam6b o ambos en un locus ectópico del genoma. En algunos aspectos, los segmentos génicos  $V_L$  y  $J_L$  de región variable humana pueden reordenarse y codificar un dominio variable de cadena ligera de inmunoglobulina humana. En algunos aspectos, el locus de cadena ligera de inmunoglobulina comprende dos segmentos génicos  $V_L$  humanos, Vk1-39 y Vk3-20. En determinados aspectos, los segmentos génicos son segmentos génicos de la línea germinal. En determinados aspectos, el ratón o la rata comprenden los segmentos génicos Jk1, Jk2, Jk3, Jk4 y Jk5. En algunos aspectos, dos o más, tres o más, cuatro o más, o cinco o más segmentos génicos  $V_L$  humanos y dos o más segmentos génicos  $J_L$  humanos están presentes en un locus de cadena ligera endógena. En algunos aspectos, al menos uno de los segmentos génicos  $V_L$  o  $J_L$  de cadena ligera humana codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos, al menos uno de los segmentos génicos VL comprende una adición o sustitución de al menos un codón no histidina, codificado por la secuencia del segmento  $V_L$  de la línea germinal humana correspondiente con un codón de histidina. En determinados aspectos, el codón de histidina añadido o sustituido está presente en la CDR3. En algunos aspectos, los segmentos génicos  $V_L$  humanos son segmentos génicos Vk1-39 y Vk3-20 humanos, y cada uno de los segmentos génicos Vk1-39 y Vk3-20 humanos comprende una sustitución de al menos un codón no histidina, codificado por un segmento génico  $V_L$  de la línea germinal humana correspondiente, con el codón de histidina. En determinados aspectos, la sustitución es de tres codones no histidina del segmento génico Vk1-39 humano, en el que la sustitución se diseña para expresar histidinas en las posiciones 106, 108 y 111. En aspectos particulares, la sustitución es de cuatro codones no histidina del segmento génico Vk1-39 humano y la sustitución se diseña para expresar histidinas en las posiciones 105, 106, 108 y 111. En aspectos particulares, la sustitución es de tres codones no histidina del segmento génico Vk3-20 humano, y la sustitución se diseña para expresar histidinas en las posiciones 105, 106 y 109. En aspectos particulares, la sustitución es de cuatro codones no histidina del segmento génico Vk3-20 humano y la sustitución se diseña para expresar histidinas en las posiciones 105, 106, 107 y 109. En algunos aspectos, se describe un ratón o una rata, en el que el ratón o la rata, después de la estimulación con un antígeno de interés, expresa una proteína de unión a antígeno que se une específicamente al antígeno, en el que la proteína de unión a antígeno comprende una secuencia de aminoácidos procedente de los segmentos génicos  $V_L$  y  $J_L$  humanos, y en el que la proteína de unión a antígeno comprende al menos un resto de histidina en una posición de aminoácido codificado por el segmento génico  $V_L$  humano. En algunos aspectos, el ratón o la rata expresa una población de proteínas de unión a antígeno en respuesta a un antígeno, en el que todas las proteínas de unión a antígeno en la población comprenden: (a) cadenas pesadas de inmunoglobulina que comprenden dominios variables de cadena pesada humana procedentes de la secuencia de ácido nucleico de región variable humana reordenada; y (b) cadenas ligeras de inmunoglobulina que comprenden dominios variables de cadena ligera de inmunoglobulina procedentes de un reordenamiento de los segmentos génicos  $V_L$  y  $J_L$  humanos en el genoma de la rata o del ratón, y en el que al menos uno de los segmentos génicos  $V_L$  humanos codifica uno o más codones de histidina que no están codificados por el segmento génico  $V_L$  de la línea germinal humana correspondiente.

Aspectos adicionales describen métodos para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado, que comprende:

- 5 (a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcional:
- (i) segmentos génicos  $V_H$ ,  $D_H$  y/o  $J_H$  de cadena pesada de inmunoglobulina funcional endógena, y
- (ii) segmentos génicos  $V_L$  y  $J_L$  de cadena ligera de inmunoglobulina funcional endógena; y
- 10 (b) colocar en el genoma modificado de la rata o del ratón:
- (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada en unión operativa con una secuencia de ácido nucleico de región constante de cadena pesada de inmunoglobulina; y
- 15 (ii) uno o más segmentos génicos de cadena ligera  $V_L$  y  $J_L$  de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, en unión operativa con una secuencia de ácido nucleico de región constante de cadena ligera de inmunoglobulina.
- 20 En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada de ratón o de rata seleccionada de un  $CH1$ , una bisagra, un  $CH2$ , un  $CH3$  y una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada y los segmentos génicos de cadena ligera  $V_L$  y  $J_L$  de inmunoglobulina humana, se colocan en, o cerca
- 25 de, una secuencia de nucleótidos correspondiente del ratón o rata de tipo silvestre. En algunos aspectos, el ratón o la rata comprende un locus de cadena pesada de inmunoglobulina que comprende un gen endógeno Adam6a, un gen Adam6b o ambos. En algunos aspectos, el ratón o la rata comprende un gen Adam6a, un gen Adam6b o ambos en un locus ectópico del genoma. En algunas realizaciones, la secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada codifica la secuencia de  $V_{H3-23}/G_Y/J_{H4-4}$  (SEQ ID NO:
- 30 137) humana. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera de inmunoglobulina es una secuencia de ácido nucleico de región constante  $C_k$  de rata o de ratón. En algunos aspectos, los segmentos génicos de cadena ligera  $V_L$  y  $J_L$  de inmunoglobulina humana son capaces de reordenarse y codificar un dominio variable de cadena ligera de inmunoglobulina humana. En algunos aspectos, la rata o el ratón comprende un locus de cadena ligera de inmunoglobulina que comprende dos segmentos génicos  $V_L$  humanos,
- 35  $V_{k1-39}$  y  $V_{k3-20}$ . En determinados aspectos, la rata o el ratón comprende los segmentos génicos  $J_{k1}$ ,  $J_{k2}$ ,  $J_{k3}$ ,  $J_{k4}$ , y  $J_{k5}$ . En algunos aspectos, dos o más, tres o más, cuatro o más, o cinco o más segmentos génicos  $V_L$  humanos y dos o más segmentos génicos  $J_L$  humanos están presentes en un locus de cadena ligera endógena. En algunos aspectos, al menos uno de los segmentos génicos de cadena ligera  $V_L$  o  $J_L$  de inmunoglobulina humana codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la
- 40 línea germinal humana correspondiente. En algunos aspectos, al menos uno de los segmentos génicos  $V_L$  comprende una adición o una sustitución de al menos un codón no de histidina codificado por la secuencia del segmento  $V_L$  de la línea germinal humana correspondiente con un codón de histidina. En determinados aspectos, el codón de histidina añadido o sustituido está presente en la CDR3. En determinados aspectos, los segmentos génicos  $V_L$  humanos son los segmentos génicos  $V_{k1-39}$  y  $V_{k3-20}$  humanos, y cada uno de los segmentos génicos
- 45  $V_{k1-39}$  y  $V_{k3-20}$  humanos comprenden la sustitución de al menos un codón no de histidina codificado por un segmento génico  $V_L$  de la línea germinal humana correspondiente con el codón de histidina. En aspectos particulares, la sustitución es de tres codones no histidina del segmento génico  $V_{k1-39}$  humano, y en el que la sustitución se diseña para expresar histidinas en las posiciones 106, 108 y 111. En aspectos particulares, la sustitución es de cuatro codones no histidina del segmento génico  $V_{k1-39}$  humano, y en el que la sustitución se diseña para expresar histidinas en las posiciones 105, 106 y 111. En aspectos particulares, la sustitución es de tres codones no histidina del segmento génico  $V_{k3-20}$  humano y en el que la sustitución se diseña para expresar histidinas en las posiciones 105, 106 y 109. En aspectos particulares, la sustitución es de cuatro codones no histidina del segmento génico  $V_{k3-20}$  humano, y la sustitución se diseña para expresar histidina en las posiciones
- 50 105, 106, 107 y 109.
- 55 Aspectos adicionales describen métodos para obtener una secuencia de ácido nucleico que codifica un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:
- 60 (a) inmunizar a un ratón o a una rata con un antígeno de interés o con un inmunógeno del mismo, en el que el ratón o la rata comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y (ii) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a
- 65 una secuencia de ácido nucleico de región constante de cadena ligera;

- (b) permitir que el ratón o la rata cree una respuesta inmunitaria;
- (c) aislar del ratón o de la rata inmunizados una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

5 (d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de la cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

En algunos aspectos, la etapa de aislamiento (c) se realiza mediante separación de células activadas por fluorescencia (FACS, siglas del inglés *fluorescence-activated cell sorting*) o citometría de flujo. En algunos aspectos, la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno, es un linfocito. En determinados aspectos, el linfocito comprende linfocitos citolíticos naturales, linfocitos T o linfocitos B. En algunos aspectos, el método comprende adicionalmente una etapa (c)' de fusionar el linfocito con una célula cancerosa. En determinados aspectos, la célula cancerosa es una célula de mieloma. En algunos aspectos, la secuencia de ácido nucleico de (d) se fusiona con una secuencia de ácido nucleico que codifica una secuencia de ácido nucleico de la región constante de inmunoglobulina. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa humana o una secuencia lambda humana. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia humana seleccionada de un CH1, una bisagra, un CH2, un CH3 y una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de (d) comprende una o más sustituciones o inserciones de codones de histidina que proceden del segmento génico  $V_L$  no reordenado en el genoma del ratón o rata.

Aspectos adicionales descritos son métodos para crear una proteína de unión a antígeno que comprende un dominio variable de cadena ligera de inmunoglobulina que puede unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

25 (a) inmunizar a una rata o a un ratón modificado genéticamente, con un primer antígeno que comprende un primer epítipo o parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma:

30 (i) una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y

35 (ii) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera de inmunoglobulina;

(b) permitir que el ratón o la rata cree una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

40 (c) aislar del ratón o de la rata una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(d) obtener de la célula de (c) la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une específicamente al primer epítipo o parte inmunógena del mismo;

45 (e) emplear la secuencia de ácido nucleico de (d) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

50 (f) expresar la secuencia de ácido nucleico de (d) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo para formar una proteína de unión a antígeno cuya cadena ligera está codificada por el ácido nucleico de (d) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

En algunos aspectos, al menos uno de los segmentos génicos de cadena ligera  $V_L$  o  $J_L$  humana codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos, el primer epítipo procede de un receptor de superficie celular. En aspectos particulares, el receptor de superficie celular es un receptor Fc. En aspectos aún más particulares, el receptor Fc es el receptor Fc neonatal (FcRn, *neonatal Fc receptor*). En algunos aspectos, el segundo antígeno o epítipo procede de un antígeno soluble. En algunos aspectos, el segundo antígeno o epítipo procede de un receptor de superficie celular. En algunos aspectos, el primer antígeno es un receptor Fc, el segundo antígeno es una proteína soluble, y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma del ratón o de la rata.

65 Aspectos adicionales descritos son métodos para obtener una secuencia de ácido nucleico que codifica un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

(a) inmunizar a un ratón o a una rata con un primer antígeno que comprende el primer epítipo o parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma: (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de la cadena pesada, y (ii) una secuencia de ácido nucleico de región variable de cadena ligera de inmunoglobulina humana no reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera;

(b) permitir que el ratón o la rata cree una respuesta inmunitaria;

(c) aislar del ratón o de la rata inmunizado una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

(d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

En algunos aspectos, la etapa de aislamiento (c) se realiza mediante separación de células activadas por fluorescencia (FACS) o citometría de flujo. En algunos aspectos, la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno es un linfocito. En determinados aspectos, el linfocito comprende linfocitos citolíticos naturales, linfocitos T o linfocitos B. En algunos aspectos, el método comprende adicionalmente: (c)' fusionar el linfocito con una célula cancerosa. En determinados aspectos, la célula cancerosa es una célula de mieloma. En algunos aspectos, la secuencia de ácido nucleico de (d) se fusiona con una secuencia de ácido nucleico que codifica una secuencia de ácido nucleico de la región constante de inmunoglobulina. En algunos aspectos, la secuencia de ácido nucleico de la región constante de la cadena ligera es una secuencia kappa o lambda humana. En algunos aspectos, la secuencia de ácido nucleico de la región constante de la cadena pesada es una secuencia humana seleccionada de un CH1, una bisagra, un CH2, un CH3 y una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de (d) comprende una o más sustituciones o inserciones de codones de histidina que proceden del segmento génico  $V_L$  no reordenado en el genoma del ratón o de la rata.

Aspectos adicionales descritos son métodos para obtener una secuencia de ácido nucleico que codifica un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

(a) inmunizar a un ratón o a una rata con un primer antígeno que comprende un primer epítipo o parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma: (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera, y (ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina unidos operativamente a una secuencia de ácido nucleico de región constante de cadena pesada;

(b) permitir que el ratón o la rata cree una respuesta inmunitaria;

(c) aislar del ratón o de la rata inmunizado una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

(d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de la cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

En algunos aspectos, la etapa de aislamiento (c) se realiza mediante separación de células activadas por fluorescencia (FACS) o citometría de flujo. En algunos aspectos, la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno es un linfocito. En determinados aspectos, el linfocito comprende linfocitos citolíticos naturales, linfocitos T o linfocitos B. En algunos aspectos, el método comprende adicionalmente: (c)' fusionar el linfocito con una célula cancerosa. En determinados aspectos, la célula cancerosa es una célula de mieloma. En algunos aspectos, la secuencia de ácido nucleico de (d) se fusiona con una secuencia de ácido nucleico que codifica una secuencia de ácido nucleico de la región constante de inmunoglobulina. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa o lambda humana. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia humana seleccionada de un CH1, una bisagra, un CH2, un CH3 o una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de (d) comprende una o más sustituciones o inserciones de codones de histidina que proceden del segmento génico  $V_L$  no reordenado en el genoma del ratón o rata.

Aspectos adicionales descritos son métodos para crear una proteína de unión a antígeno que comprende un dominio variable de cadena ligera de inmunoglobulina que puede unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

(a) inmunizar a un ratón o a una rata modificado genéticamente con un primer antígeno que comprende un primer epítipo o una parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma:

5 (i) una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y

(ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana no reordenada unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera;

10 (b) permitir que el ratón o la rata cree una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

15 (c) aislar del ratón o de la rata una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique el dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

20 (e) emplear la secuencia de ácido nucleico de (d) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

25 (f) expresar la secuencia de ácido nucleico de (d) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo para formar una proteína de unión a antígeno cuya cadena ligera está codificada por el ácido nucleico de (d) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

30 En algunos aspectos, al menos uno de los segmentos génicos de cadena ligera  $V_L$  o  $J_L$  humana codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos, el primer epítipo procede de un receptor de superficie celular. En determinados aspectos, el receptor de superficie celular es un receptor Fc. En aspectos particulares, el receptor Fc es FcRn. En algunos aspectos, el segundo antígeno o epítipo procede de un antígeno soluble. En algunos aspectos, el segundo antígeno o epítipo procede de un receptor de superficie celular. En algunos aspectos, el primer antígeno es un receptor Fc, el segundo antígeno es una proteína soluble y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma del ratón o la rata.

40 Aspectos adicionales descritos son métodos para crear una proteína de unión a antígeno que comprende un dominio variable de cadena ligera de inmunoglobulina que puede unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

45 (a) inmunizar a un ratón o a una rata modificado genéticamente con un primer antígeno que comprende un primer epítipo o parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma:

(i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera; y

50 (ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana no reordenada unidos operativamente a una secuencia de ácido nucleico de región constante de cadena pesada;

55 (b) permitir que el ratón o la rata cree una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

(c) aislar del ratón o de la rata una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

60 (d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique el dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(e) emplear la secuencia de ácido nucleico de (d) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

65 (f) expresar la secuencia de ácido nucleico de (d) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo para formar

una proteína de unión a antígeno cuya cadena ligera esté codificada por el ácido nucleico de (d) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

5 En algunos aspectos, al menos uno de los segmentos génicos  $V_L$  o  $J_L$  de la cadena ligera humana codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos el primer epítipo procede de un receptor de superficie celular. En determinados aspectos, el receptor de superficie celular es un receptor Fc. En aspectos particulares, el receptor Fc es FcRn. En algunos aspectos, el segundo antígeno o epítipo procede de un antígeno soluble. En algunos aspectos, el segundo antígeno o epítipo procede de un receptor de superficie celular. En algunos aspectos, el primer antígeno es un receptor Fc, el segundo antígeno es una proteína soluble y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma del ratón o de la rata.

### 15 Breve descripción de los dibujos

La **FIG. 1** ilustra esquemas de la construcción de un mini locus ("mini locus UHC") de dominio variable de cadena pesada reordenada, que comprende una secuencia de nucleótidos de región variable ( $V_{H3-23/D/J_{H4}}$ ; SEQ ID NO: 136) de inmunoglobulina humana reordenada y un intrón de  $J_{H4}$  (SEQ ID NO: 140), que están unidos operativamente a un promotor  $V_{H3-23}$  humano (SEQ ID NO: 139). El mini locus UHC se flanqueó en 5' y 3' con brazos de homología de ratón. En la Etapa 1 (Ligamiento I-Ceul/Spel (Amp+Espec)), entre los sitios I-Ceul y Spel, cadena arriba del promotor, se introdujo un casete de selección de espectinomina para generar pJSh0038 (minilocus UHC).

25 La **FIG. 2** ilustra esquemas de (A) direccionamiento de un casete de selección de higromicina (EM7-HIG) en el extremo 5' del clon MAID 1115 BAC (2. BHR (Hig+Kan)); y (B) direccionamiento del minilocus UHC (pJSh0038) en el clon VI432 BAC (3. BHR (Espec+Hig)), en el locus de IgM cadena arriba.

30 La **FIG. 3** ilustra esquemas de (A) direccionamiento de la construcción pDBa0049 que comprende un casete de cloranfenicol en el extremo 3' del clon VI421, que comprende, de 5' a 3', un gen Adam6a (presente en una dirección 3' a 5'); un casete de neomicina (presente en una dirección de 3' a 5') flanqueado por sitios FRT; un gen Adam6b (presente en una dirección 3' a 5'); la Región de Control Intergénica 1 (IGCR1, *Intergenic Control Region 1*; una región reguladora de recombinación V(D)J clave); y un casete de espectinomina (presente en una dirección 5' a 3') (4. BHR (Cm+Kan)); y (B) direccionamiento del locus genómico del clon VI444 BAC que contiene los genes Adam6a y 6b en el locus genómico de cadena pesada universal (UHC, *universal heavy chain*) cadena arriba del clon VI443 BAC entre los sitios I-Ceul y Ascl mediante digestión con enzimas de restricción y ligamiento (ligamiento 5. I-Ceul/Ascl (Hig+Kan)).

40 La **FIG. 4** ilustra esquemas de (A) direccionamiento de la construcción final (ADN de MAID6031 BAC) en células ES aisladas del ratón heterocigoto 1661; y muestra (B) la localización genómica de las sondas y de los cebadores utilizados en los ensayos de exploración.

45 La **FIG. 5** muestra un listado de anticuerpos en la base de datos ASAP de Regeneron Pharmaceuticals que contienen secuencias de CDR3 similares a las de la secuencia CDR3 de UHC (AKDYSNYYFDY, SEQ ID NO: 143).

50 La **FIG. 6** ilustra la organización genómica del ADN del cromosoma artificial bacteriano (BAC, *Bacterial Artificial Chromosome*) 6031 y células ES heterocigotas 6031, y la localización genómica de los cebadores y de las sondas utilizados en los ensayos de exploración.

La **FIG. 7** muestra un listado de cebadores y sondas utilizados para confirmar una pérdida de alelo (LOA, *loss of allele*), una ganancia de alelo (GOA, *gain of allele*) o un alelo parental (parental) en los ensayos de exploración.

55 La **FIG. 8** muestra secuencias de cebadores y sondas utilizados en los ensayos de exploración.

La **FIG. 9** ilustra la estructura genómica del locus de cadena pesada de inmunoglobulina de ratones F0 modificados genéticamente, que contiene una copia del alelo diana (incluyendo los genes Adam6a/6b y la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada ( $hV_{H3-23}D/J_{H4}$ )). **(A)** MAID 6031 het: un ratón F0 heterocigoto que comprende un locus de cadena pesada de inmunoglobulina modificado genéticamente con un casete de selección; **(B)** MAID 6032 het: un ratón F0 heterocigoto que comprende un locus de cadena pesada de inmunoglobulina modificado genéticamente sin un casete de selección.

65 La **FIG. 10** muestra el resultado del análisis de separación de células activadas por fluorescencia (FACS) de células de la médula ósea aisladas de un ratón heterocigoto 6032 o de tipo silvestre. Panel superior: células de médula ósea aisladas de un ratón heterocigoto 6032 F0 o de tipo silvestre se seleccionaron en singletes y se

separaron basándose en la expresión de CD19 (un marcador de linfocitos B) y de CD3 (un marcador de linfocitos T). Panel inferior: linfocitos B seleccionados CD19<sup>+</sup> se separaron basándose en la presencia de anticuerpos IgM<sup>b</sup> (anticuerpos producidos a partir de un alelo de tipo silvestre; alelo B6) o anticuerpos IgM<sup>a</sup> (anticuerpos producidos a partir del alelo modificado genéticamente (alelo 129) que codifica un dominio variable de cadena pesada reordenada (hV<sub>H</sub>3-23(D)<sub>J<sub>H</sub>4</sub>).

La **FIG. 11** muestra el resultado del análisis FACS de esplenocitos aislados de un ratón heterocigoto 6032 o de tipo silvestre. Panel superior: esplenocitos aislados de un ratón heterocigoto 6032 F0 o de tipo silvestre se seleccionaron en singletes y se separaron basándose en la expresión de CD19 (un marcador de linfocitos B) y de CD3 (un marcador de linfocitos T). Panel inferior: linfocitos B seleccionados CD19<sup>+</sup> se separaron basándose en la presencia de anticuerpos IgM<sup>b</sup> (anticuerpos producidos a partir de un alelo de tipo silvestre; alelo B6) o anticuerpos IgM<sup>a</sup> (anticuerpos producidos a partir del alelo modificado genéticamente (alelo 129) que codifica un dominio variable de cadena pesada reordenada (hV<sub>H</sub>3-23(D)<sub>J<sub>H</sub>4</sub>).

La **FIG. 12** muestra el resultado del análisis FACS de las células sanguíneas aisladas de un ratón heterocigoto 6032 o de tipo silvestre. Panel superior: células sanguíneas aisladas de un ratón heterocigoto 6032 F0 o de tipo silvestre se seleccionaron en singletes y se separaron basándose en la expresión de CD19 (un marcador de linfocitos B) y de CD3 (un marcador de linfocitos T). Panel inferior: linfocitos B seleccionados CD19<sup>+</sup> se separaron basándose en la presencia de anticuerpos IgM<sup>b</sup> (anticuerpos producidos a partir de un alelo de tipo silvestre; alelo B6) o anticuerpos IgM<sup>a</sup> (anticuerpos producidos a partir del alelo modificado genéticamente (alelo 129) que codifica un dominio variable de cadena pesada reordenada (hV<sub>H</sub>3-23(D)<sub>J<sub>H</sub>4</sub>).

La **FIG. 13A** muestra los resultados del análisis FACS del número total de linfocitos B inmaduros CD19<sup>+</sup>(CD19+IgD<sup>int</sup>Ig-M<sup>hi</sup>) y maduros (CD19+IgM<sup>lo</sup>IgD<sup>hi</sup>) en bazos extraídos de ratones de tipo silvestre (TS) y homocigotos (6032HO) de una secuencia de nucleótidos de región variable de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4). Panel superior: células de bazo aisladas de un ratón de tipo silvestre u homocigoto 6032 F2 se seleccionaron en singletes y se separaron basándose en la expresión de CD19 (un marcador de linfocitos B) y de CD3 (un marcador de linfocitos T). El panel inferior muestra gráficas de contorno representativas de esplenocitos seleccionados en CD19<sup>+</sup> y teñidos para inmunoglobulina D (IgD) e inmunoglobulina M (IgM) de un ratón de tipo silvestre (TS) y un ratón homocigoto para una secuencia de nucleótidos de región variable de inmunoglobulina humana de cadena pesada reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4). Dentro de cada región seleccionada se muestra el porcentaje de células.

La **FIG. 13B** muestra el número total de linfocitos B (CD19<sup>+</sup>), linfocitos B maduros (CD19+IgD<sup>hi</sup> IgM<sup>lo</sup>) y linfocitos B inmaduros (CD19+IgD<sup>int</sup>IgM<sup>hi</sup>) en bazos extraídos de ratones de tipo silvestre (TS) y homocigotos para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4).

La **FIG. 13C** muestra gráficas de contorno representativas de esplenocitos Igλ<sup>+</sup> e Igκ<sup>+</sup> seleccionados en CD19<sup>+</sup> de un ratón de tipo silvestre (TS) y un ratón homocigoto (6032HO) para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4).

La **FIG. 13D** muestra el número total de linfocitos B (CD19<sup>+</sup>), linfocitos B Igκ<sup>+</sup> (CD19+Igkappa<sup>+</sup>) y linfocitos B Igλ<sup>+</sup> (CD19+Ig-lambda<sup>+</sup>) en bazos extraídos de ratones de tipo silvestre (TS) y homocigotos para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4).

La **FIG. 13E** muestra el desarrollo de linfocitos B periféricos en un ratón de tipo silvestre y ratones homocigotos para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4). La primera columna (izquierda) de las gráficas de contorno muestran esplenocitos CD93<sup>+</sup> y B220<sup>+</sup> seleccionados en CD19<sup>+</sup> indicando linfocitos B inmaduros y maduros. La segunda columna (central) de la gráfica de contorno muestra la expresión de IgM<sup>+</sup> y CD23<sup>+</sup> en linfocitos B inmaduros indicando poblaciones de linfocitos B T1 (IgD-IgM+CD21<sup>lo</sup>CD23<sup>-</sup>), T2 (IgD<sup>hi</sup>IgM<sup>hi</sup>CD21<sup>mid</sup>CD23<sup>+</sup>) y T3. La tercera columna (derecha) de la gráfica de contorno muestra la expresión de linfocitos B maduros CD21<sup>+</sup>(CD35<sup>+</sup>) e IgM<sup>+</sup> indicando una primera población más pequeña que da lugar a linfocitos B de la zona marginal y a una segunda población que da lugar a linfocitos B foliculares (FO). Dentro de cada región seleccionada se muestra el porcentaje de células.

La **FIG. 14A** muestra gráficas de contorno representativas de médula ósea teñida para linfocitos B y T (CD19<sup>+</sup> y CD3<sup>+</sup>, respectivamente) de un ratón de tipo silvestre (TS) y un ratón homocigoto para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4).

La **FIG. 14B** muestra el número absoluto de células (izquierda), el número total de células (centro) y el número total de linfocitos B (CD19<sup>+</sup>) (derecha) en médula ósea aislada de los fémures de ratones de tipo silvestre (TS) y homocigotos para la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4).

- 5 La **FIG. 14C** muestra gráficas de contorno representativas de médula ósea seleccionada en singletes teñida para inmunoglobulina M (IgM) y B220 de un ratón de tipo silvestre (TS) y un ratón homocigoto para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana (V<sub>H</sub>3-23/D/J<sub>H</sub>4) reordenada. En cada una de las gráficas de contorno se indican linfocitos B inmaduros, maduros y pro/pre linfocitos B.
- 10 La **FIG. 14D** muestra el número total de linfocitos B maduros (B220<sup>hi</sup>IgM+) e inmaduros (B220<sup>int</sup>IgM+) en médula ósea aislada de los fémures de ratones de tipo silvestre (TS) y homocigotos para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4).
- 15 La **FIG. 14E** muestra gráficas de contorno representativas de médula ósea seleccionada en singletes teñida para inmunoglobulina M (IgM) y B220 de un ratón de tipo silvestre (TS) y de ratones homocigotos para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4) En cada una de las gráficas de contorno se ubican linfocitos B inmaduros, maduros y pro/pre linfocitos B.
- 20 La **FIG. 14F** muestra gráficas de contorno representativas de médula ósea seleccionada en linfocitos B inmaduros (B220<sup>int</sup>IgM+) y maduros (B220<sup>hi</sup>IgM+) teñida para la expresión de Igκ e Igλ aislada de los fémures de un ratón de tipo silvestre (TS) y de ratones homocigotos para una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (V<sub>H</sub>3-23/D/J<sub>H</sub>4).
- 25 La **FIG. 15** muestra los niveles de mIgG específicas de antígeno en los sueros de ratón (F0 y F1 de tipo silvestre o 6031 HET) los días 15 y 24 después de la inmunización en las almohadillas plantares.
- La **FIG. 16** muestra la secuencia de nucleótidos optimizada con codones y la secuencia de aminoácidos deducida de hV<sub>H</sub>3-23 (D4-4\_Fase de Lectura 3)J<sub>H</sub>6 (SEQ ID NO: 145).
- La **FIG. 17** muestra la secuencia de nucleótidos optimizada con codones y la secuencia de aminoácidos deducida de hV<sub>H</sub>3-23 (D4-4\_Fase de Lectura 2)J<sub>H</sub>6 (SEQ ID NO: 146).
- 30 La **FIG. 18** muestra la secuencia de nucleótidos optimizada con codones y la secuencia de aminoácidos deducida de hV<sub>H</sub>3-23 (D4-4\_Fase de Lectura 3)J<sub>H</sub>4 (SEQ ID NO: 147).
- La **FIG. 19** muestra la secuencia de nucleótidos optimizada con codones y la secuencia de aminoácidos deducida de hV<sub>H</sub>3-23 (D4-4\_Fase de Lectura 2)J<sub>H</sub>4 (SEQ ID NO: 148).
- 35 La **FIG. 20** muestra ejemplos de dos locus de doble cadena ligera (DLC, *dual light chain*) modificados genéticamente. El locus de la parte superior (DLC-5J) contiene un fragmento de ADN humano modificado por ingeniería genética que contiene dos segmentos génicos Vk humanos y cinco segmentos génicos Jk humanos. El locus de la parte inferior (DLC-1J) contiene un fragmento de ADN humano modificado por ingeniería genética que contiene dos segmentos génicos Vk humanos y un segmento génico Jk humano. Cada locus puede reordenarse para formar una región Vk humana unida operativamente a una región constante de cadena ligera endógena (por ejemplo, una Cκ). Se muestran promotores de inmunoglobulina (P, flecha blanca encima del locus), exones líder (L, flechas blancas cortas) y los dos segmentos génicos Vk humanos (flechas blancas largas), todos ellos flanqueados cadena arriba (5') por un casete de neomicina que contiene sitios de recombinación Frt. Las secuencias señal de recombinación modificadas por ingeniería genética con cada uno de los segmentos génicos (Vk y Jk) humanos se indican con óvalos blancos yuxtapuestos con cada segmento génico. En la mayoría de los aspectos, a menos que se indique de otra manera, las formas oscuras y las líneas continuas representan secuencias de ratón, y las formas blancas y las líneas dobles representan secuencias humanas. Los diagramas no se presentan a escala.
- 40
- 45
- 50 Las **FIGS. 21A-21C** muestran una estrategia general de construcción de un vector diana para la modificación por ingeniería genética de un locus kappa de inmunoglobulina que comprende dos segmentos Vk humanos (hVk1-39 y hVk3-20) y un segmento Jk humano (JK5), así como potenciadores de ratón y un brazo de IgκC. La **FIG. 21D** muestra la introducción de este vector diana en células ES y la generación de ratones heterocigotos con el mismo; mientras que la **FIG. 21E** muestra la delección del casete de selección en células ES utilizando la enzima FLP. En la mayoría de los aspectos, salvo que se indique de otra manera, las formas oscuras y las líneas continuas representan secuencias de ratón, y las formas blancas y las líneas dobles representan las secuencias humanas. Los diagramas no se presentan a escala.
- 55
- 60 Las **FIGS. 22A-22D** muestran la secuencia de nucleótidos (SEQ ID NO: 82) de la parte del locus κ de inmunoglobulina modificada por ingeniería genética que comprende dos segmentos Vk humanos (hVk1-39 y hVk3-20) y un segmento Jk humano; la secuencia de nucleótidos abarca la secuencia humana modificada por ingeniería genética y comprende 100 pares de bases de la secuencia de ratón endógena en ambos extremos 5' y 3'. La parte inferior de la **FIG. 22D** explica diferentes fuentes utilizadas para representar las diversas secuencias.
- 65 Las **FIGS. 23A-23B** muestran una estrategia general para la construcción de un vector diana para la modificación por ingeniería genética de un locus kappa de inmunoglobulina que comprende dos segmentos Vk humanos

(hVκ1-39 y hVκ3-20) y cinco segmentos Jk humanos, así como potenciadores de ratón y un brazo de IgκC. La **FIG. 23C** muestra la introducción de este vector diana en células ES y la generación de ratones heterocigotos con el mismo; mientras que la **FIG. 23D** muestra la delección del casete de selección en células ES utilizando la enzima FLP. En la mayoría de los aspectos, salvo que se indique de otra manera, las formas oscuras y las líneas continuas representan secuencias de ratón, y las formas blancas y las líneas dobles representan secuencias humanas. Los diagramas no se presentan a escala.

Las **FIGS. 24A-24D** muestran la secuencia de nucleótidos (SEQ ID NO: 83) del locus κ de inmunoglobulina modificado por ingeniería genética que comprende dos segmentos Vκ humanos (hVκ1-39 y hVκ3-20) y cinco segmentos Jk humanos; la secuencia de nucleótidos abarca la secuencia de ratón endógena modificada por ingeniería genética y de 100 pares de bases en ambos extremos 5' y 3'. La parte inferior de la **FIG. 24D** explica diferentes fuentes utilizadas para representar diversas secuencias.

La **FIG. 25A**, en el panel superior, muestra gráficas de contorno representativas de médula ósea teñida para linfocitos B y T (CD19+y CD3+, respectivamente) de un ratón de tipo silvestre (TS) y un ratón homocigoto para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J). El panel inferior muestra gráficas de contorno representativas de médula ósea seleccionada en CD19+ y teñida para ckit+ y CD43+ de un ratón de tipo silvestre (TS) y un ratón homocigoto para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J). En las gráficas de contorno del panel inferior se observan pro y pre linfocitos B.

La **FIG. 25B** muestra el número de pro (CD19+CD43+ckit+) y pre (CD19+CD43-ckit-) linfocitos B en médula ósea extraída de los fémures de ratones de tipo silvestre (TS) y homocigotos para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J).

La **FIG. 26A** muestra gráficas de contorno representativas de médula ósea seleccionada en singletes teñida para inmunoglobulina M (IgM) y B220 de un ratón de tipo silvestre (TS) y de un ratón homocigoto para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J). En cada una de las gráficas de contorno se indican los linfocitos B inmaduros, maduros y pro/pre.

La **FIG. 26B** muestra el número total de linfocitos B (CD19+), linfocitos B inmaduros (B220<sup>int</sup>IgM+) y linfocitos B maduros (B220<sup>hi</sup>IgM+) en médula ósea aislada de los fémures de ratones de tipo silvestre (TS) y ratones homocigotos para dos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J).

La **FIG. 27A** muestra gráficas de contorno representativas de médula ósea seleccionada en singletes teñidas para inmunoglobulina M (IgM) y B220 de un ratón de tipo silvestre (TS) y de un ratón homocigoto para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J). En cada una de las gráficas de contorno se indican los linfocitos B inmaduros, maduros y pro/pre.

La **FIG. 27B** muestra gráficas de contorno representativas de médula ósea seleccionada en linfocitos B inmaduros (B220<sup>int</sup>IgM+) y maduros (B220<sup>hi</sup>IgM+) teñidos para la expresión de Igκ e Igλ aislados de los fémures de un ratón de tipo silvestre (TS) y de un ratón homocigoto para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J).

La **FIG. 28A**, en el panel superior, muestra gráficas de contorno representativas de esplenocitos seleccionados en singletes y teñidos para linfocitos B y T (CD19+y CD3+, respectivamente) de un ratón de tipo silvestre (TS) y de un ratón homocigoto para dos segmentos génicos Vκ humanos y cinco Jk humanos (DLC-5J). El panel inferior muestra gráficas de contorno representativas de esplenocitos seleccionados en CD19+ y teñidos para inmunoglobulina D (IgD) e inmunoglobulina M (IgM) de un ratón de tipo silvestre (TS) y de uno homocigoto para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J). En cada una de las gráficas de contorno se indican los linfocitos B maduros (54 para TS, 56,9 para DLC-5J) y transicionales (23,6 para TS, 25,6 para DLC-5J).

La **FIG. 28B** muestra el número total de linfocitos B CD19+, linfocitos B transicionales (CD19+IgM<sup>hi</sup>IgD<sup>lo</sup>) y linfocitos B maduros (CD19+IgM<sup>ol</sup>IgD<sup>hi</sup>) en bazo extraído de ratones de tipo silvestre (TS) y homocigotos para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J).

La **FIG. 29A** muestra gráficas de contorno representativas de esplenocitos Igλ+ e Igκ+ seleccionados en CD19+ de un ratón de tipo silvestre (TS) y de un ratón homocigoto para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J).

La **FIG. 29B** muestra el número total de linfocitos B (CD19+), linfocitos B Igκ+ (CD19+Igκ+) y linfocitos B Igλ+ (CD19+Igλ+) en bazo extraído de ratones de tipo silvestre (TS) y homocigotos para dos segmentos génicos Vκ humanos y cinco segmentos génicos Jk humanos (DLC-5J).

La **FIG. 30A** muestra el desarrollo de linfocitos B periféricos en ratones homocigotos para dos segmentos génicos Vκ humanos y cinco Jk humanos. La primera gráfica de contorno (extremo izquierdo) muestra

esplenocitos CD93+ y B220+ seleccionados en CD19+ que indica linfocitos B inmaduros (39,6) y maduros (57,8). La segunda gráfica de contorno (parte superior central) muestra la expresión de IgM+ y CD23+ en linfocitos B inmaduros indicando poblaciones de linfocitos T1 (33,7; IgD-IgM+CD21<sup>lo</sup>CD23-), T2 (21,2; IgD<sup>hi</sup>IgM<sup>hi</sup>CD21<sup>mid</sup>CD23+) y T3 (29,1). La tercera gráfica de contorno (parte inferior central) muestra la expresión de linfocitos B maduros CD21+(CD35+) e IgM+ que indica una pequeña población (14,8) que da lugar a linfocitos B de la zona marginal y una segunda población (70,5) que da lugar a linfocitos B foliculares (FO). La cuarta gráfica de contorno (parte superior derecha) muestra la expresión de B220+ y CD23+ en linfocitos B maduros indicando poblaciones de linfocitos B de la zona marginal (90,5; ZM) y precursores de la zona marginal (7,3; IgM<sup>hi</sup>IgD<sup>hi</sup>CD21<sup>hi</sup>CD23+). La quinta gráfica de contorno (inferior derecha) muestra la expresión de IgD+ e IgM+ en linfocitos B maduros indicando poblaciones de linfocitos B FO-I (79,0; IgD<sup>hi</sup>IgM<sup>lo</sup>CD21<sup>mid</sup>CD23+) y FO-II (15,1; IgD<sup>hi</sup>IgM<sup>hi</sup>CD21<sup>mid</sup>CD23+). Dentro de cada región seleccionada se muestra el porcentaje de células.

La **FIG. 30B** muestra el desarrollo de linfocitos B periféricos en ratones de tipo silvestre. La primera gráfica de contorno (extremo izquierdo) muestra esplenocitos CD93+y B220+ seleccionados en CD19+ indicando linfocitos B inmaduros (31,1) y maduros (64,4). La segunda gráfica de contorno (central superior) muestra la expresión de IgM+ y CD23+ en linfocitos B inmaduros indicando poblaciones de linfocitos B T1 (28,5; IgD-IgM+CD21<sup>lo</sup>CD23-), T2 (28,7; IgD<sup>hi</sup>IgM<sup>hi</sup>CD21<sup>mid</sup>CD23+) y T3 (30,7). La tercera gráfica de contorno (central inferior) muestra la expresión de CD21+ (CD35+) e IgM+ de linfocitos B maduros indicando una pequeña población (7,69) que da lugar a linfocitos B de la zona marginal y una segunda población (78,5) que da lugar a linfocitos B foliculares (FO). La cuarta gráfica de contorno (derecha superior) muestra la expresión de B220+ y CD23+ en linfocitos B maduros indicando poblaciones de linfocitos B de la zona marginal (79,9; MZ) y precursores de la zona marginal (19,4; IgM<sup>hi</sup>IgD<sup>hi</sup>CD21<sup>hi</sup>CD23+). La quinta gráfica de contorno (derecha inferior) muestra la expresión de IgD+ e IgM+ en linfocitos B maduros indicando poblaciones de linfocitos B FO-I (83,6; IgD<sup>hi</sup>IgM<sup>lo</sup>CD21<sup>mid</sup>CD23+) y FO-II (13,1; IgD<sup>hi</sup>IgM<sup>hi</sup>CD21<sup>mid</sup>CD23+). Dentro de cada región seleccionada se muestra el porcentaje de células.

La **FIG. 31** muestra el número total de poblaciones de linfocitos B transicionales, de la zona marginal y folicular en bazo extraído de ratones de tipo silvestre (TS) y homocigotos para dos segmentos génicos Vk humanos y cinco Jk humanos (DLC-5J).

La **FIG. 32** muestra la expresión relativa de ARNm en médula ósea (eje y) de cadenas ligeras procedentes de Vk3-20 y de Vk1-39 en un ensayo de PCR cuantitativa utilizando sondas específicas para los segmentos génicos Vk3-20 o Vk1-39 en ratones homocigotos para un reemplazo de los segmentos génicos Vk y Jk endógenos con segmentos génicos Vk y Jk humanos (Hk) (cadena ligera humana de un ratón VELOCIMMUNE®), en ratones de tipo silvestre (TS), en ratones homocigotos para dos segmentos génicos Vk humanos y cinco segmentos génicos Jk humanos (DLC-5J) y en ratones homocigotos para dos segmentos génicos Vk humanos y un segmento génico Jk humano (DLC-1J). Las señales se normalizan con respecto a la expresión de Ck de ratón. ND: no detectado.

La **FIG. 33** muestra la expresión relativa de ARNm en bazo completo (eje y) de las cadenas ligeras procedentes de Vk3-20 y Vk1-39 en un ensayo de PCR cuantitativa utilizando sondas específicas para los segmentos génicos Vk3-20 o Vk1-39 en ratones homocigotos para un reemplazo de los segmentos génicos Vk y Jk endógenos con los segmentos génicos Vk y Jk humanos (Hk) (cadena ligera humana de un ratón VELOCIMMUNE®), en ratones de tipo silvestre (TS), en ratones homocigotos para dos segmentos génicos Vk humanos y cinco segmentos génicos Jk humanos (DLC-5J) y en ratones homocigotos para dos segmentos génicos Vk humanos y un segmento génico Jk humano (DLC-1J). Las señales se normalizan con respecto a la expresión de Ck de ratón. ND: no detectado

La **FIG. 34** muestra la secuencia y las propiedades (% de contenido GC, N, % de emparejamiento erróneo, Tf) de cebadores seleccionados por mutagénesis utilizados para modificar por ingeniería genética cuatro restos de histidina en las CDR3 de la secuencia de la cadena ligera Vk1-39 y Vk3-20 humana. En la Tabla indicada más adelante se incluyen las SEQ ID NO de estos cebadores utilizados en el Listado de Secuencias. F = cebador directo (*forward*), R = cebador inverso (*reverse*).

La **FIG. 35A** muestra la introducción de un vector diana que comprende dos segmentos de la cadena ligera Vk humana, cada uno de ellos sustituido con cuatro restos de histidina y cinco Jk humanos en células ES y la generación de ratones heterocigotos con el mismo; mientras que la **FIG. 35B** muestra la selección del casete de selección en células ES utilizando la enzima FLPO. En la mayoría de los aspectos, salvo que se indique de otra manera, las formas oscuras y las líneas continuas representan secuencias de ratón y las formas blancas y las líneas dobles representan secuencias humanas. Los diagramas no se presentan a escala.

La **FIG. 36** muestra la secuencia y las propiedades (% de contenido GC, N, % de emparejamiento erróneo, Tf) de cebadores seleccionados por mutagénesis utilizados para modificar por ingeniería genética tres restos de histidina en las CDR3 de la secuencia de cadena ligera Vk1-39 y Vk3-20. En la Tabla indicada más adelante se incluyen las SEQ ID NO de estos cebadores utilizados en el Listado de Secuencias. F = cebador directo, R = cebador inverso.

La **FIG. 37A** muestra la introducción de un vector diana que comprende dos segmentos de cadena ligera Vk humana, cada uno de ellos sustituido con tres restos de histidina y cinco Jk humanos en células ES y la generación de ratones heterocigotos con el mismo; mientras que la **FIG. 37B** muestra la delección del casete de selección en células ES utilizando la enzima FLPo. En la mayoría de los aspectos, salvo que se indique de otra manera, las formas oscuras y las líneas continuas representan secuencias de ratón, y las formas blancas y las líneas dobles representan secuencias humanas. Los diagramas no se presentan a escala.

La **FIG. 38A** muestra el alineamiento de la secuencia de aminoácidos codificada por la secuencia Vk3-20 de la línea germinal humana (secuencia inferior) con traducción de aminoácidos ejemplares de la secuencia variable de la cadena kappa ligera de IgM expresada en un ratón que comprende dos segmentos kappa V (Vk3-20 y Vk1-39), cada uno de ellos sustituido con 3 restos de histidina en la secuencia de la CDR3 (secuencia superior); el alineamiento muestra la secuencia variable de cadena kappa de IgM expresada en un ratón que conservó las tres sustituciones de histidina introducidas en la secuencia de la línea germinal. La **FIG. 38B** muestra el alineamiento de la secuencia de aminoácidos codificada por la secuencia Vk1-39 de la línea germinal humana (secuencia inferior en cada alineamiento) con traducción de aminoácidos ejemplares de la secuencia variable de la cadena kappa ligera de IgM expresada en un ratón que comprende dos segmentos kappa V (Vk3-20 y Vk1-39), cada uno de ellos sustituido con 3 restos de histidina en la secuencia de la CDR3 (secuencia superior en cada alineamiento); el alineamiento superior muestra la secuencia variable de la cadena kappa de IgM expresada en un ratón que conservó las tres modificaciones de histidina introducidas en la secuencia de la línea germinal, el alineamiento inferior muestra una secuencia variable de la cadena kappa de IgM expresada en un ratón que conservó dos de las tres modificaciones de histidina introducidas en la secuencia de la línea germinal. En algunos aspectos, la histidina introducida en la última posición de Vk puede perderse durante el reordenamiento V-J.

La **FIG. 39** ilustra la estructura genómica de ratones F2 modificados genéticamente que comprenden una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de la cadena pesada (MAID6032; "ratón UHC") y que adicionalmente comprende locus de cadena ligera modificados genéticamente que contienen dos segmentos génicos Vk humanos (por ejemplo, un segmento génico Vk1-39 humano y Vk3-20 humano) y cinco segmentos génicos Jk humanos (hJk1-5; DLC-5J) (MAID 1912HO).

La **FIG. 40A**, en el panel superior, muestra gráficas de contorno representativas de esplenocitos seleccionados en singletes y teñidos para linfocitos B y T (CD19+ y CD3+, respectivamente), de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y ratones homocigotos para una secuencia de ácido nucleico de la región variable de la cadena pesada reordenada en el locus de la cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de la cadena ligera (MAID 1912HO 6032 HO; DLC x UHC). El panel inferior muestra gráficas de contorno representativas de esplenocitos seleccionados en CD19+ y teñidos para inmunoglobulina D (IgD) e inmunoglobulina M (IgM) de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de la cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de la cadena ligera (MAID 1912HO 6032 HO; DLC x UHC). En cada una de las gráficas de contorno se indican los linfocitos B maduros e inmaduros.

La **FIG. 40B** muestra el número total de linfocitos B CD19+, linfocitos B maduros (CD 19+IgM<sup>lo</sup>Ig<sup>hi</sup>) y linfocitos B inmaduros (CD19+Ig-M<sup>hi</sup>IgD<sup>int</sup>) en bazo extraídos de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de la cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de la cadena ligera (MAID 1912HO 6032 HO; DLC x UHC).

La **FIG. 41A** muestra gráficas de contorno representativas de esplenocitos Ig $\lambda$ <sup>+</sup> e Ig $\kappa$ <sup>+</sup> seleccionados en CD19<sup>+</sup> de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y de ratones homocigotos para una secuencia de ácido nucleico de la región variable de la cadena pesada reordenada en los locus de la cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de la cadena ligera (MAID 1912HO 6032 HO; DLC x UHC).

La **FIG. 41B** muestra el número total de linfocitos B (CD19<sup>+</sup>), linfocitos B Ig $\kappa$ <sup>+</sup>(CD19+Ig $\kappa$ <sup>+</sup>) y linfocitos B Ig $\lambda$ <sup>+</sup> (CD19<sup>+</sup>Ig $\lambda$ <sup>+</sup>) en bazo extraídos de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y de ratones homocigotos para una secuencia de ácido nucleico de la región variable de la cadena pesada reordenada en los locus de la cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de la cadena ligera (MAID 1912HO 6032 HO; DLC x UHC).

La **FIG. 42** muestra el análisis de citometría de flujo de la expresión en superficie de IgM en linfocitos B en bazo extraídos de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de la cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de la cadena ligera (MAID 1912HO 6032 HO; DLC x UHC). Las células se tiñeron con un anticuerpo fluorescente (conjugado con PE-Cy7) contra IgM.

La **FIG. 43A** muestra el desarrollo de linfocitos B periféricos en ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y en ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de cadena ligera (MAID 1912HO 6032 HO; DLC x UHC). La primera gráfica de contorno (extremo izquierdo) muestra esplenocitos CD93<sup>+</sup> y B220<sup>+</sup> seleccionados en CD19<sup>+</sup> indicando linfocitos B inmaduros y maduros. La segunda gráfica de contorno (central) muestra la expresión de IgM<sup>+</sup> y CD23<sup>+</sup> en linfocitos B inmaduros indicando poblaciones de linfocitos B T1 (IgD-IgM<sup>+</sup>CD21<sup>lo</sup>CD23<sup>+</sup>), T2 (IgD<sup>hi</sup>IgM<sup>hi</sup>CD21<sup>mid</sup>CD23<sup>+</sup>) y T3. La tercera gráfica de contorno (derecha) muestra la expresión de CD21<sup>+</sup> (CD35<sup>+</sup>) e IgM<sup>+</sup> de linfocitos B maduros indicando las primeras poblaciones más pequeñas que dan lugar a linfocitos B de la zona marginal y segundas poblaciones más grandes que dan lugar a linfocitos B foliculares (FO). En cada región seleccionada se muestra el porcentaje de células.

La **FIG. 43B** muestra el desarrollo de linfocitos B periféricos en ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y en ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de la cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humanos en el locus de la cadena ligera (MAID 1912HO 6032 HO; DLC x UHC). La primera gráfica de contorno (izquierda) muestra la expresión de CD21<sup>+</sup>(CD35<sup>+</sup>) e IgM<sup>+</sup> de linfocitos B maduros, indicando una población pequeña que da lugar a linfocitos B de la zona marginal y una segunda población que da lugar a linfocitos B foliculares (FO). La segunda gráfica de contorno (central) muestra la expresión de B220<sup>+</sup> y CD23<sup>+</sup> en linfocitos B maduros indicando poblaciones de linfocitos B de la zona marginal (MZ) y precursores de la zona marginal (IgM<sup>hi</sup>IgD<sup>hi</sup>CD21<sup>hi</sup>CD23<sup>+</sup>). La tercera gráfica de contorno (derecha) muestra la expresión de IgD<sup>+</sup> e IgM<sup>+</sup> en linfocitos B maduros indicando poblaciones de linfocitos B FO-I (IgD<sup>hi</sup>IgM<sup>int</sup>CD21<sup>int</sup>CD23<sup>+</sup>) y FO-II (IgD<sup>hi</sup>IgM<sup>int</sup>CD21<sup>int</sup>CD23<sup>+</sup>). En cada región seleccionada se muestra el porcentaje de células.

La **FIG. 44A** muestra gráficas de contorno representativas de médula ósea teñida para linfocitos B y T (CD19<sup>+</sup> y CD3<sup>+</sup>, respectivamente) de un ratón de control modificado genéticamente (VI3; 1293HO 1460HO) y un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y dos segmentos génicos Vk humanos y cinco Jk humano en los locus de cadena ligera (MAID 1912HO 6032 HO; DLC x UHC).

La **FIG. 44B** muestra el porcentaje de linfocitos, el número total de células/fémur y el número de linfocitos B CD19<sup>+</sup> en médula ósea extraída de los fémures de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y en dos segmentos génicos Vk humanos y en cinco segmentos Jk humanos en los locus de cadena ligera (MAID 1912HO 6032 HO; DLC x UHC).

La **FIG. 45A** muestra gráficas de contorno representativas de médula ósea seleccionada en CD19<sup>+</sup> y teñida para ckit<sup>+</sup> y CD43<sup>+</sup> de un ratón de control modificado genéticamente (VI3; 1293HO 1460HO) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y en dos segmentos génicos Vk humanos y en cinco segmentos Jk humanos en los locus de cadena ligera (MAID 1912HO 6032 HO; DLC x UHC). En las gráficas de contorno se indican los pro y pre linfocitos B.

La **FIG. 45B** muestra el número de linfocitos Pre (CD19<sup>+</sup>CD43<sup>-</sup>ckit<sup>+</sup>) y Pro (CD19<sup>+</sup>CD43<sup>+</sup>ckit<sup>+</sup>) linfocitos B en médula ósea extraída de los fémures de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y en dos segmentos génicos Vk humanos y Jk humanos en los locus de cadena ligera (MAID 1912HO 6032HO; DLC x UHC).

La **FIG. 46A** muestra gráficas de contorno representativas de médula ósea seleccionada en singletes teñida para inmunoglobulina M (IgM) y B220 de un ratón de control modificado genéticamente (VI3; 1293HO 1460HO) y un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y en dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de cadena ligera (MAID 1912HO 6032 HO; DLC x UHC). En cada una de las gráficas de contorno se muestran linfocitos B inmaduros, maduros y pro/pre.

La **FIG. 46B** muestra el número total de células/fémur, linfocitos B inmaduros (B220<sup>int</sup>IgM<sup>+</sup>) y maduros (B220<sup>hi</sup>IgM<sup>+</sup>) en médula ósea aislada de los fémures de ratones de control modificados genéticamente (VI3; 1293HO 1460HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y en dos segmentos génicos Vk humanos y cinco Jk humanos en los locus de cadena ligera (MAID 1912HO 6032 HO; DLC x UHC).

La **FIG. 47** muestra gráficas de contorno representativas de médula ósea seleccionada en linfocitos B inmaduros (B220<sup>int</sup>IgM<sup>+</sup>) y maduros (B220<sup>hi</sup>IgM<sup>+</sup>) teñidos para la expresión de Igλ e Igκ aislados de fémures de un ratón de control modificado genéticamente (VI3; 1293HO 1460HO) y de un ratón homocigoto para una secuencia de ácido

nucleico de región variable de cadena pesada reordenada en los locus de cadena pesada y en dos segmentos génicos V<sub>k</sub> humanos y cinco J<sub>k</sub> humanos en los locus de cadena ligera (MAID 1912HO 6032 HO; DLC x UHC).

La **FIG. 48** muestra los niveles de mIgG específicas de antígeno en los sueros de ratón (HET de tipo silvestre o 1912HO 6031 (DLC homocigoto x UHC heterocigoto)) antes de la inmunización en las almohadillas plantares, 23 días después de una 1ª ronda de inmunización en las almohadilla plantares, 5 semanas después de la 1ª ronda de inmunización en las almohadillas plantares y después de una 2ª ronda de inmunización en las almohadillas plantares.

La **FIG. 49** ilustra la estructura genómica de ratones F1 modificados genéticamente que contienen secuencias de ácido nucleico de región variable de cadena pesada reordenadas en los locus de cadena ligera kappa (es decir, una secuencia VDJ de cadena pesada reordenada unida operativamente a una secuencia de ácido nucleico constante de cadena ligera kappa).

La **FIG. 50A** muestra, en el panel superior, gráficas de contorno representativas de médula ósea teñida para linfocitos B y T (CD19<sup>+</sup> y CD3<sup>+</sup>, respectivamente) de un ratón de tipo silvestre (TS) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa. En el panel inferior se muestran gráficas de contorno representativas de médula ósea seleccionada en CD19<sup>+</sup> y teñida para ckit<sup>+</sup> y CD43<sup>+</sup> de un ratón de tipo silvestre (TS) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en los locus de cadena ligera kappa. En las gráficas de contorno del panel inferior se indican los Pro y Pre linfocitos B.

La **FIG. 50B** muestra el número de Pro (CD19<sup>+</sup>CD43<sup>+</sup>ckit<sup>+</sup>) y Pre (CD19<sup>+</sup>CD43<sup>+</sup>ckit<sup>-</sup>) linfocitos B en médula ósea extraída de los fémures de ratones de tipo silvestre (TS) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa.

La **FIG. 51A** muestra gráficas de contorno representativas de médula ósea seleccionada en singletes teñida para inmunoglobulina M (IgM) y B220 de un ratón de tipo silvestre (TS) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de la cadena ligera kappa. En cada una de las gráficas de contorno se indican los linfocitos B inmaduros, maduros y pro/pre.

La **FIG. 51B** muestra el número total de linfocitos B (CD19<sup>+</sup>) y pro/pre B (IgM-B220<sup>+</sup>) en médula ósea aislada de los fémures de ratones de tipo silvestre (TS) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa.

La **FIG. 51C** muestra el número de linfocitos B inmaduros (B220<sup>int</sup>IgM<sup>+</sup>) y B maduros (B220<sup>hi</sup>IgM<sup>+</sup>) en médula ósea aislada de los fémures de ratones de tipo silvestre (TS) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa.

La **FIG. 52** muestra gráficas de contorno representativas de médula ósea seleccionada en linfocitos B inmaduros (B220<sup>int</sup>IgM<sup>+</sup>) y maduros (B220<sup>hi</sup>IgM<sup>+</sup>) teñidos para la expresión de Igλ e Igκ aislados de los fémures de ratones de tipo silvestre (TS) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa.

La **FIG. 53A**, muestra, en el panel superior, gráficas de contorno representativas de esplenocitos seleccionados en singletes y teñidos para linfocitos B y T (CD19<sup>+</sup> y CD3<sup>+</sup>, respectivamente) de un ratón de tipo silvestre (TS) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa. En el panel inferior se muestran gráficas de contorno representativas de esplenocitos seleccionados en CD19<sup>+</sup> y teñidos para inmunoglobulina D (IgD) e inmunoglobulina M (IgM) de un ratón de tipo silvestre (TS) y un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa. En cada una de las gráficas de contorno se muestran linfocitos B maduros (56,9 para TS, 43 para hV<sub>H3</sub>-23/D/J<sub>H4</sub> en kappa) y transicionales (26,8 para TS, 34 para hV<sub>H3</sub>-23/D/J<sub>H4</sub> en kappa).

La **FIG. 53B** muestra el número total de linfocitos B CD19<sup>+</sup>, linfocitos B maduros (CD19<sup>+</sup>IgM<sup>lo</sup>Ig<sup>hi</sup>) y linfocitos B transicionales (CD19<sup>+</sup>IgM<sup>hi</sup>IgD<sup>int</sup>) en bazos extraídos de ratones de tipo silvestre (TS) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa.

La **FIG. 54A** muestra gráficas de contorno representativas de esplenocitos Igλ<sup>+</sup> e Igκ<sup>+</sup> seleccionados en CD19<sup>+</sup> de un ratón de tipo silvestre (TS) y una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa.

La **FIG. 54B** muestra el número total de linfocitos B (CD19<sup>+</sup>), linfocitos B Igκ<sup>+</sup> (CD19<sup>+</sup>Igκ<sup>+</sup>) y linfocitos B Igλ<sup>+</sup> (CD19<sup>+</sup>Igλ<sup>+</sup>) en bazos extraídos de ratones de tipo silvestre (TS) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H3</sub>-23/D/J<sub>H4</sub>) en el locus de cadena ligera kappa.

La **FIG. 55** muestra el desarrollo de linfocitos B periféricos en el compartimiento esplénico de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada (hV<sub>H</sub>3-23/D/J<sub>H</sub>4) en el locus de cadena ligera kappa en comparación con ratones de tipo silvestre. La primera gráfica de contorno (izquierda) muestra esplenocitos CD93<sup>+</sup> y B220<sup>+</sup> seleccionados en CD19<sup>+</sup> indicando linfocitos B inmaduros y maduros. La segunda gráfica de contorno (central) muestra la expresión de IgM<sup>+</sup> y CD23<sup>+</sup> en linfocitos B inmaduros, indicando poblaciones T1, T2 y T3 de linfocitos B. La tercera gráfica de contorno (derecha) muestra la expresión de CD21<sup>+</sup>(CD35<sup>+</sup>) e IgM<sup>+</sup> de linfocitos B maduros indicando una primera pequeña población que da lugar a linfocitos B de la zona marginal y una segunda población más grande que da lugar a linfocitos B foliculares (FO). En cada región seleccionada se muestra el porcentaje de células.

La **FIG. 56** ilustra la estructura genómica de ratones F2 homocigotos modificados genéticamente, de una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa (MAID 6079HO; "ratón UHC en cadena kappa" homocigoto) y de ratones homocigotos de una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (MAID 1994HO; ratón con cadena kappa en cadena pesada ("KoH", *kappa on heavy*)).

La **FIG. 57A** muestra, en el panel superior, gráficas de contorno representativas de médula ósea teñida para linfocitos B y T (CD19<sup>+</sup> y CD3<sup>+</sup>, respectivamente) de un ratón VELOCIMMUNE® (VI3) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigoto para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada. El panel inferior muestra gráficas de contorno representativas de médula ósea seleccionada en CD19<sup>+</sup> y teñida para ckit<sup>+</sup> y CD43<sup>+</sup> de un ratón VELOCIMMUNE® (VI3) y un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigoto para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada. En las gráficas de contorno del panel inferior se indican linfocitos Pro y Pre B.

La **FIG. 57B** muestra el número total de linfocitos Pro B (CD19<sup>+</sup>) y los números de linfocitos Pro (CD19<sup>+</sup>CD43<sup>+</sup>ckit<sup>+</sup>) y Pre (CD19<sup>+</sup>CD43<sup>+</sup>ckit<sup>-</sup>) B en médula ósea extraída de los fémures de ratones VELOCIMMUNE® (1242HO 1640HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (1994HO 6079HO). Los números se presentan como números absolutos de linfocitos por fémur y porcentaje celular.

La **FIG. 58A** muestra gráficas de contorno representativas de médula ósea seleccionada en singletes teñida para inmunoglobulina M (IgM) y B220 de un ratón VELOCIMMUNE® (1242HO 1640HO) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de la cadena ligera kappa y homocigoto para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (1994HO 6079HO). En cada una de las gráficas de contorno se muestran los linfocitos B inmaduros, maduros y pro/pre.

La **FIG. 58B** muestra el número de linfocitos B inmaduros (B220<sup>int</sup>IgM<sup>+</sup>) y maduros (B220<sup>hi</sup>IgM<sup>+</sup>) en médula ósea aislada de los fémures de ratones VELOCIMMUNE® (1242HO 1640HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigotos para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (1994HO 6079HO). Los números se presentan tanto como en números absolutos de linfocitos por fémur como en porcentaje celular.

La **FIG. 59** muestra gráficas de contorno representativas de médula ósea seleccionada en linfocitos B inmaduros (B220<sup>int</sup>IgM<sup>+</sup>) y maduros (B220<sup>hi</sup>IgM<sup>+</sup>) teñidos para la expresión de Igλ e Igκ aislados de los fémures de ratones VELOCIMMUNE® (1242HO 1640HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigotos para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un el locus de cadena pesada (1994HO 6079HO).

La **FIG. 60A** muestra gráficas de contorno representativas de esplenocitos seleccionados en CD19<sup>+</sup> y teñidos para inmunoglobulina D (IgD) e inmunoglobulina M (IgM) de un ratón VELOCIMMUNE® (VI3; 1242HO 1640HO) y de un ratón homocigoto para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigoto para una secuencia de ácido nucleico de la región variable de cadena ligera kappa en un locus de cadena pesada (1994HO 6079HO). En cada una de las gráficas de contorno se indican linfocitos B maduros y transicionales/inmaduros.

La **FIG. 60B** muestra el número total de linfocitos B CD19<sup>+</sup>, linfocitos B maduros (CD19<sup>+</sup>IgM<sup>lo</sup>Ig<sup>hi</sup>) y linfocitos B transicionales (CD19<sup>+</sup>Ig-M<sup>hi</sup>IgD<sup>lo</sup>) en bazos extraídos de ratones VELOCIMMUNE® (1242HO 1640HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigotos para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (1994HO 6079HO).

La **FIG. 61** muestra el número total de linfocitos B (CD19<sup>+</sup>), linfocitos B Igk<sup>+</sup> (CD19<sup>+</sup>Igk<sup>+</sup>) y linfocitos B Igλ<sup>+</sup> (CD19<sup>+</sup>Igλ<sup>+</sup>) en bazo extraídos de ratones VELOCIMMUNE® (1242HO 1640HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigotos para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (1994HO 6079HO). Los números se presentan tanto como en números absolutos de linfocitos como en porcentaje celular de linfocitos.

La **FIG. 62** muestra el desarrollo de linfocitos B periféricos en el compartimiento esplénico de ratones VELOCIMMUNE® (1242HO 1640HO) y de ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa y homocigotos para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (1994HO) 6079HO). La gráfica de contorno superior muestra esplenocitos CD93<sup>+</sup> y B220<sup>+</sup> seleccionados en CD19<sup>+</sup> indicando los linfocitos B inmaduros y maduros. La gráfica de contorno inferior muestra la expresión de IgM<sup>+</sup> y CD23<sup>+</sup> en linfocitos B inmaduros indicando poblaciones T1, T2 y T3 de linfocitos B. En cada región seleccionada se muestra el porcentaje de linfocitos.

### Descripción detallada

La presente invención no se limita a métodos particulares y a las condiciones experimentales descritas, ya que dichos métodos y condiciones pueden variar. También debe entenderse que la terminología que se utiliza en el presente documento tiene la finalidad de describir aspectos particulares únicamente y no pretende ser limitante, dado que el alcance de la presente invención está definido por las reivindicaciones.

A menos que se defina de otra manera, todos los términos y frases que se utilizan en el presente documento incluyen los significados que los términos y frases se entienden en la técnica, a menos que se indique claramente lo contrario o sea claramente obvio en el contexto en el que se utilice el término o la frase. Aunque en la práctica o ensayo de la presente invención puede utilizarse cualquiera de los métodos y materiales similares o equivalentes a los descritos en el presente documento, ahora se describen métodos y materiales particulares.

El término "anticuerpo", como se usa en el presente documento, incluye moléculas de inmunoglobulina que comprenden cuatro cadenas polipeptídicas, dos cadenas pesadas (H, *heavy*) y dos cadenas ligeras (L, *light*) interconectadas por enlaces disulfuro. Cada cadena pesada comprende un dominio variable de cadena pesada y una región constante de cadena pesada (C<sub>H</sub>). La región constante de cadena pesada comprende tres dominios, C<sub>H1</sub>, C<sub>H2</sub> y C<sub>H3</sub>. Cada cadena ligera comprende un dominio variable de cadena ligera y una región constante de cadena ligera (C<sub>L</sub>). Los dominios variables de cadena ligera y pesada pueden subdividirse además en regiones de hipervariabilidad, denominadas regiones determinantes de la complementariedad (CDR, *complementarity determining regions*), intercaladas con regiones que son más conservadas, denominadas regiones estructurales (FR, del inglés *framework*). Cada dominio variable de cadena pesada y ligera comprende tres CDR y cuatro FR, que se disponen desde el extremo amino al extremo carboxilo en el siguiente orden: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4 (las CDR de cadena pesada pueden abreviarse como HCDR1, HCDR2 y HCDR3; las CDR de la cadena ligera pueden abreviarse como LCDR1, LCDR2 y LCDR3). La expresión anticuerpo de "alta afinidad" se refiere a un anticuerpo que tiene una K<sub>D</sub> con respecto a su epítipo diana de aproximadamente 10<sup>-9</sup> M o inferior (por ejemplo, de aproximadamente 1 x 10<sup>-9</sup> M, 1 x 10<sup>-10</sup> M, 1 x 10<sup>-11</sup> M, o de aproximadamente 1 x 10<sup>-12</sup> M). En un aspecto, la K<sub>D</sub> se mide mediante resonancia de plasmón superficial, por ejemplo, BIACORE™; en otro aspecto, la K<sub>D</sub> se mide por ELISA.

La frase "anticuerpo biespecífico" incluye un anticuerpo que es capaz de unirse selectivamente a dos o más epítomos. Generalmente, los anticuerpos biespecíficos comprenden dos cadenas pesadas no idénticas, uniéndose cada cadena pesada específicamente con un epítipo diferente bien en dos moléculas diferentes (por ejemplo, diferentes epítomos en dos inmunógenos diferentes) o en la misma molécula (por ejemplo, diferentes epítomos en el mismo inmunógeno). Si un anticuerpo biespecífico es capaz de unirse selectivamente a dos epítomos diferentes (un primer epítipo y un segundo epítipo), la afinidad de la primera cadena pesada por el primer epítipo generalmente será de al menos uno a dos o tres o cuatro o más órdenes de magnitud menor que la afinidad de la primera cadena pesada por el segundo epítipo, y viceversa. Los epítomos que se unen específicamente por el anticuerpo biespecífico pueden ser en la misma diana o en una diana diferente (por ejemplo, en la misma proteína o en una proteína diferente). Como ejemplos de anticuerpos biespecíficos se incluyen aquellos con una primera cadena pesada específica para un antígeno tumoral y con una segunda cadena pesada específica para un marcador citotóxico, por ejemplo, un receptor Fc (por ejemplo, FcγRI, FcγRII, FcγRIII, etc.) o un marcador de linfocitos T (por ejemplo, CD3, CD28, etc.). Además, el segundo dominio variable de cadena pesada puede sustituirse con un dominio variable de cadena pesada que tenga una especificidad deseada diferente. Por ejemplo, un anticuerpo biespecífico con una primera cadena pesada específica para un antígeno tumoral y una segunda cadena pesada específica para una toxina pueden emparejarse para suministrar una toxina (por ejemplo, saporina, alcaloides de la vinca, etc.) a una célula tumoral. Otros ejemplos de anticuerpos biespecíficos incluyen aquellos con una primera cadena pesada específica para un receptor activador (por ejemplo, receptor de linfocitos B, FcγRI, FcγRIIA, FcγRIIIA, FcαRI, receptor de linfocitos T etc.) y una segunda cadena pesada específica para un receptor inhibidor

(por ejemplo, FcγRIIB, CD5, CD22, CD72, CD300a, etc.). Dichos anticuerpos biespecíficos pueden construirse para afecciones terapéuticas asociadas a la activación celular (por ejemplo, alergia y asma). Los anticuerpos biespecíficos pueden crearse, por ejemplo, combinando cadenas pesadas que reconozcan diferentes epítopos del mismo inmunógeno. Por ejemplo, las secuencias de ácido nucleico que codifican secuencias variables de cadena pesada que reconocen diferentes epítopos del mismo inmunógeno pueden fusionarse con secuencias de ácido nucleico que codifican las mismas, o diferentes, regiones constantes de cadena pesada, y dichas secuencias pueden expresarse en una célula que expresa una cadena ligera de inmunoglobulina. Un anticuerpo biespecífico típico tiene dos cadenas pesadas, teniendo cada una de ellas tres CDR de cadena pesada, seguidas (del extremo N al extremo C), de un dominio C<sub>H</sub>1, una bisagra, un dominio C<sub>H</sub>2 y un dominio C<sub>H</sub>3, y una cadena ligera de inmunoglobulina que no confiere especificidad de unión con el epítipo sino que puede asociarse con cada cadena pesada, o que puede asociarse con cada cadena pesada y que puede unirse a uno o más de los epítopos unidos por las regiones de unión a epítipo de cadena pesada, o que puede asociarse con cada cadena pesada y permitir la unión de una, o de las dos cadenas pesadas, con uno o ambos epítopos. De manera similar, la expresión “anticuerpo triespecífico” incluye un anticuerpo que es capaz de unirse selectivamente a tres o más epítopos.

El término “célula” incluye cualquier célula que sea adecuada para la expresión de una secuencia de ácido nucleico recombinante. Las células incluyen células procariotas y eucariotas (unicelulares o multicelulares), bacterianas (por ejemplo cepas de *E. coli*, *Bacillus* spp., *Streptomyces* spp., etc.), células de micobacterias, células fúngicas, células de levadura (por ejemplo, *S. cerevisiae*, *S. pombe*, *P. pastoris*, *P. methanolica*, etc.), células vegetales, células de insecto (por ejemplo SF-9, SF-21, células de insecto infectadas con baculovirus, *Trichoplusia ni*, etc.), células de animales no humanos, células humanas o fusiones celulares, tales como, por ejemplo, hibridomas o cuadromas. En algunos aspectos, la célula es una célula de ser humano, de mono, simio, hámster, rata o ratón. En algunos aspectos, la célula es eucariota y se selecciona de las siguientes células: CHO (por ejemplo, CHO K1, DXB-11 CHO, Veggie-CHO), COS (por ejemplo, COS-7), célula retiniana, Vero, CV1, de riñón (por ejemplo, HEK293, 293 EBNA, MSR 293, MDCK, HaK, BHK), HeLa, HepG2, WI38, MRC 5, Colo205, HB 8065, HL-60, (por ejemplo, BHK21), Jurkat, Daudi, A431 (epidérmica), CV-1, U937, 3T3, célula L, célula C127, SP2/0, NS-0, MMT 060562, célula de Sertoli, célula BRL 3A, célula HT1080, célula de mieloma, célula tumoral y una línea celular procedente de una células anteriormente mencionada. En algunos aspectos, la célula comprende uno o más genes víricos, por ejemplo, una célula retiniana que expresa un gen vírico (por ejemplo, una célula PER.C6™).

La frase “región determinante de la complementariedad” o el término “CDR”, incluye una secuencia de aminoácidos codificada por una secuencia de ácido nucleico de genes de inmunoglobulina de un organismo que normalmente (es decir, en un animal de tipo silvestre) aparece entre dos regiones estructurales en una región variable de una cadena ligera o pesada de una molécula de inmunoglobulina (por ejemplo, un anticuerpo o un receptor de linfocito T). Una CDR puede estar codificada por ejemplo, por una secuencia de línea germinal o por una secuencia reordenada o no reordenada y, por ejemplo, por un linfocito B virgen o maduro o por un linfocito T. Una CDR puede estar somáticamente mutada (por ejemplo, variar de una secuencia codificada en una línea germinal animal), humanizada y/o modificada con sustituciones, adiciones o deleciones de aminoácidos. En algunas circunstancias (por ejemplo, para una CDR3), las CDR pueden estar codificadas por dos o más secuencias (por ejemplo, secuencias de la línea germinal) que no son contiguas (por ejemplo, en una secuencia de ácido nucleico no reordenada), pero que son contiguas en una secuencia de ácido nucleico de linfocito B, por ejemplo, como resultado del corte y empalme o conexión de las secuencias (por ejemplo, recombinación V-D-J para formar una CDR3 de cadena pesada).

El término “conservativa”, cuando se utiliza para describir una sustitución de aminoácido conservativa, incluye la sustitución de un resto de aminoácido por otro resto de aminoácido que tiene un grupo R en la cadena lateral con propiedades químicas (por ejemplo, carga o hidrofobicidad) similares. En general, una sustitución de aminoácidos conservativa no cambiará sustancialmente las propiedades funcionales de interés de una proteína, por ejemplo, la capacidad de una región variable para unirse específicamente a un epítipo diana con una afinidad deseada. Como ejemplos de grupos de aminoácidos que tienen cadenas laterales con propiedades químicas similares se incluyen cadenas laterales alifáticas, tales como glicina, alanina, valina, leucina e isoleucina; cadenas laterales de hidroxilo alifáticas, tales como, serina y treonina; cadenas laterales que contienen amida, tales como, asparagina y glutamina; cadenas laterales aromáticas, tales como, fenilalanina, tirosina y triptófano; cadenas laterales básicas, tales como, lisina, arginina e histidina; cadenas laterales ácidas, tales como ácido aspártico y ácido glutámico; y cadenas laterales que contienen azufre, tales como, cisteína y metionina. Como grupos de sustitución de aminoácidos conservativa se incluyen, por ejemplo, valina/leucina/isoleucina, fenilalanina/tirosina, lisina/arginina, alanina/valina, glutamato/aspartato y asparagina/glutamina. En algunos aspectos, una sustitución de aminoácidos conservativa puede ser una sustitución de cualquier resto nativo en una proteína con alanina, tal y como se utiliza, por ejemplo, en la mutagénesis de barrido con alanina. En algunos aspectos, se realiza una sustitución conservativa que tiene un valor positivo en la matriz de probabilidad logarítmica PAM250 desvelada en Gonnet *et al.* (1992) Exhaustive Matching of the Entire Protein Sequence Database, Science 256: 1443-45. En algunos aspectos, la sustitución es una sustitución moderadamente conservativa en la que la sustitución tiene un valor no negativo en la matriz de probabilidad logarítmica PAM250.

En algunos aspectos, las posiciones de los restos en una cadena pesada o ligera de inmunoglobulina, difieren en una o más sustituciones de aminoácidos conservativas. En algunos aspectos, las posiciones de los restos en una cadena ligera de inmunoglobulina o fragmento funcional de la misma (por ejemplo, un fragmento que permita la

expresión y la secreción, por ejemplo, de un linfocito B) no son idénticas a las de una cadena ligera cuya secuencia de aminoácidos se enumere en el presente documento, aunque difieren en una o más sustituciones de aminoácidos conservativas.

5 La frase “proteína de unión a epítipo” incluye una proteína que tiene al menos una CDR que es capaz de reconocer selectivamente un epítipo, por ejemplo, que es capaz de unirse a un epítipo con una  $K_D$  que es aproximadamente de un micromolar o inferior (por ejemplo, una  $K_D$  que es de aproximadamente  $1 \times 10^{-6}$  M,  $1 \times 10^{-7}$  M,  $1 \times 10^{-9}$  M,  $1 \times 10^{-9}$  M,  $1 \times 10^{-10}$  M,  $1 \times 10^{-11}$  M, o de aproximadamente  $1 \times 10^{-12}$  M). Las proteínas de unión a epítipo terapéuticas (por ejemplo, anticuerpos terapéuticos) frecuentemente requieren una  $K_D$  que esté en el intervalo nanomolar o  
10 picomolar.

La frase “fragmento funcional” incluye fragmentos de proteínas de unión a epítipo que pueden expresarse, secretarse y unirse específicamente a un epítipo con una  $K_D$  en el intervalo micromolar, nanomolar o picomolar. El reconocimiento específico incluye tener una  $K_D$  que esté al menos en el intervalo micromolar, nanomolar o  
15 picomolar.

La expresión “línea germinal”, en referencia a una secuencia de ácido nucleico de inmunoglobulina, incluye una secuencia de ácido nucleico que pueda transferirse a la descendencia.

20 La expresión “cadena pesada” o “cadena pesada de inmunoglobulina” incluye una secuencia de cadena pesada de inmunoglobulina, incluyendo una secuencia de región constante de cadena pesada de inmunoglobulina, de cualquier organismo. Los dominios variables de cadena pesada incluyen tres CDR de cadena pesada, y cuatro regiones FR, a menos que se especifique de otra manera. Los fragmentos de cadenas pesadas incluyen las CDR, CDR y FR y combinaciones de las mismas. Una cadena pesada típica tiene, después del dominio variable (desde el extremo N al extremo C), un dominio  $C_H1$ , una bisagra, un dominio  $C_H2$  y un dominio  $C_H3$ . Un fragmento funcional de una cadena  
25 pesada incluye un fragmento que es capaz de reconocer específicamente un epítipo (por ejemplo, reconocer el epítipo con una  $K_D$  en el intervalo micromolar, nanomolar o picomolar), que es capaz de expresarse y secretarse de una célula, y que comprende al menos una CDR. Un dominio variable de cadena pesada está codificado por una secuencia génica de región variable, que generalmente comprende segmentos  $V_H$ ,  $D_H$  y  $J_H$  procedentes de un repertorio de segmentos  $V_H$ ,  $D_H$  y  $J_H$  presentes en la línea germinal. Las secuencias, localizaciones y nomenclatura de los segmentos de cadena pesada para V,  $D_H$  y J de varios organismos pueden encontrarse en la base de datos de IMGT, [www.imgt.org](http://www.imgt.org).

35 Cuando el término “identidad” se usa en relación con una secuencia, incluye identidad determinada por diversos algoritmos diferentes conocidos en la técnica que pueden utilizarse para medir la identidad de las secuencias de nucleótidos y/o de aminoácidos. En algunos aspectos descritos en el presente documento, las identidades se determinan utilizando un alineamiento ClustalW v. 1.83 (lento) que emplea una penalización por apertura de hueco de 10,0, una penalización por hueco extendido de 0,1 y que utiliza una matriz de similitud de Gonnet (MACVECTOR™ 10.0.2, MacVector Inc., 2008). La longitud de las secuencias comparada con respecto a la  
40 identidad de secuencias, dependerá de las secuencias particulares, pero en el caso de un dominio constante de cadena ligera, la longitud debe contener secuencias de suficiente longitud para doblarse en un dominio constante de cadena ligera que sea capaz de auto asociarse para formar un dominio constante de cadena ligera canónico, por ejemplo, capaz de formar dos láminas beta que comprenden cadenas beta y capaz de interactuar con al menos un dominio  $C_H1$  de un ser humano o un ratón. En el caso de un dominio  $C_H1$ , la longitud de la secuencia debe contener  
45 secuencias de suficiente longitud para plegarse en un dominio  $C_H1$  que sea capaz de formar dos láminas beta que comprendan cadenas beta y que sea capaz de interactuar al menos con un dominio constante de cadena ligera de ratón o de ser humano.

La frase “molécula de inmunoglobulina” incluye dos cadenas pesadas de inmunoglobulina y dos cadenas ligeras de  
50 inmunoglobulina. Las cadenas pesadas y ligeras pueden ser idénticas o diferentes.

La frase “cadena ligera” incluye una secuencia de cadena ligera de inmunoglobulina de cualquier organismo, y a menos que se especifique de otra manera, incluye cadenas ligeras, lambda y kappa, humanas y una VpreB, así como cadenas ligeras sustitutivas. Los dominios variables de cadena ligera, incluyen típicamente tres CDR de  
55 cadena ligera, y cuatro regiones estructurales (FR), a menos que se especifique de otra manera. Generalmente, una cadena ligera de longitud completa incluye, desde el extremo amino al extremo carboxilo, un dominio variable que incluye FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 y una región constante de cadena ligera. Un dominio variable de cadena ligera está codificado por una secuencia génica de región variable de cadena ligera, que generalmente comprende segmentos  $V_L$  y  $J_L$ , procedentes de un repertorio de segmentos V y J presentes en la línea germinal.  
60 Las secuencias, localizaciones y nomenclatura para los segmentos de cadena ligera V y J de diversos organismos, pueden encontrarse en la base de datos de IMGT, [www.imgt.org](http://www.imgt.org). Las cadenas ligeras incluyen, por ejemplo, las que no se unen selectivamente ni a un primer ni a un segundo epítipo unidos selectivamente por la proteína de unión a epítipo en la que aparece. Las cadenas ligeras también incluyen las que se unen y reconocen, o ayudan a la cadena pesada con la unión y reconocimiento, uno o más epítopos selectivamente unidos mediante la proteína de  
65 unión a epítipo en la que aparecen. Las cadenas ligeras comunes o universales incluyen las procedentes de un gen humano Vk1-39Jk5 o de un gen humano Vk3-20Jk1, e incluyen versiones somáticamente mutadas (por ejemplo,

maduradas por afinidad) de las mismas. Las dobles cadenas ligeras (DLC, *dual light chains*) incluyen las procedentes de un locus de cadena ligera que comprende, a lo sumo, dos segmentos  $V_k$  humanos, por ejemplo, un segmento génico  $V_{k1-39}$  humano y un segmento génico  $V_{k3-20}$  humano, e incluyen versiones somáticamente mutadas (por ejemplo, maduradas por afinidad) de las mismas.

5 La frase “somáticamente hipermutada” incluye la referencia a una secuencia de ácido nucleico de un linfocito B que ha experimentado un intercambio de clase, en la que la secuencia de ácido nucleico de una región variable de inmunoglobulina (por ejemplo una secuencia de nucleótidos que codifica un dominio variable de cadena pesada o que incluye una secuencia CDR o FR de cadena pesada) en el linfocito B con intercambio de clase, no es idéntica a la secuencia de ácido nucleico en el linfocito B antes del intercambio de clase, tal como, por ejemplo, una diferencia en una secuencia de ácido nucleico de una CDR o estructural entre un linfocito B que no ha experimentado un intercambio de clase y un linfocito B que ha experimentado un intercambio de clase. “Somáticamente mutada” incluye la referencia a secuencias de ácido nucleico de linfocitos B madurados por afinidad que no son idénticas a las regiones de región variable de inmunoglobulina correspondientes en linfocitos B que no están madurados por afinidad (es decir, secuencias en el genoma de las células de línea germinal). La frase “somáticamente mutada” también incluye la referencia a una secuencia de ácido nucleico de región variable de inmunoglobulina de un linfocito B después de la exposición del linfocito B a un epítipo de interés, en el que la secuencia de ácido nucleico difiere de la secuencia de ácido nucleico correspondiente antes de exponer al linfocito B al epítipo de interés. La frase “somáticamente mutada” se refiere a secuencias de anticuerpos que se han generado en un animal, por ejemplo, un ratón que tiene secuencias de ácido nucleico de región variable de inmunoglobulina humana, en respuesta a una exposición a un inmunógeno, y que es el resultado de los procesos de selección extrínsecamente operativos en dicho animal.

La expresión “no reordenada”, con referencia a una secuencia de ácido nucleico, incluye secuencias de ácido nucleico que existen en la línea germinal de una célula animal.

La frase “dominio variable” incluye una secuencia de aminoácidos de una cadena ligera o pesada de inmunoglobulina (modificada según se desee) que comprende las siguientes regiones de aminoácidos, en secuencia desde el extremo N al extremo C (a menos que se indique de otra manera): FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4.

La expresión “unido(a) operativamente” se refiere a una relación en la que los componentes funcionan unidos operativamente en su manera deseada. En un caso, una secuencia de ácido nucleico que codifica una proteína puede estar unida operativamente a secuencias reguladoras (por ejemplo, una secuencia promotora, potenciadora, silenciadora, etc.) para conservar la regulación transcripcional apropiada. En un caso, una secuencia de ácido nucleico de una región variable de inmunoglobulina (o segmentos V(D)J) puede estar unida operativamente a una secuencia de ácido nucleico de una región constante de inmunoglobulina para permitir que se produzca la recombinación adecuada entre las secuencias en una secuencia de cadena ligera o pesada de inmunoglobulina.

“Funcional”, como se usa en el presente documento, por ejemplo, en referencia a un polipéptido funcional, incluye un polipéptido que conserva al menos una actividad biológica normalmente asociada a la proteína nativa. En otro caso, un segmento génico de inmunoglobulina funcional puede incluir un segmento génico variable que pueda reordenarse de manera productiva para generar una secuencia génica de inmunoglobulina reordenada.

Un “pH neutro” incluye un pH entre aproximadamente 7,0 y aproximadamente 8,0, por ejemplo, un pH entre aproximadamente 7,0 y aproximadamente 7,4, por ejemplo, entre aproximadamente 7,2 y aproximadamente 7,4, por ejemplo, un pH fisiológico. Un “pH ácido” incluye un pH de 6,0 o inferior, por ejemplo, un pH entre aproximadamente 5,0 y aproximadamente 6,0, un pH entre aproximadamente 5,75 y aproximadamente 6,0, por ejemplo, un pH de los compartimentos endosómicos o lisosómicos.

La expresión “variante polimórfica” como se usa en el presente documento, incluye una secuencia en la que uno o más nucleótidos o aminoácidos se han sustituido por nucleótidos o aminoácidos diferentes en comparación con la secuencia determinada. Los alelos polimórficos de los segmentos génicos variables de cadena pesada (genes  $V_H$ ) de inmunoglobulina humana, han sido en gran parte el resultado de la inserción/delección de segmentos génicos y de diferencias de un solo nucleótido dentro de las regiones codificantes, teniendo ambas el potencial de tener consecuencias funcionales en la molécula de inmunoglobulina. En la técnica se conocen bien ejemplos de alelos polimórficos comunes de los genes  $V_H$  de inmunoglobulina humana (véase, por ejemplo, el documento US 13/653.456).

Cuando las expresiones “sustancial” o “sustancialmente todos”, se utilizan con referencia a una cantidad de segmentos génicos (por ejemplo, “sustancialmente todos” los segmentos génicos V, D o J), incluyen segmentos génicos tanto funcionales como no funcionales e incluyen, en diversos aspectos, por ejemplo, el 80 % o más, el 85 % o más, el 90 % o más, el 95 % o más, el 96 % o más, el 97 % o más, el 98 % o más, o el 99 % o más, de todos los segmentos génicos V, D, o J. En diversos aspectos, “sustancialmente todos” los segmentos génicos incluyen, por ejemplo, al menos el 95 %, 96 %, 97 %, 98 % o 99 % de los segmentos génicos funcionales (es decir, no pseudogén).

**Animales no humanos que comprenden una secuencia génica de región variable de cadena ligera reordenada y opcionalmente un repertorio limitado de segmentos génicos variables de cadena ligera no reordenada**

5 Aunque se han desarrollado diversos anticuerpos biespecíficos con dobles propiedades de unión a antígeno, en los anticuerpos biespecíficos convencionales, la especificidad y afinidad de las regiones variables de cadena ligera o cadena pesada ha tenido que sacrificarse hasta cierto grado porque, en los anticuerpos biespecíficos convencionales, cualquiera de una región variable de cadena ligera o de cadena pesada solo contribuye a la unión  
10 de cada uno de los determinantes antigénicos, mientras que en los anticuerpos regulares, las regiones variables tanto de cadena pesada como de cadena ligera, pueden contribuir a la unión con el mismo determinante antigénico.

Por lo tanto, la generación de regiones variables de cadena ligera que tienen la capacidad de unirse a un antígeno independientemente de una región variable de cadena pesada, puede ser útil para crear dominios variables de  
15 cadena ligera ( $V_L$ ) para su uso en moléculas de unión a antígeno (por ejemplo, moléculas de unión biespecíficas que comprenden una región constante de cadena pesada (por ejemplo, seleccionada de un  $C_H1$ , una bisagra, un  $C_H2$ , un  $C_H3$  y una combinación de los mismos) fusionada con  $V_L$ ), particularmente aquellas que no comprenden un dominio variable de cadena pesada, incluyendo los heterodímeros que tienen la misma región constante de cadena pesada o similar pero las  $V_L$  con diferentes especificidades y/o afinidades.

Una estrategia para producir dichos dominios variables de cadena ligera que pueden unirse a un antígeno independientemente de una región variable de cadena pesada, es aplicar una presión selectiva sobre las secuencias de nucleótidos que codifican una región variable o dominio de una cadena ligera ( $V_L$ ) para generar las CDR3 de  
20 cadena ligera con un repertorio de unión antigénica más diverso. Como se desvela en el presente documento, esto puede realizarse generando genéticamente un ratón o una rata que contenga, en su genoma, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. Dado que la secuencia de cadena pesada está restringida a secuencias comunes o universales (es decir, la misma o una muy similar) en estos ratones y ratas, las secuencias de nucleótidos (es decir, genes) de región variable de cadena ligera se verán forzadas a crear las CDR3 de cadena ligera con propiedades de unión antigénicas más diversas y eficaces, que  
25 pueden unirse a un determinante antigénico independientemente de las regiones variables de cadena pesada. Así mismo, como se desvela en el presente documento, el reemplazo exacto de los segmentos génicos de región variable de la línea germinal (por ejemplo, por direccionamiento génico mediado por recombinación homóloga) permite la creación de ratones y ratas que tienen locus de inmunoglobulina parcialmente humanos. Dado que normalmente los locus de inmunoglobulina parcialmente humanos se reordenan, hipermutan y mutan somáticamente (por ejemplo, intercambio de clase), los locus de inmunoglobulina parcialmente humanos generan en el animal anticuerpos que comprenden regiones variables humanas. Estos ratones o ratas exhiben un sistema  
30 inmunario humoral que es sustancialmente similar al de los ratones o ratas de tipo silvestre, y presentan poblaciones de células normales y estructuras de órganos linfoides normales, incluso cuando los animales carecen de un repertorio completo de segmentos génicos de la región variable humana. La inmunización de estos animales (por ejemplo, ratones) da como resultado fuertes respuestas humorales que presentan una amplia diversidad de uso de los segmentos génicos variables. Las secuencias de nucleótidos que codifican las regiones variables pueden identificarse y clonarse, después fusionarse (por ejemplo, en un sistema *in vitro*) con cualquiera de las secuencias de elección, por ejemplo, con cualquier isotipo de inmunoglobulina adecuado para un uso particular, dando como resultado un anticuerpo o una proteína de unión a antígeno procedente de secuencias completamente humanas.

Además, utilizando ratones o ratas que tengan un repertorio restringido (limitado) de segmentos génicos de región variable de cadena ligera, por ejemplo, un repertorio restringido de segmentos variables de cadena ligera, que comprende uno o más de los segmentos génicos  $V_L$  humanos, pero menos que el número de segmentos de tipo silvestre (por ejemplo, una doble cadena ligera o "DLC", Publicación de Solicitud de Patente de Estados Unidos n.º  
35 2013/0198880) en combinación con la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada descrita anteriormente, puede producirse un dominio variable de cadena ligera de inmunoglobulina que puede emparejarse de un modo más eficaz con un dominio variable de cadena pesada de inmunoglobulina. Así mismo, introduciendo codones de histidina, por ejemplo, mediante la adición de uno o más codones de histidina o la sustitución de uno o más codones no histidina con codones de histidina, en los segmentos  
40 génicos variables de cadena ligera limitada en el genoma de los ratones o ratas descritos en el presente documento, pueden generarse secuencias de aminoácidos de región variable de cadena ligera que pueden conferir reciclabilidad dependiente de pH mejorada a las proteínas de unión a antígeno (por ejemplo, anticuerpos biespecíficos o triespecíficos).

En algunos aspectos, los ratones o ratas modificados genéticamente, como se describe en el presente documento, proporcionan una mayor producción de anticuerpos, aunque al mismo tiempo limitando la diversidad, aumentando de este modo la probabilidad de emparejamiento satisfactorio de cadenas ligeras con cadenas pesadas generadas en una rata o un ratón que comprende una sola región variable de cadena ligera reordenada (por ejemplo, un ratón de Cadena Ligera Universal ("ULC" *Universal Light Chain*); véase, por ejemplo, la publicación previa a la concesión  
45 de patente de Estados Unidos 2013/0185821). En algunos aspectos, las propias cadenas ligeras pueden presentar propiedades de unión a antígeno. En algunos aspectos, el ratón o la rata puede inducirse para que produzcan

proteínas de unión a antígeno que presenten especificidad antigénica que resida en sus cadenas ligeras (por ejemplo, limitando un repertorio de cadenas pesadas de inmunoglobulina de rata o de ratón; por ejemplo, reemplazando el locus de cadena pesada de ratón o rata con un locus que comprenda una sola secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada). En algunos aspectos, las proteínas de unión a antígeno (por ejemplo, anticuerpos) producidas en dichos animales serán específicas para un primer epítipo particular (por ejemplo, antígenos efectores, moléculas citotóxicas, receptores Fc, toxinas, receptores activadores o inhibidores, marcadores de linfocitos T, transportadores de inmunoglobulina, etc.) a través de su unión a la cadena ligera. Dichas cadenas ligeras humanas específicas de epítipo procedentes de estos ratones o ratas puede coexpresarse con cadenas pesadas humanas procedentes de un ratón con un repertorio limitado de cadenas ligeras, por ejemplo, un ratón o una rata de ULC, en el que la cadena pesada se selecciona en función de su capacidad para unirse a un segundo epítipo (por ejemplo, un segundo epítipo en un antígeno diferente).

En diversos aspectos, se describe un ratón o una rata que en su genoma de línea germinal comprende un locus de cadena pesada de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, una secuencia VDJ de cadena pesada reordenada). La secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia de región constante de cadena pesada humana o no humana endógena. En algunos aspectos, un dominio variable de cadena pesada de inmunoglobulina codificado por la secuencia de nucleótidos de región variable de cadena pesada reordenada no es inmunógeno al ratón o a la rata. En algunos aspectos, el ratón o la rata se modifica para que comprenda una secuencia de nucleótidos que codifique dos copias, tres copias, cuatro copias o más, del dominio variable de cadena pesada reordenada, unidas a un dominio constante de cadena pesada. En algunos aspectos, la secuencia de nucleótidos codifica una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. Por ejemplo, la secuencia de nucleótidos puede codificar al menos una, dos, tres, cuatro, cinco copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. En algunos aspectos, la secuencia de nucleótidos codifica 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. En algunos aspectos, el locus comprende una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unidas operativamente a una secuencia génica de dominio constante de cadena pesada.

También se desvela un ratón o una rata que está modificado genéticamente para que contenga un locus de cadena ligera de inmunoglobulina que codifique un dominio variable de cadena pesada reordenada (es decir, un locus de cadena ligera que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada) unido operativamente a una secuencia génica de la región constante de cadena ligera humana o no humana. Por ejemplo, en algunos aspectos, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, una región VDJ previamente diseñada; es decir, una secuencia de cadena pesada común o universal) puede unirse operativamente a una secuencia génica de región constante de cadena ligera direccionando la secuencia de cadena pesada reordenada en un locus de cadena ligera tanto kappa como lambda, de ratón o de rata. Por tanto, en algunos aspectos, la secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está presente en el genoma de la línea germinal del ratón o de la rata. En algunos aspectos, el dominio variable de cadena pesada reordenada expresado por el ratón o la rata modificado genéticamente no es inmunógeno al ratón o la rata. En algunos aspectos, el ratón o la rata se modifica para que comprenda una secuencia de nucleótidos que codifique dos copias, tres copias, cuatro copias o más de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a un dominio constante de cadena ligera. En algunos aspectos, la secuencia de nucleótidos puede codificar una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. Por ejemplo, la secuencia de nucleótidos codifica al menos 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. En algunos aspectos, el locus comprende una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia génica de dominio constante de cadena ligera.

En diversos aspectos, el locus de la cadena ligera de inmunoglobulina de los ratones y las ratas descritos en el presente documento comprende un repertorio limitado de segmentos génicos variables de la cadena ligera, por ejemplo, uno o más segmentos génicos  $V_L$  humanos, pero menos que el número de segmentos de tipo silvestre; y uno o más segmentos génicos  $J_L$ , unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera no humana. Por tanto, se describen ratones y ratas modificados genéticamente que en sus genomas comprenden: (i) un locus de cadena pesada de inmunoglobulina que comprende una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada humana o no humana; y (ii) un locus de cadena ligera de inmunoglobulina que comprende dos o más segmentos génicos  $V_L$  y  $J_L$  variables de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidas operativamente a una secuencia de ácido nucleico de región constante de cadena ligera. En algunos aspectos, la región constante de cadena ligera es una región constante de una rata o de un ratón, por ejemplo, una región constante  $C_k$  de una rata o de un ratón. En algunos aspectos, los segmentos génicos de la región variable humana pueden reordenarse y

codificar dominios variables humanos de un anticuerpo, y el ratón o la rata no comprende ningún segmento génico  $V_L$  endógeno. En algunos aspectos, el ratón o la rata comprende cinco segmentos génicos  $J_k$  humanos, por ejemplo, los segmentos génicos  $J_{k1}$ ,  $J_{k2}$ ,  $J_{k3}$ ,  $J_{k4}$  y  $J_{k5}$ . En algunos aspectos, el locus de cadena ligera de inmunoglobulina comprende dos segmentos génicos  $V_L$  humanos,  $V_{k1-39}$  y  $V_{k3-20}$ . En algunos aspectos, uno o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos  $V_L$  humanos y dos o más segmentos génicos  $J_L$  humanos, están presentes en un locus de cadena ligera endógena, por ejemplo, en un locus de cadena ligera kappa endógena. En algunos aspectos, el ratón comprende un locus de cadena ligera  $\lambda$  funcional. En algunos aspectos, el ratón comprende un locus de cadena ligera  $\lambda$  no funcional. En algunos aspectos, el uno o más segmentos génicos  $V_H$  humanos, uno o más segmentos génicos  $D_H$  humanos y uno o más segmentos génicos  $J_H$  humanos, están unidos operativamente a una secuencia de región constante de cadena pesada de ratón o rata.

En algunos aspectos, los ratones modificados genéticamente que comprenden en sus genomas (i) un locus de cadena pesada de inmunoglobulina que comprende una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina reordenada unida a una secuencia de ácido nucleico de región constante de cadena pesada humana o no humana y (ii) un locus de cadena ligera de inmunoglobulina que comprende dos o más segmentos génicos  $V_L$  y  $J_L$  variables de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera, muestran números de linfocitos B CD19+ y números de linfocitos B maduros, que son sustancialmente iguales a los números observados en ratones de tipo silvestre o en ratones que contienen otras modificaciones de sus locus de inmunoglobulina (es decir, ratones de control modificados genéticamente; por ejemplo, ratones VELOCIMMUNE®, en los que el sistema inmunitario humoral del ratón funciona como el de un ratón de tipo silvestre). En algunos aspectos, dichos ratones demuestran un aumento en el número de linfocitos B inmaduros en el bazo en comparación con ratones de control modificados genéticamente. En aspectos específicos, dichos ratones demuestran un aumento de aproximadamente 2, de aproximadamente 3, de aproximadamente 4, o de aproximadamente 5 veces o más, en los números de linfocitos B inmaduros en el bazo en comparación con ratones de control modificados genéticamente. En algunos aspectos, dichos ratones también son sustancialmente similares a ratones de tipo silvestre o ratones de control modificados genéticamente con respecto al uso de la cadena ligera kappa y gamma en linfocitos B esplénicos. En algunos aspectos, dichos ratones demuestran un aumento en la superficie aumentada en linfocitos B esplénicos (es decir, más expresión en la superficie de IgM por célula) en comparación con ratones de control modificados genéticamente. En algunos aspectos, dichos ratones demuestran un desarrollo alterado de linfocitos B periféricos a través de varios estadios de desarrollo de linfocitos B en el compartimento esplénico en comparación con los ratones de control modificados genéticamente, por ejemplo un aumento en linfocitos B inmaduros, T1 y/o de la zona marginal. En algunos aspectos, dichos ratones demuestran números de linfocitos B CD19+ en el compartimento de la médula ósea que es sustancialmente similar a los números demostrados en los ratones de control modificados genéticamente. En algunos aspectos, dichos ratones demuestran algunos linfocitos pro-B en la médula ósea en comparación con ratones de control modificados genéticamente. En aspectos específicos, los números de linfocitos pro-B en el compartimento de la médula ósea se reducen a aproximadamente 2 veces, aproximadamente 5 veces, aproximadamente 10 veces, aproximadamente 15 veces, aproximadamente 20 veces, aproximadamente 25 veces o más, en comparación con ratones de control modificados genéticamente. En algunos aspectos, dichos ratones demuestran aproximadamente 2 veces, aproximadamente 3 veces, aproximadamente 4 veces, aproximadamente 5 veces, etc., algunos linfocitos B inmaduros y/o maduros en la médula ósea en comparación con ratones de control modificados genéticamente. En algunos aspectos, dichos ratones presentan una ligera preferencia (por ejemplo, un aumento de 2 veces) de uso de genes de cadena ligera lambda en el compartimento de la médula ósea, en comparación con los ratones de control modificados genéticamente.

También se desvela un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente que comprende: (a) una primera secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en el que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera; y (b) una segunda secuencia de nucleótidos que codifica un dominio variable de cadena ligera humana (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada. Por ejemplo, en algunos aspectos, una cadena pesada reordenada de una región VDJ prediseñada (es decir, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada; es decir, una secuencia de cadena pesada común o universal) puede unirse operativamente a una secuencia génica de región constante de cadena ligera direccionando la secuencia de cadena pesada reordenada en un locus de cadena ligera kappa o lambda de ratón. Por tanto, como en otros aspectos, este locus de inmunoglobulina modificado genéticamente puede estar presente en el genoma de la línea germinal del ratón o de la rata. Los ratones y ratas modificados genéticamente comprenden secuencias de nucleótidos de región variable de cadena ligera de inmunoglobulina humana en unión operativa con secuencias génicas de región constante de cadena pesada descritas en la publicación previa a la concesión de patente de Estados Unidos 2012/0096572. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida a una secuencia génica de región constante de cadena ligera kappa. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está

unida operativamente a una secuencia génica de región constante de la cadena ligera  $\kappa$  de rata o de ratón. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$  humana. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$ . En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$  de ratón o de rata. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$  humana.

También se desvela un ratón modificado genéticamente que comprende un locus de cadena ligera de inmunoglobulina que contiene una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada y un locus de cadena pesada de inmunoglobulina que contiene secuencias de dominio variable de cadena ligera de inmunoglobulina humana no reordenada (por ejemplo, genes de cadena ligera kappa) que presenta frecuencias de linfocitos B CD19+ y pre B en la médula ósea que están alteradas con respecto a un ratón de tipo silvestre o a un ratón modificado genéticamente con otras modificaciones en un locus de inmunoglobulina (por ejemplo, ratones de control modificados genéticamente; por ejemplo, ratones VELOCIMMUNE®, en los que el sistema inmunitario humoral de ratón funciona como el de un ratón de tipo silvestre). En aspectos específicos, los números de linfocitos pre B y linfocitos B CD19+ en la médula ósea son 2 veces menores, 3 veces menores, 4 veces menores o 5 veces menores en comparación con un ratón de tipo silvestre o con un ratón de control con locus de inmunoglobulina modificado genéticamente. En aspectos específicos, el número de linfocitos B inmaduros en la médula ósea es 2 veces menor, 3 veces menor, 4 veces menor o 5 veces menor en comparación con un ratón de tipo silvestre o con un ratón de control con locus de inmunoglobulina modificado genéticamente. En algunos aspectos, un ratón modificado genéticamente comprende un locus de cadena ligera de inmunoglobulina que contiene una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada y un locus de cadena pesada de inmunoglobulina que contiene secuencias de dominio variable de cadena ligera de inmunoglobulina humana no reordenada (por ejemplo, genes de cadena ligera kappa) que no expresa o no expresa esencialmente genes de cadena ligera lambda en las células de médula ósea. En algunos aspectos, un ratón modificado genéticamente que comprende un locus de cadena ligera de inmunoglobulina contiene una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada y un locus de cadena pesada de inmunoglobulina que contiene secuencias de dominio variable de cadena ligera de inmunoglobulina humana no reordenada (por ejemplo, genes de cadena ligera kappa) tiene niveles reducidos de linfocitos B esplénicos en comparación con un ratón de tipo silvestre o con un ratón de control con locus de inmunoglobulina modificado genéticamente. En aspectos específicos, los niveles de linfocitos B esplénicos y de linfocitos B maduros son 2 veces menores, 3 veces menores, 4 veces menores o 5 veces menores en comparación con un ratón de tipo silvestre o con un ratón de control con locus de inmunoglobulina modificado genéticamente. En algunos aspectos, un ratón modificado genéticamente que comprende un locus de cadena ligera de inmunoglobulina que contiene una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada y un locus de cadena pesada de inmunoglobulina que contiene secuencias de dominio variable de cadena ligera de inmunoglobulina humana no reordenada (por ejemplo, genes de cadena ligera kappa) no expresa o esencialmente no expresa genes de cadena ligera lambda en células B esplénicas. En aspectos específicos, un ratón modificado genéticamente que comprende un locus de cadena ligera de inmunoglobulina que contiene una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada y un locus de cadena pesada de inmunoglobulina que contiene secuencias de dominio variable de cadena ligera de inmunoglobulina humana no reordenada (por ejemplo genes de cadena ligera kappa) tiene una frecuencia aumentada de células en la fase T1 en el bazo en comparación con un ratón de tipo silvestre o un ratón de control con locus de inmunoglobulina modificado genéticamente.

Aunque en el presente documento se analizan extensamente aspectos que emplean un dominio variable de cadena pesada humana reordenada en un ratón (es decir, un ratón con un locus de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), también se describen ratas que comprenden un locus de inmunoglobulina modificado genéticamente que codifica un dominio variable de cadena pesada humana reordenada. Como métodos para modificar un genoma animal no humano (por ejemplo, un genoma de un cerdo, vaca, roedor, gallina, etc.) se incluyen, por ejemplo, el empleo de una nucleasa de dedo de zinc (ZFN) o una nucleasa efectora de tipo activador de transcripción (TALEN, *transcription activator-like effector nuclease*) para modificar un genoma que incluya una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada.

El animal empleado es un ratón o una rata. En algunos aspectos, el animal es un ratón.

En algunos aspectos, el ratón o la rata es un ratón de una cepa C57BL seleccionada de C57BL/A, C57BL/An, C57BL/GrFa, C57BL/KaLwN, C57BL/6, C57BL/6J, C57BL/6ByJ, C57BL/6N, C57BL/6NJ, C57BL/10, C57BL/10ScSn, C57BL/10Cr y C57BL/Ola. En otro aspecto, el ratón es una cepa 129. En algunos aspectos, la cepa 129 se selecciona del grupo que consiste en 129P1, 129P2, 129P3, 129X1, 129S1, 129S1 (por ejemplo, 129S1/SV, 129S1/SvIm), 129S2, 129S4, 129S5, 129S9/SvEvH, 129S6 (129/SvEvTac), 129S7, 129S8, 129T1, 129T2 (véase, por ejemplo, Festing *et al.* (1999) Revised nomenclature for strain 129 mice, *Mammalian Genome* 10: 836, véase también,

Auerbach *et al.* (2000) Establishment and Chimera Analysis of 129/SvEv- and C57BL/6-Derived Mouse Embryonic Stem Cell Lines). En algunos aspectos, el ratón modificado genéticamente es una mezcla de una cepa 129 mencionada anteriormente y de una cepa C57BL/6 mencionada anteriormente (por ejemplo, una cepa C57BL/6). En otro aspecto, el ratón es una mezcla de las cepas 129 mencionadas anteriormente, o una mezcla de las cepas C57BL/6 mencionadas anteriormente. En algunos aspectos, la cepa 129 de la mezcla es una cepa 129S6 (129/SvEvTac). En otro aspecto, el ratón es una mezcla de una cepa procedente de 129/SvEv- y una cepa C57BL/6. En un aspecto específico, el ratón es una mezcla de una cepa 129/SvEv- y una cepa C57BL/6 como se describe en Auerbach *et al.* 2000 *BioTechniques* 29: 1024-1032. En otro aspecto, el ratón es una cepa BALB, por ejemplo, una cepa BALB/c. En otro aspecto, el ratón es una mezcla de una cepa BALB (por ejemplo cepa BALB/c) y de otra cepa mencionada anteriormente.

En algunos aspectos, el animal no humano es una rata. En algunos aspectos, la rata se selecciona de una rata Wistar, una cepa LEA, una cepa Sprague Dawley, una cepa Fischer, F344, F6, ACI y Dark Agouti (DA). En algunos aspectos, la cepa de rata es una mezcla de dos o más de una cepa seleccionada del grupo que consiste en Wistar, LEA, Sprague Dawley, Fischer, F344, F6, ACI y Dark Agouti (DA).

En algunos aspectos, se describe un ratón modificado genéticamente que en su genoma de línea germinal comprende un locus de cadena pesada de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que genera poblaciones de linfocitos B esplénicos maduros e inmaduros que son esencialmente normales con respecto a un ratón de tipo silvestre. En algunos aspectos, dicho ratón modificado genéticamente tiene una ligera disminución en el uso de las secuencias génicas lambda de cadena ligera con respecto al tipo silvestre en linfocitos B esplénicos. En aspectos específicos, dicho ratón modificado genéticamente utiliza secuencias génicas lambda de cadena ligera con una frecuencia menor de 2 veces, 3 veces, 4 veces o 5 veces en comparación con el tipo silvestre en linfocitos B esplénicos. En algunos aspectos, dicho ratón modificado genéticamente tiene una ligera disminución en la población T1 de linfocitos B esplénicos y un aumento de linfocitos B esplénicos en la zona marginal con respecto al tipo silvestre. En algunos aspectos, dicho ratón modificado genéticamente tiene poblaciones de linfocitos B casi normales en la médula ósea. En algunos aspectos, dicho ratón modificado genéticamente utiliza secuencias génicas lambda con una frecuencia que es la mitad o menor que la mitad de la frecuencia de las secuencias génicas lambda utilizadas en el tipo silvestre.

En diversos aspectos, como se describe en el presente documento, el dominio variable de cadena pesada reordenada (es decir, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada) procede de una secuencia o segmento génico V, D y J humana. En algunos aspectos, el dominio variable de cadena pesada reordenada procede de un segmento J de la línea germinal humana, de un segmento D de la línea germinal humana y de un segmento V de la línea germinal humana. En algunos aspectos, el segmento V<sub>H</sub> humano corresponde a variantes observadas en la población humana.

En algunos aspectos, la secuencia de región constante de cadena pesada de un animal humano o no humano, comprende una secuencia seleccionada de un C<sub>H</sub>1, una bisagra, un C<sub>H</sub>2, un C<sub>H</sub>3 y una combinación de los mismos. En aspectos específicos, la secuencia de región constante comprende un C<sub>H</sub>1, una bisagra, un C<sub>H</sub>2 y un C<sub>H</sub>3. En diversos aspectos, como se describe en el presente documento, el segmento génico J humano se selecciona del grupo que consiste en J<sub>H</sub>1, J<sub>H</sub>2, J<sub>H</sub>3, J<sub>H</sub>4, J<sub>H</sub>5, J<sub>H</sub>6 y una variante polimórfica de los mismos. La secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada codifica la secuencia de V<sub>H</sub>3-23/GY/J<sub>H</sub>4-4 (SEQ ID NO: 137) humana. En algunos aspectos, el dominio variable de cadena pesada reordenada codificado por y expresado de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada comprende la secuencia de V<sub>H</sub>3-23/X<sub>1</sub>X<sub>2</sub>/J humana (en la que X<sub>1</sub> y X<sub>2</sub> es cualquier aminoácido). En algunos aspectos, X<sub>1</sub> es Gly y X<sub>2</sub> es Tyr. En algunos aspectos, el dominio variable de cadena pesada reordenada comprende la secuencia de V<sub>H</sub>3-23/X<sub>1</sub>X<sub>2</sub>/J<sub>H</sub>4-4 humana (en la que X<sub>1</sub> y X<sub>2</sub> es cualquier aminoácido). En algunos aspectos, X<sub>2</sub> es un aminoácido que comprende un grupo fenilo. En aspectos específicos, X<sub>2</sub> se selecciona de Tyr y Phe.

En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada comprende un segmento D humano que no es autorreactivo (no inmunógeno) en el animal. En algunos aspectos, la secuencia de nucleótidos comprende un segmento D humano que puede expresarse en una secuencia variable de cadena pesada de un linfocito B maduro de un ratón. En algunos aspectos, el segmento D es un segmento que se ha expresado en un ratón que comprende un locus de inmunoglobulina humanizado que comprende un segmento V<sub>H</sub> humano, D<sub>H</sub> humano y J<sub>H</sub> humano

Diversos aspectos utilizan o engloban características o información de secuencia procedente del ratón humanizado VELOCIMMUNE®. Los ratones VELOCIMMUNE® humanizados contienen un reemplazo exacto, a gran escala de regiones variables de la línea germinal de cadena pesada de inmunoglobulina (IgH) de ratón y de cadena ligera de inmunoglobulina (por ejemplo, cadena ligera κ, Igκ) con correspondientes regiones variables de inmunoglobulina humana, en los locus endógenos (véanse, por ejemplo, los documentos US 6.596.541 y US 8.502.018). En total, aproximadamente seis megabases de locus de ratón se reemplazan con aproximadamente 1,5 megabases de secuencia genómica humana. Este reemplazo exacto da como resultado un ratón con locus de inmunoglobulina

híbridos que hacen que las cadenas pesada y ligera tengan una región variable humana y una región constante de ratón. El reemplazo exacto de los segmentos  $V_H$ -D- $J_H$  y  $V_k$ - $J_k$  de ratón dejan intactas secuencias de ratón flanqueantes y funcionales en los locus de inmunoglobulina híbridos. El sistema inmunitario humoral del ratón funciona como el del ratón de tipo silvestre. El desarrollo de linfocitos B no se ve obstaculizado en ningún aspecto significativo y se genera una rica diversidad de regiones variables humanas en el ratón después de la exposición antigénica. Por otra parte, los ratones VELOCIMMUNE® humanizados presentan una respuesta esencialmente normal, de tipo silvestre, a la inmunización que difiere solo en un aspecto significativo respecto de los ratones de tipo silvestre en los que las regiones variables generadas en respuesta a la humanización son completamente humanas. Los ratones VELOCIMMUNE® humanizados son posibles porque los segmentos génicos de inmunoglobulina para las cadenas pesada y ligera  $\kappa$  se reordenan de manera similar en seres humanos y en ratones. Aunque los locus no son idénticos, son lo suficientemente similares para que la humanización del locus génico variable de cadena pesada pueda realizarse reemplazando aproximadamente tres millones de pares de bases de secuencia de ratón contiguas que contienen todos los segmentos génicos  $V_H$ , D y  $J_H$  cubriendo aproximadamente un millón de bases de secuencia génica humana contigua básicamente la secuencia equivalente de un locus de inmunoglobulina humana. Por ejemplo, en algunos aspectos, los segmentos D proceden de una cadena pesada expresada en un linfocito B maduro de un ratón VELOCIMMUNE® humanizado inmunizado con un antígeno, en el que el segmento D contribuye a no más de dos aminoácidos en la secuencia CDR3 de cadena pesada.

En aspectos particulares, se proporciona un ratón VELOCIMMUNE® que comprende un locus de cadena pesada de inmunoglobulina que codifica un dominio variable de cadena pesada reordenada (es decir, que comprende un locus de cadena pesada de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada). Un ratón VELOCIMMUNE® modificado de esta manera comprende un reemplazo de segmentos génicos variables de cadena pesada de inmunoglobulina de ratón con una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, una secuencia de Cadena Pesada Universal en un locus de cadena pesada endógena) y un reemplazo de segmentos génicos variables de cadena ligera  $\kappa$  de inmunoglobulina de ratón con al menos 40 segmentos génicos  $V_k$  humanos y cinco segmentos génicos  $J_k$  humanos. En algunos aspectos, los segmentos génicos  $V_k$  humanos se seleccionan del grupo que consiste en  $V_{k1-5}$ ,  $V_{k1-6}$ ,  $V_{k1-8}$ ,  $V_{k1-9}$ ,  $V_{k1-12}$ ,  $V_{k1-13}$ ,  $V_{k1-16}$ ,  $V_{k1-17}$ ,  $V_{k1-22}$ ,  $V_{k1-27}$ ,  $V_{k1-32}$ ,  $V_{k1-33}$ ,  $V_{k1-35}$ ,  $V_{k1-37}$ ,  $V_{k1-39}$ ,  $V_{k1D-8}$ ,  $V_{k1D-12}$ ,  $V_{k1D-13}$ ,  $V_{k1D-16}$ ,  $V_{k1D-17}$ ,  $V_{k1D-22}$ ,  $V_{k1D-27}$ ,  $V_{k1D-32}$ ,  $V_{k1D-33}$ ,  $V_{k1D-35}$ ,  $V_{k1D-37}$ ,  $V_{k1D-39}$ ,  $V_{k1D-42}$ ,  $V_{k1D-43}$ ,  $V_{k1-NL1}$ ,  $V_{k2-4}$ ,  $V_{k2-10}$ ,  $V_{k2-14}$ ,  $V_{k2-18}$ ,  $V_{k2-19}$ ,  $V_{k2-23}$ ,  $V_{k2-24}$ ,  $V_{k2-26}$ ,  $V_{k2-28}$ ,  $V_{k2-29}$ ,  $V_{k2-30}$ ,  $V_{k2-36}$ ,  $V_{k2-38}$ ,  $V_{k2-40}$ ,  $V_{k2D-10}$ ,  $V_{k2D-14}$ ,  $V_{k2D-18}$ ,  $V_{k2D-19}$ ,  $V_{k2D-23}$ ,  $V_{k2D-24}$ ,  $V_{k2D-26}$ ,  $V_{k2D-28}$ ,  $V_{k2D-29}$ ,  $V_{k2D-30}$ ,  $V_{k2D-36}$ ,  $V_{k2D-38}$ ,  $V_{k2D-40}$ ,  $V_{k3-7}$ ,  $V_{k3-11}$ ,  $V_{k3-15}$ ,  $V_{k3-20}$ ,  $V_{k3-25}$ ,  $V_{k3-31}$ ,  $V_{k3-34}$ ,  $V_{k3D-7}$ ,  $V_{k3D-7}$ ,  $V_{k3D-11}$ ,  $V_{k3D-15}$ ,  $V_{k3D-15}$ ,  $V_{k3D-20}$ ,  $V_{k3D-25}$ ,  $V_{k3D-31}$ ,  $V_{k3D-34}$ ,  $V_{k3-NL1}$ ,  $V_{k3-NL2}$ ,  $V_{k3-NL3}$ ,  $V_{k3-NL4}$ ,  $V_{k3-NL5}$ ,  $V_{k4-1}$ ,  $V_{k5-2}$ ,  $V_{k6-21}$ ,  $V_{k6D-21}$ ,  $V_{k6D-41}$  y  $V_{k7-3}$ . En algunos aspectos, los segmentos génicos  $V_k$  humanos comprenden  $V_{k4-1}$ ,  $V_{k5-2}$ ,  $V_{k7-3}$ ,  $V_{k2-4}$ ,  $V_{k1-5}$  y  $V_{k1-6}$ . En un aspecto, los segmentos génicos  $V_k$  comprenden  $V_{k3-7}$ ,  $V_{k1-8}$ ,  $V_{k1-9}$ ,  $V_{k2-10}$ ,  $V_{k3-11}$ ,  $V_{k1-12}$ ,  $V_{k1-13}$ ,  $V_{k2-14}$ ,  $V_{k3-15}$  y  $V_{k1-16}$ . En algunos aspectos, los segmentos génicos  $V_k$  comprenden  $V_{k1-17}$ ,  $V_{k2-18}$ ,  $V_{k2-19}$ ,  $V_{k3-20}$ ,  $V_{k6-21}$ ,  $V_{k1-22}$ ,  $V_{k1-23}$ ,  $V_{k2-24}$ ,  $V_{k3-25}$ ,  $V_{k2-26}$ ,  $V_{k1-27}$ ,  $V_{k2-28}$ ,  $V_{k2-29}$ , y  $V_{k2-30}$ . En algunos aspectos, los segmentos génicos  $V_k$  comprenden  $V_{k3-31}$ ,  $V_{k3-32}$ ,  $V_{k1-33}$ ,  $V_{k3-34}$ ,  $V_{k1-35}$ ,  $V_{k2-36}$ ,  $V_{k1-37}$ ,  $V_{k2-38}$ ,  $V_{k1-39}$  y  $V_{k2-40}$ . En algunos aspectos, los segmentos génicos  $V_k$  comprenden segmentos génicos de inmunoglobulina  $\kappa$  humanos contiguos que abarcan el locus de cadena ligera  $\kappa$  de inmunoglobulina humana desde  $V_{k4-1}$  a  $V_{k2-40}$ , y los segmentos génicos  $J_k$  comprenden los segmentos génicos contiguos que abarcan el locus de cadena ligera  $\kappa$  de inmunoglobulina humana de  $J_{k1}$  a  $J_{k5}$ . En algunos aspectos, la secuencia de nucleótidos de dominio variable de cadena pesada humana reordenada está unida operativamente a una secuencia de región constante de cadena pesada de ratón. Un ratón VELOCIMMUNE® comprende un locus de cadena pesada de inmunoglobulina que codifica un dominio variable de cadena pesada reordenada (es decir, que comprende un locus de cadena pesada de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada) que puede utilizarse en cualquiera de los aspectos, realizaciones, métodos etc. descritos en el presente documento.

En diversos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada humana o de ratón (por ejemplo, una secuencia génica de región constante de cadena pesada que codifica un isotipo de inmunoglobulina seleccionado de IgM, IgD, IgA, IgE, IgG y combinaciones de los mismos). Por ejemplo, se describen ratones o ratas modificados genéticamente que comprenden locus de inmunoglobulina en los que: (a) una primera secuencia de nucleótidos codifica un dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada humana o no humana; y (b) una segunda secuencia de nucleótidos codifica un dominio variable de cadena ligera (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera humana o no humana. En algunos aspectos, la secuencia génica de región constante de cadena pesada humana se selecciona de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y combinaciones de los mismos. En algunos aspectos, la secuencia génica de región constante de cadena pesada de ratón se selecciona de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y combinaciones de los mismos. En algunos aspectos, el reemplazo adicional de

determinadas secuencias génicas de la región constante de animales no humanos con secuencias génicas humanas (por ejemplo, reemplazo de la secuencia C<sub>H1</sub> de ratón por la secuencia C<sub>H1</sub> humana, y el reemplazo de la secuencia C<sub>L</sub> de ratón por la secuencia C<sub>L</sub> humana) da como resultado animales no humanos modificados genéticamente con locus de inmunoglobulina híbridos que crean anticuerpos que tienen regiones variables humanas y regiones constantes parcialmente humanas, adecuadas, por ejemplo, para crear fragmentos de anticuerpo completamente humanos, por ejemplo, Fab' completamente humanos. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada de rata. En algunos aspectos, la secuencia génica de región constante de cadena pesada de rata se selecciona de un C<sub>H1</sub>, una bisagra, un C<sub>H2</sub>, un C<sub>H3</sub> y combinaciones de los mismos. En diversos aspectos, el locus de cadena pesada de inmunoglobulina modificado genéticamente del ratón o de la rata comprende dos copias, tres copias, cuatro copias o más de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia génica de dominio constante de cadena pesada. En un aspecto particular, el locus comprende una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia génica de dominio constante de cadena pesada.

En diversos aspectos, la secuencia de nucleótidos de región constante de cadena pesada comprende una modificación en un C<sub>H2</sub> o en un C<sub>H3</sub>, en el que la modificación aumenta la afinidad de la secuencia de aminoácidos de región constante de cadena pesada por FcRn en un entorno ácido (por ejemplo, en un endosoma en donde el pH varía de aproximadamente 5,5 a aproximadamente 6,0). En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de región constante de cadena pesada humana que comprende una modificación en la posición 250 según la numeración EU (263 según la numeración Kabat) (por ejemplo, E o Q); 250 según la numeración EU (263 según la numeración Kabat) y 428 según la numeración EU (459 según la numeración Kabat) (por ejemplo, L o F); 252 según la numeración EU (265 según la numeración Kabat) (por ejemplo, L/Y/F/W o T), 254 según la numeración EU (267 según la numeración Kabat) (por ejemplo, S o T), y 256 según la numeración EU (269 según la numeración Kabat) (por ejemplo, S/R/Q/E/D o T); o una modificación en la posición 428 según la numeración EU (459 según la numeración Kabat) y/o 433 según la numeración EU (464 según la numeración Kabat) (por ejemplo, L/R/S/P/Q o K) y/o 434 según la numeración EU (465 según la numeración Kabat) (por ejemplo, H/F o Y); o una modificación en la posición 250 según la numeración EU (263 según la numeración Kabat) y/o 428 según la numeración EU (459 según la numeración Kabat); o una modificación en la posición 307 según la numeración EU (326 según la numeración Kabat) o 308 según la numeración EU (327 según la numeración Kabat) (por ejemplo, 308F, V308F) y 434 según la numeración EU (465 según la numeración Kabat). En un aspecto, la modificación comprende una modificación 428L (por ejemplo, M428L) y 434S (por ejemplo, N434S) según la numeración EU (una 459, por ejemplo, M459L y 465S (por ejemplo, N465S) modificación por numeración Kabat); una modificación 428L, 259I (por ejemplo, V259I) y 308F (por ejemplo, V308F) según la numeración EU (una 459L, 272I (por ejemplo, V272I) y 327F (por ejemplo, V327F) modificación por numeración Kabat; una modificación 433K (por ejemplo, H433K) ) y 434 (por ejemplo, 434Y) según la numeración EU (una modificación 464K (por ejemplo, H464K) y una modificación 465 (por ejemplo, 465Y) según la numeración Kabat; una modificación 252, 254 y 256 (por ejemplo, 252Y, 254T y 256E)) según la numeración EU (una modificación 265, 267, 269 (por ejemplo, 265Y, 267T y 269E) por numeración Kabat; una modificación 250Q y 428L (por ejemplo, T250Q y M428L) según la numeración EU (una modificación 263Q y 459L, por ejemplo, T263Q y M459L, según la numeración Kabat), y una modificación 307 y/o 308 (por ejemplo, 307F o 308P) según la numeración EU (modificación 326 y/o 327, por ejemplo, 326F o 308P, según la numeración Kabat), en la que la modificación aumenta la afinidad de la secuencia de aminoácidos de región constante de cadena pesada por FcRn en un entorno ácido (por ejemplo, en un endosoma en donde el pH varía de aproximadamente 5,5 a aproximadamente 6,0). En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de C<sub>H2</sub> humana que comprende al menos una modificación entre los restos de aminoácido en las posiciones 252 y 257 según la numeración EU (es decir, al menos una modificación entre las posiciones de aminoácidos 265 y 270 según la numeración Kabat), en la que la modificación aumenta la afinidad de la secuencia de aminoácidos de C<sub>H2</sub> humana por FcRn en un entorno ácido (por ejemplo, en un endosoma donde el pH varía de aproximadamente 5,5 a aproximadamente 6,0). En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de C<sub>H2</sub> humana que comprende al menos una modificación entre los restos de aminoácido en las posiciones 307 y 311 (es decir, al menos una modificación entre las posiciones de aminoácidos 326 y 330 según la numeración Kabat), en la que la modificación aumenta la afinidad de la secuencia de aminoácidos de C<sub>H2</sub> por FcRn en un entorno ácido (por ejemplo, en un endosoma en donde el pH varía de aproximadamente 5,5 a aproximadamente 6,0). En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de C<sub>H3</sub> humana, en la que la secuencia de aminoácidos C<sub>H3</sub> comprende al menos una modificación entre los restos de aminoácido en las posiciones 433 y 436 según la numeración EU (es decir, al menos una modificación entre los restos de aminoácido en las posiciones 464 y 467 según la numeración Kabat), en la que la modificación aumenta la afinidad de la secuencia de aminoácidos de C<sub>H3</sub> por FcRn en un entorno ácido (por ejemplo, en un endosoma donde el pH varía de aproximadamente 5,5 a aproximadamente 6,0). En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de región constante de cadena pesada humana que comprende una mutación seleccionada del grupo que consiste en M428L según la numeración EU (459 según la numeración Kabat), N434S según la numeración EU (465 según la numeración Kabat)

y una combinación de las mismas. En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de región constante de cadena pesada humana que comprende una mutación seleccionada del grupo que consiste en M428L según numeración EU (M459L según numeración Kabat), V259I según numeración EU (V272I según numeración Kabat), V308F según numeración EU (V327 según numeración Kabat) y una combinación de las mismas. En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de región constante de cadena pesada humana que comprende una mutación N434A según numeración EU (una mutación N465A según numeración Kabat). En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de región constante de cadena pesada humana que comprende una mutación seleccionada del grupo que consiste en M252Y según numeración EU (M265Y según numeración Kabat), S254T según numeración EU (S267T según numeración Kabat), T256E según numeración EU (T269E según numeración Kabat) y una combinación de las mismas. En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de región constante de cadena pesada humana que comprende una mutación seleccionada del grupo que consiste en T250Q según numeración EU (T263Q según numeración Kabat), M428L según numeración EU (M459L según numeración Kabat), o ambas. En algunos aspectos, la secuencia de nucleótidos de región constante de cadena pesada codifica una secuencia de aminoácidos de región constante de cadena pesada humana que comprende una mutación seleccionada del grupo que consiste en H433K según numeración EU (H464K según numeración Kabat), N434Y según numeración EU (N465Y según numeración Kabat), o ambas.

En algunos aspectos, un locus de inmunoglobulina modificado genéticamente comprende: (1) un primer alelo, en el que la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada como se describe en el presente documento está unida operativamente a una primera secuencia de nucleótidos de región constante de cadena pesada que codifica una primera secuencia de aminoácidos de  $C_{H3}$  de una IgG humana seleccionada de IgG1, IgG2, IgG4 y una combinación de las mismas; y (2) un segundo alelo, en el que la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada como se describe en el presente documento está unida operativamente a una segunda secuencia de nucleótidos de región constante de cadena pesada que codifica una segunda secuencia de aminoácidos  $C_{H3}$  de la IgG humana seleccionada de IgG1, IgG2, IgG4 y una combinación de los mismos, y en la que la segunda secuencia de aminoácidos de  $C_{H3}$  comprende una modificación que reduce o elimina la unión de la segunda secuencia de aminoácidos de  $C_{H3}$  con la Proteína A (véase, por ejemplo, la Patente de Estados Unidos n.º 8.586.713). En algunos aspectos, la segunda secuencia de aminoácidos de  $C_{H3}$  comprende una modificación H95R (según numeración de exones de IMGT; H435R según numeración EU). En un aspecto, la segunda secuencia de aminoácidos de  $C_{H3}$  comprende adicionalmente una modificación Y96F (según numeración de exones IMGT, H436F según numeración EU). En otro aspecto, la segunda secuencia de aminoácidos de  $C_{H3}$  comprende tanto una modificación H95R (según numeración de exones IMGT, H435R según numeración EU) como una modificación Y96F (según numeración de exones IMGT, H436F según EU). En algunos aspectos, la segunda secuencia de aminoácidos de  $C_{H3}$  es de una IgG1 humana modificada y comprende adicionalmente una mutación seleccionada del grupo que consiste en D16E, L18M, N44S, K52N, V57M y V82I (IMGT; D356E, L38M, N384S, K392N, V397M y V422I según EU). En algunos aspectos, la segunda secuencia de aminoácidos de  $C_{H3}$  es de una IgG2 humana modificada y comprende adicionalmente una mutación seleccionada del grupo que consiste en N44S, K52N y V82I (IMGT: N384S, K392N y V422I por EU). En algunos aspectos, la segunda secuencia de aminoácidos de  $C_{H3}$  es de una IgG4 humana modificada y comprende adicionalmente una mutación seleccionada del grupo que consiste en Q15R, N44S, K52N, V57M, R69K, E79Q y V82I (IMGT: Q355R, N384S, K392N, V397M, R409K, E419Q y V422I según EU). En algunos aspectos, la secuencia de aminoácidos de región constante de cadena pesada es una secuencia de aminoácidos de región constante no humana, y la secuencia de aminoácidos de región constante de cadena pesada comprende uno o más de cualquiera de los tipos de modificaciones descritas anteriormente.

En diversos aspectos, los dominios Fc se modifican para tener la unión al receptor Fc alterada, que a su vez afecta a la función efectora. En algunos aspectos, una región constante de cadena pesada ( $C_H$ ) modificada por ingeniería genética, que incluye el dominio Fc, es quimérica. Como tal, una región  $C_H$  quimérica combina dominios  $C_H$  procedentes de más de un isotipo de inmunoglobulina. Por ejemplo, una región  $C_H$  quimérica comprende parte o todo de un dominio  $C_{H2}$  procedente de una molécula de IgG1 humana, IgG2 humana o IgG4 humana, combinada con parte o todo de un dominio  $C_{H3}$  procedente de una molécula de IgG1 humana, IgG2 humana o IgG4 humana. En algunos aspectos, una región  $C_H$  quimérica contiene una región bisagra quimérica. Por ejemplo, una bisagra quimérica puede comprender una secuencia de aminoácidos de "bisagra superior" (restos de aminoácido de las posiciones 216 a 227 según la numeración EU; los restos de aminoácidos de las posiciones 226 a 240 según la numeración de Kabat) procedente de una región bisagra de una IgG1 humana, IgG2 humana o IgG4 humana, combinada con una secuencia de "bisagra inferior" (restos de aminoácidos de las posiciones 228 a 236 según la numeración EU; posiciones de aminoácidos de las posiciones 241 a 249 según la numeración Kabat) procedente de una región bisagra de IgG1 humana, IgG2 humana o IgG4 humana. En algunos aspectos, la región bisagra quimérica comprende los restos de aminoácido procedentes de una bisagra superior de IgG1 humana o IgG4 humana y restos de aminoácidos procedentes de una bisagra inferior de IgG2 humana.

En algunos aspectos, el dominio Fc puede modificarse por ingeniería genética para activar todas, algunas, o ninguna, de las funciones efectoras normales del dominio Fc, sin afectar a las propiedades farmacocinéticas

deseadas de la proteína (por ejemplo, anticuerpos) que contiene dicho dominio Fc. Para ejemplos de proteínas que comprenden regiones C<sub>H</sub> quiméricas y que tienen funciones efectoras alteradas, véase la Solicitud Provisional de Estados Unidos n.º 61/759.578, presentada el 1 de febrero de 2013.

- 5 El genoma de los ratones o ratas se modifica (i) para delecionar o hacer no funcional (por ejemplo, mediante la inserción de una secuencia de nucleótidos (por ejemplo, una secuencia de nucleótidos exógena)) en el locus de inmunoglobulina o mediante una reordenación no funcional o inversión de todos los segmentos V<sub>H</sub>, D, J<sub>H</sub> endógenos de inmunoglobulina funcionales; y (ii) para que comprenda una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, en la que la secuencia de nucleótidos está presente en un
- 10 locus endógeno (es decir, en el que la secuencia de nucleótidos está localizada en un ratón o una rata de tipo silvestre). En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está en un locus ectópico en el genoma (por ejemplo, en un locus diferente del locus de la cadena de inmunoglobulina endógena en su genoma, o dentro de su locus endógeno, por ejemplo, dentro de un locus variable de inmunoglobulina, en el que el locus endógeno se coloca o se traslada a una
- 15 localización diferente en el genoma). En algunos aspectos, por ejemplo, aproximadamente el 80 % o más, aproximadamente el 85 % o más, aproximadamente el 90 % o más, aproximadamente el 95 % o más, aproximadamente el 96 % o más, aproximadamente el 97 % o más, aproximadamente el 98 % o más o aproximadamente el 99 % o más de todos los segmentos génicos V, D o J de cadena pesada funcional endógena están delecionados o se han hecho no funcionales. En algunos aspectos, por ejemplo, al menos el 95 %, 96 %, 97 %, 98 % o 99 % de los segmentos génicos V, D o J de cadena pesada funcional endógena están delecionados o se han hecho no funcionales. La secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada endógena.
- 25 En algunos aspectos, el ratón o la rata modificado genéticamente comprende una modificación que deleciona o hace no funcionales a los segmentos génicos variables de cadena pesada V<sub>H</sub>, D y J<sub>H</sub> funcionales endógenos y a los segmentos génicos V<sub>L</sub> y J<sub>L</sub> variables de cadena ligera funcionales endógenos; y comprende (i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada y (ii) una secuencia de nucleótidos que codifica segmentos génicos V de cadena ligera (V<sub>L</sub>) de inmunoglobulina humana no reordenada y
- 30 segmentos génicos J de cadena ligera (J<sub>L</sub>) de inmunoglobulina humana no reordenada (es decir, en la que la secuencia de nucleótidos se localiza en un animal no humano de tipo silvestre) o dentro de su locus endógeno, por ejemplo, dentro de un locus de región variable de inmunoglobulina, en el que el locus endógeno se coloca o se traslada a una localización diferente en el genoma). En algunos aspectos, el ratón o la rata modificado genéticamente comprende una modificación que deleciona o hace no funcionales a segmentos génicos variables de
- 35 cadena pesada V<sub>H</sub>, D y J<sub>H</sub> funcionales endógenos y a segmentos génicos V<sub>L</sub> y J<sub>L</sub> variables de cadena ligera funcionales; y comprende (i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada y (ii) uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre (V<sub>L</sub> y J<sub>L</sub>) en una localización endógena (es decir en la que la secuencia de nucleótidos se localiza en un ratón o una rata de tipo silvestre) o dentro de su locus endógeno, por ejemplo, dentro de un locus de región variable de inmunoglobulina, en el que el locus endógeno se coloca o se traslada a una localización diferente en el genoma). La secuencia de nucleótidos de región variable de
- 40 cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada endógena. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera humana o no humana, bien kappa o lambda.
- 45

Diversos aspectos incluyen dominios variables de cadena ligera. Las secuencias de ácido nucleico que codifican dominios variables de cadena ligera pueden utilizarse en la creación de los ratones y ratas modificados genéticamente descritos en el presente documento, pueden expresarse por dichos animales y/o pueden codificar aminoácidos presentes en los anticuerpos representados producidos por, (o procedentes de secuencias diversificadas por) dichos animales. En algunos aspectos, el dominio variable de cadena ligera es un dominio variable de cadena ligera κ humana. En algunos aspectos, el dominio variable de cadena ligera es un dominio variable de cadena ligera κ de ratón. En algunos aspectos, el dominio variable de cadena ligera es un dominio variable de cadena ligera κ de rata. En algunos aspectos, el dominio variable de cadena ligera es un dominio

50 variable de cadena ligera λ humana. En algunos aspectos, el dominio variable de cadena ligera es un dominio variable de cadena ligera λ de ratón. En algunos aspectos, el dominio variable de cadena ligera es un dominio variable de cadena ligera λ de rata.

En diversos aspectos, los dominios variables de cadena ligera producidos por los ratones y ratas modificados genéticamente descritos en el presente documento, están codificados por uno o más segmentos génicos variables de cadena ligera κ de inmunoglobulina humana o de ratón. En algunos aspectos, el uno o más segmentos génicos variables de cadena ligera κ de inmunoglobulina de ratón comprenden aproximadamente tres megabases del locus de cadena ligera κ de inmunoglobulina de ratón. En algunos aspectos, el uno o más segmentos génicos variables de cadena ligera κ de inmunoglobulina de ratón comprende al menos 137 segmentos génicos Vκ, al menos cinco

60 segmentos génicos Jκ o una combinación de los mismos del locus de cadena ligera κ de inmunoglobulina de ratón. En algunos aspectos, el uno o más segmentos génicos variables de cadena ligera κ de inmunoglobulina humana

65

comprende aproximadamente una mitad de megabases de un locus de cadena ligera  $\kappa$  de inmunoglobulina humana. En aspectos específicos, el uno o más segmentos génicos variables de cadena ligera  $\kappa$  de inmunoglobulina humana comprenden la repetición proximal (con respecto a la región constante  $\kappa$  de inmunoglobulina) de un locus de cadena ligera  $\kappa$  de inmunoglobulina humana. En algunos aspectos, el uno o más segmentos génicos variables de cadena ligera  $\kappa$  de inmunoglobulina humana comprenden al menos 40 segmentos génicos  $V_{\kappa}$ , al menos cinco segmentos génicos  $J_{\kappa}$  o una combinación de los mismos de un locus de cadena ligera  $\kappa$  de inmunoglobulina humana.

En aspectos particulares, el ratón o la rata modificados genéticamente comprende adicionalmente una secuencia de nucleótidos que codifica un segmento génico de cadena ligera ( $V_L$ ) de inmunoglobulina humana no reordenada y un segmento génico de cadena ligera ( $J_L$ ) de inmunoglobulina humana no reordenada. En algunos aspectos, la secuencia de nucleótidos codifica el segmento génico  $V$  de cadena ligera no reordenada y el segmento génico  $J$  de cadena ligera no ordenada está unido operativamente a una secuencia génica de región constante de cadena ligera de inmunoglobulina. En otros aspectos, la secuencia de nucleótidos codifica el segmento génico  $V$  de cadena ligera no reordenada y el segmento génico  $J$  de cadena ligera no reordenada está unido operativamente a una secuencia génica de región constante de cadena pesada de inmunoglobulina. En algunos aspectos, el segmento génico  $V$  de cadena ligera ( $V_L$ ) de inmunoglobulina humana no reordenada y el segmento génico  $J$  ( $J_L$ ) de inmunoglobulina humana no reordenada están unidos operativamente, en un locus endógeno de rata o de ratón, con un gen de región constante de cadena ligera de inmunoglobulina de rata o de ratón; por ejemplo, un gen de región constante de cadena ligera  $\kappa$  o  $\lambda$ .

En diversos aspectos, los segmentos génicos de región variable humana no reordenados (por ejemplo, segmentos génicos  $V_{\kappa}$  humanos) pueden reordenarse y codificar dominios variables humanos de un anticuerpo. En algunos aspectos, el ratón o la rata no comprende un segmento génico  $V_L$  endógeno. En algunos aspectos, los segmentos génicos  $V_{\kappa}$  humanos expresados por los ratones y las ratas se seleccionan del grupo que consiste en  $V_{\kappa 1-5}$ ,  $V_{\kappa 1-6}$ ,  $V_{\kappa 1-8}$ ,  $V_{\kappa 1-9}$ ,  $V_{\kappa 1-12}$ ,  $V_{\kappa 1-13}$ ,  $V_{\kappa 1-16}$ ,  $V_{\kappa 1-17}$ ,  $V_{\kappa 1-22}$ ,  $V_{\kappa 1-27}$ ,  $V_{\kappa 1-32}$ ,  $V_{\kappa 1-33}$ ,  $V_{\kappa 1-35}$ ,  $V_{\kappa 1-37}$ ,  $V_{\kappa 1-39}$ ,  $V_{\kappa 1D-8}$ ,  $V_{\kappa 1D-12}$ ,  $V_{\kappa 1D-13}$ ,  $V_{\kappa 1D-16}$ ,  $V_{\kappa 1D-17}$ ,  $V_{\kappa 1D-22}$ ,  $V_{\kappa 1D-27}$ ,  $V_{\kappa 1D-32}$ ,  $V_{\kappa 1D-33}$ ,  $V_{\kappa 1D-35}$ ,  $V_{\kappa 1D-37}$ ,  $V_{\kappa 1D-39}$ ,  $V_{\kappa 1D-42}$ ,  $V_{\kappa 1D-43}$ ,  $V_{\kappa 1-NL1}$ ,  $V_{\kappa 2-4}$ ,  $V_{\kappa 2-10}$ ,  $V_{\kappa 2-14}$ ,  $V_{\kappa 2-18}$ ,  $V_{\kappa 2-19}$ ,  $V_{\kappa 2-23}$ ,  $V_{\kappa 2-24}$ ,  $V_{\kappa 2-26}$ ,  $V_{\kappa 2-28}$ ,  $V_{\kappa 2-29}$ ,  $V_{\kappa 2-30}$ ,  $V_{\kappa 2-36}$ ,  $V_{\kappa 2-38}$ ,  $V_{\kappa 2-40}$ ,  $V_{\kappa 2D-10}$ ,  $V_{\kappa 2D-14}$ ,  $V_{\kappa 2D-18}$ ,  $V_{\kappa 2D-19}$ ,  $V_{\kappa 2D-23}$ ,  $V_{\kappa 2D-24}$ ,  $V_{\kappa 2D-26}$ ,  $V_{\kappa 2D-28}$ ,  $V_{\kappa 2D-29}$ ,  $V_{\kappa 2D-30}$ ,  $V_{\kappa 2D-36}$ ,  $V_{\kappa 2D-38}$ ,  $V_{\kappa 2D-40}$ ,  $V_{\kappa 3-7}$ ,  $V_{\kappa 3-11}$ ,  $V_{\kappa 3-15}$ ,  $V_{\kappa 3-20}$ ,  $V_{\kappa 3-25}$ ,  $V_{\kappa 3-31}$ ,  $V_{\kappa 3-34}$ ,  $V_{\kappa 3D-7}$ ,  $V_{\kappa 3D-11}$ ,  $V_{\kappa 3D-15}$ ,  $V_{\kappa 3D-20}$ ,  $V_{\kappa 3D-25}$ ,  $V_{\kappa 3D-31}$ ,  $V_{\kappa 3D-34}$ ,  $V_{\kappa 3-NL1}$ ,  $V_{\kappa 3-NL2}$ ,  $V_{\kappa 3-NL3}$ ,  $V_{\kappa 3-NL4}$ ,  $V_{\kappa 3-NL5}$ ,  $V_{\kappa 4-1}$ ,  $V_{\kappa 5-2}$ ,  $V_{\kappa 6-21}$ ,  $V_{\kappa 6D-21}$ ,  $V_{\kappa 6D-41}$  y  $V_{\kappa 7-3}$ . En algunos aspectos, los ratones o las ratas modificados genéticamente descritos en el presente documento, expresan todos los genes  $V_{\kappa}$  humanos funcionales. En algunos aspectos, los segmentos génicos  $V_{\kappa}$  humanos comprenden  $V_{\kappa 4-1}$ ,  $V_{\kappa 5-2}$ ,  $V_{\kappa 7-3}$ ,  $V_{\kappa 2-4}$ ,  $V_{\kappa 1-5}$  y  $V_{\kappa 1-6}$ . En un aspecto, los segmentos génicos  $V_{\kappa}$  comprenden  $V_{\kappa 3-7}$ ,  $V_{\kappa 1-8}$ ,  $V_{\kappa 1-9}$ ,  $V_{\kappa 2-10}$ ,  $V_{\kappa 3-11}$ ,  $V_{\kappa 1-12}$ ,  $V_{\kappa 1-13}$ ,  $V_{\kappa 2-14}$ ,  $V_{\kappa 3-15}$  y  $V_{\kappa 1-16}$ . En algunos aspectos, los segmentos génicos  $V_{\kappa}$  humanos comprenden  $V_{\kappa 1-17}$ ,  $V_{\kappa 2-18}$ ,  $V_{\kappa 2-19}$ ,  $V_{\kappa 3-20}$ ,  $V_{\kappa 6-21}$ ,  $V_{\kappa 1-22}$ ,  $V_{\kappa 1-23}$ ,  $V_{\kappa 2-24}$ ,  $V_{\kappa 3-25}$ ,  $V_{\kappa 2-26}$ ,  $V_{\kappa 1-27}$ ,  $V_{\kappa 2-28}$ ,  $V_{\kappa 2-29}$ , y  $V_{\kappa 2-30}$ . En algunos aspectos, los segmentos génicos  $V_{\kappa}$  humanos comprenden  $V_{\kappa 3-31}$ ,  $V_{\kappa 3-32}$ ,  $V_{\kappa 1-33}$ ,  $V_{\kappa 3-34}$ ,  $V_{\kappa 1-35}$ ,  $V_{\kappa 2-36}$ ,  $V_{\kappa 1-37}$ ,  $V_{\kappa 2-38}$ ,  $V_{\kappa 1-39}$  y  $V_{\kappa 2-40}$ . En diversos aspectos, el animal no humano comprende cinco segmentos génicos  $J_{\kappa}$  humanos, por ejemplo, los segmentos génicos  $J_{\kappa 1}$ ,  $J_{\kappa 2}$ ,  $J_{\kappa 3}$ ,  $J_{\kappa 4}$ , y  $J_{\kappa 5}$ . En aspectos específicos, los segmentos génicos  $V_{\kappa}$  comprenden segmentos génicos  $\kappa$  de inmunoglobulina humana contiguos que abarcan el locus de cadena ligera  $\kappa$  de inmunoglobulina humana desde  $V_{\kappa 4-1}$  a  $V_{\kappa 2-40}$ , y los segmentos génicos  $J_{\kappa}$  comprenden segmentos génicos contiguos que abarcan el locus de cadena ligera  $\kappa$  de inmunoglobulina humana de  $J_{\kappa 1}$  a  $J_{\kappa 5}$ . En algunos aspectos, el locus de cadena ligera de inmunoglobulina comprende dos segmentos génicos  $V_L$  humanos,  $V_{\kappa 1-39}$  y  $V_{\kappa 3-20}$ . En algunos aspectos, uno o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos  $V_L$  humanos y dos o más segmentos génicos  $J_L$  humanos, están presentes en un locus de cadena ligera endógena, por ejemplo, en un locus de cadena ligera kappa endógena. En algunos aspectos, el ratón o la rata modificado genéticamente, es un ratón que comprende un locus de cadena ligera  $\lambda$  funcional. En otros aspectos, el ratón comprende un locus de cadena ligera  $\lambda$  no funcional. En algunos aspectos, el uno o más segmentos génicos  $V_H$ ,  $D_H$  y  $V_H$  humanos, están unidos operativamente a una secuencia de región constante de cadena pesada de ratón o rata (es decir, el uno o más segmentos génicos  $V_L$  humanos y dos o más segmentos génicos  $J_L$  humanos, están presentes en un locus de cadena pesada endógena).

En algunos aspectos, un ratón o una rata modificado genéticamente como se describe en el presente documento expresa una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, produce una proteína de unión a antígeno que comprende un dominio variable de cadena pesada reordenada) y uno o más, dos o más, tres o más, cuatro o más, cinco o más etc. dominios variables de cadena ligera codificados por genes  $V_{\kappa}$  seleccionados del grupo que consiste en  $V_{\kappa 1-5}$ ,  $V_{\kappa 1-6}$ ,  $V_{\kappa 1-8}$ ,  $V_{\kappa 1-9}$ ,  $V_{\kappa 1-12}$ ,  $V_{\kappa 1-13}$ ,  $V_{\kappa 1-16}$ ,  $V_{\kappa 1-17}$ ,  $V_{\kappa 1-22}$ ,  $V_{\kappa 1-27}$ ,  $V_{\kappa 1-32}$ ,  $V_{\kappa 1-33}$ ,  $V_{\kappa 1-35}$ ,  $V_{\kappa 1-37}$ ,  $V_{\kappa 1-39}$ ,  $V_{\kappa 1D-8}$ ,  $V_{\kappa 1D-12}$ ,  $V_{\kappa 1D-13}$ ,  $V_{\kappa 1D-16}$ ,  $V_{\kappa 1D-17}$ ,  $V_{\kappa 1D-22}$ ,  $V_{\kappa 1D-27}$ ,  $V_{\kappa 1D-32}$ ,  $V_{\kappa 1D-33}$ ,  $V_{\kappa 1D-35}$ ,  $V_{\kappa 1D-37}$ ,  $V_{\kappa 1D-39}$ ,  $V_{\kappa 1D-42}$ ,  $V_{\kappa 1D-43}$ ,  $V_{\kappa 1-NL1}$ ,  $V_{\kappa 2-4}$ ,  $V_{\kappa 2-10}$ ,  $V_{\kappa 2-14}$ ,  $V_{\kappa 2-18}$ ,  $V_{\kappa 2-19}$ ,  $V_{\kappa 2-23}$ ,  $V_{\kappa 2-24}$ ,  $V_{\kappa 2-26}$ ,  $V_{\kappa 2-28}$ ,  $V_{\kappa 2-29}$ ,  $V_{\kappa 2-30}$ ,  $V_{\kappa 2-36}$ ,  $V_{\kappa 2-38}$ ,  $V_{\kappa 2-40}$ ,  $V_{\kappa 2D-10}$ ,  $V_{\kappa 2D-14}$ ,  $V_{\kappa 2D-18}$ ,  $V_{\kappa 2D-19}$ ,  $V_{\kappa 2D-23}$ ,  $V_{\kappa 2D-24}$ ,  $V_{\kappa 2D-26}$ ,  $V_{\kappa 2D-28}$ ,  $V_{\kappa 2D-29}$ ,  $V_{\kappa 2D-30}$ ,  $V_{\kappa 2D-36}$ ,  $V_{\kappa 2D-38}$ ,  $V_{\kappa 2D-40}$ ,  $V_{\kappa 3-7}$ ,  $V_{\kappa 3-11}$ ,  $V_{\kappa 3-15}$ ,  $V_{\kappa 3-20}$ ,  $V_{\kappa 3-25}$ ,  $V_{\kappa 3-31}$ ,  $V_{\kappa 3-34}$ ,  $V_{\kappa 3D-7}$ ,  $V_{\kappa 3D-11}$ ,  $V_{\kappa 3D-15}$ ,  $V_{\kappa 3D-20}$ ,  $V_{\kappa 3D-25}$ ,  $V_{\kappa 3D-31}$ ,  $V_{\kappa 3D-34}$ ,  $V_{\kappa 3-NL1}$ ,  $V_{\kappa 3-NL2}$ ,  $V_{\kappa 3-NL3}$ ,  $V_{\kappa 3-NL4}$ ,  $V_{\kappa 3-NL5}$ ,  $V_{\kappa 4-1}$ ,  $V_{\kappa 5-2}$ ,  $V_{\kappa 6-21}$ ,  $V_{\kappa 6D-21}$ ,  $V_{\kappa 6D-41}$  y  $V_{\kappa 7-3}$ .

En diversos aspectos, al menos uno de los segmentos génicos de región variable de cadena ligera (por ejemplo, segmentos génicos de región variable de cadena ligera humana) codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos, el dominio variable de cadena ligera como se describe en el presente documento, exhibe una

5 disminución en la semivida ( $t_{1/2}$ ) disociadora a un pH ácido, en comparación con un pH neutro, de al menos aproximadamente 2 veces, de al menos aproximadamente 3 veces, de al menos aproximadamente 4 veces, de al menos aproximadamente 5 veces, de al menos aproximadamente 10 veces, de al menos aproximadamente 15 veces, de al menos aproximadamente 20 veces, de al menos aproximadamente 25 veces o de al menos aproximadamente 30 veces. En algunos aspectos, la disminución en la  $t_{1/2}$  a un pH ácido, en comparación con un

10 pH neutro, es de aproximadamente 30 veces o mayor. En algunos aspectos, al menos uno de los segmentos génicos  $V_L$  comprende una sustitución de al menos un codón no de histidina, codificado por la secuencia del segmento  $V_L$  de la línea germinal humana correspondiente, con un codón de histidina. En algunos aspectos, la sustitución es de uno, dos, tres o cuatro codones (por ejemplo tres o cuatro codones). En algunos aspectos, la sustitución es en el codón o codones de la CDR3. En algunos aspectos, los segmentos génicos  $V_L$  humanos son

15 segmentos génicos Vk1-39 y Vk3-20 humanos, y cada uno de los segmentos génicos Vk1-39 y Vk3-20 humanos comprende una sustitución de al menos un codón no de histidina, codificado por un segmento génico  $V_L$  de la línea germinal humana correspondiente, con el codón de histidina. En algunos aspectos, cada uno de los segmentos génicos Vk1-39 y Vk3-20 humanos comprende una sustitución de tres o cuatro codones de histidina. En algunos aspectos, las tres o cuatro sustituciones están en la región CDR3. En algunos aspectos, la sustitución es de tres

20 codones no de histidina del segmento génico Vk1-39 humano, en el que la sustitución se diseña para expresar histidinas en las posiciones 106, 108 y 111. En algunos aspectos, la sustitución es de cuatro codones no de histidina del segmento génico Vk1-39 humano y la sustitución se diseña para expresar histidinas en las posiciones 105, 106, 108 y 111 (véanse, por ejemplo, los documentos US 2013/0247234A1 y WO 2013/138680). En algunos aspectos, la sustitución es de tres codones no de histidina del segmento génico Vk3-20 humano, y la sustitución se diseña para

25 expresar histidinas en las posiciones 105, 106 y 109. En otros aspectos adicionales, la sustitución es de cuatro codones no de histidina del segmento génico Vk3-20 humano y la sustitución se diseña para expresar histidinas en las posiciones 105, 106, 107 y 109. En algunos aspectos, el locus de cadena ligera de inmunoglobulina comprende uno o más segmentos génicos  $V_L$  humanos, pero menos que el número de segmentos de tipo silvestre, y uno o más, por ejemplo, dos o más, segmentos génicos  $J_L$  humanos, en el que cada uno de los segmentos génicos  $V_L$

30 humanos comprende al menos un codón de histidina que no está codificado por el segmento génico  $V_L$  de la línea germinal humana correspondiente. En diversos aspectos, el ratón o la rata que comprende los locus de inmunoglobulina modificados genéticamente como se describe en el presente documento, después de la estimulación con un antígeno de interés, expresan una proteína de unión a antígeno que comprende una secuencia de aminoácidos procedente de los segmentos génicos  $V_L$  humanos, en el que la proteína de unión a antígeno

35 conserva al menos un resto de histidina en una posición de aminoácido codificada por al menos un codón de histidina introducido en el segmento génico  $V_L$  humano. En algunos aspectos, el ratón o la rata expresa una población de proteínas de unión a antígeno en respuesta a un antígeno, en el que todas las proteínas de unión a antígeno en la población comprenden (a) dominios variables de cadena ligera de inmunoglobulina procedentes de una reordenación de los segmentos génicos  $V_L$  humanos y los segmentos génicos  $J_L$ , en el que al menos uno de los

40 segmentos génicos  $V_L$  humanos codifica uno o más codones de histidina que no están codificados por el segmento génico  $V_L$  de la línea germinal humana correspondiente, y (b) cadenas pesadas de inmunoglobulina que comprenden dominios variables de cadena pesada humana codificados por la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada.

45 Diversos aspectos abarcan secuencias de región constante de cadena ligera. En algunos aspectos, por ejemplo, una primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada) está unida operativamente a una secuencia génica de región constante de

50 cadena pesada y una segunda secuencia de nucleótidos que codifica el dominio variable de cadena ligera humana (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de cadena ligera de inmunoglobulina humana no reordenada) está unida operativamente a una secuencia génica de región constante de cadena ligera. En algunos aspectos, una primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada) está unida

55 operativamente a una secuencia génica de región constante de cadena ligera, y a una segunda secuencia de nucleótidos que codifica el dominio variable de cadena ligera humana (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de cadena ligera de inmunoglobulina humana no reordenada) está unida operativamente a una secuencia génica de región constante de cadena pesada. En diversos aspectos, la secuencia de región constante de cadena ligera unida operativamente a la secuencia de nucleótidos de región

60 variable de cadena pesada de inmunoglobulina humana reordenada es una secuencia de región constante de cadena ligera  $\kappa$  humana. En algunos aspectos, la secuencia de región constante de cadena ligera unida operativamente al dominio variable de cadena pesada reordenada es una secuencia de región constante de cadena ligera  $\kappa$  de ratón. En algunos aspectos, la secuencia de región constante de cadena ligera unida operativamente al dominio variable de cadena pesada reordenada es una secuencia de región constante de cadena ligera  $\lambda$  humana. En algunos aspectos,

65 algunos aspectos, la secuencia de región constante de cadena ligera unida operativamente al dominio variable de cadena pesada reordenada es una secuencia de región constante de cadena ligera  $\lambda$  humana. En algunos aspectos,

la secuencia de región constante de cadena ligera unida operativamente al dominio variable de cadena pesada reordenada es una secuencia de región constante de cadena ligera  $\lambda$  de ratón. En algunos aspectos, la secuencia de región constante de cadena ligera unida operativamente al dominio variable de cadena pesada reordenada es una secuencia de región constante de cadena ligera  $\lambda$  de rata.

5 En diversos aspectos, se describen ratones y ratas que comprenden un locus de inmunoglobulina modificado genéticamente que codifica un dominio variable de cadena pesada reordenada (es decir, en la que un locus de inmunoglobulina comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en el que el dominio variable de cadena pesada reordenada comprende una secuencia variable de cadena pesada ( $V_H$ ) que está unida operativamente, mediante un espaciador, a una secuencia de segmento J de cadena pesada ( $J_H$ ), en la que el espaciador comprende al menos un resto de aminoácido. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada no humana. En algunos aspectos, la secuencia génica de región constante de cadena pesada no humana es una secuencia génica de región constante de ratón o de rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada humana. En algunos aspectos, la región constante de cadena pesada comprende una secuencia seleccionada de un  $C_H1$ , una bisagra, un  $C_H2$ , un  $C_H3$  y una combinación de los mismos. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera no humana. En algunos aspectos, la secuencia génica de región constante de cadena ligera no humana es una secuencia génica de región constante de ratón o de rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera humana. En algunos aspectos, el espaciador tiene un solo resto de aminoácido. En algunos aspectos, el espaciador tiene dos restos de aminoácido. En algunos aspectos, el espaciador tiene tres restos de aminoácido. En algunos aspectos, el espaciador tiene cuatro restos de aminoácido. En algunos aspectos, el espaciador tiene cinco restos de aminoácido. En algunos aspectos, el espaciador tiene seis restos de aminoácido.

30 En otro aspecto, se proporcionan ratones y ratas modificados genéticamente y métodos para crear dichos animales, en los que los ratones y las ratas comprenden un locus de inmunoglobulina de cadena ligera universal ("ULC", del inglés *universal light chain*) funcional. En algunos aspectos, dichos ratones y ratas comprenden adicionalmente un locus de dominio variable de cadena pesada reordenada (es decir, un locus de inmunoglobulina de dominio variable de cadena pesada que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada). Una ULC es una cadena ligera común que puede utilizarse en un formato biespecífico que contiene una función, por ejemplo, una modificación que influye en la unión al FcRn para mejorar una semivida, por ejemplo, un formato biespecífico que comprende una cadena pesada que se une a un antígeno y una cadena ligera que se une al FcRn. Por ejemplo, los ratones modificados genéticamente, como se describe en el presente documento, están inmunizado con FcRN, para obtener anticuerpos que se unen al FcRN exclusivamente a través de las cadenas ligeras. Estas cadenas ligeras producidas por el ratón o la rata modificados genéticamente, se usan como ULC que ayudan al anticuerpo biespecífico a asociarse con un FcRn, ayudando de este modo a aumentar la semivida. El resto del anticuerpo (por ejemplo, o una segunda cadena ligera diferente, o una cadena pesada que se une a un antígeno diferente de FcRn) se selecciona para realizar una segunda función. Una CLU como se usa en los aspectos descritos en el presente documento, también puede utilizarse para generar secuencias de cadena variable de anticuerpo cuya diversidad da como resultado principalmente los procesos de mutación somática (por ejemplo, hipermutación), elucidando de este modo secuencias de cadena variable de anticuerpo cuya capacidad de unión a antígeno se beneficie de acontecimientos posgenómicos.

50 Diversos aspectos incluyen ratones y ratas modificados genéticamente que en sus genomas comprenden una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada, y adicionalmente en sus genomas comprenden un ácido nucleico que codifica un dominio variable de cadena ligera como se describe en el presente documento, clonado sobre una secuencia de ácido nucleico de región constante seleccionada de una región constante kappa, una región constante lambda, una región constante de cadena pesada (por ejemplo, seleccionada del grupo que consiste en un  $C_H1$ , una bisagra, un  $C_H2$ , un  $C_H3$  y una combinación de los mismos). En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena ligera está clonada sobre una primera secuencia de ácido nucleico de región constante de cadena pesada humana, y un segundo dominio variable de cadena ligera está clonado sobre una segunda secuencia de ácido nucleico de región constante de cadena pesada humana; en la que la primera y la segunda secuencia de ácido nucleico de región constante de cadena pesada humana son iguales, uniéndose el primer dominio variable de cadena ligera específicamente a un primer antígeno y uniéndose el segundo dominio variable de cadena ligera a un segundo antígeno. En estos aspectos, se forma un dímero de dos polipéptidos, en el que cada uno de los dominios variables de cadena ligera, fusionados con la región constante de cadena pesada, muestra distinta especificidad de unión con el antígeno.

65 En otro aspecto, se describe un ratón o una rata modificado genéticamente (por ejemplo, un ratón) que es capaz de producir una cadena ligera que se une a un receptor, o a otra fracción, que atraviese la barrera hematoencefálica, por ejemplo, al receptor de transferrina. Estudios previos han demostrado que los anticuerpos de baja afinidad dirigidos contra el receptor de transferrina, atravesarán la barrera hematoencefálica y se liberarán debido a la baja

afinidad. Por tanto, en algunos aspectos, los ratones o las ratas modificados genéticamente (por ejemplo, ratones) descritos en el presente documento, se utilizan para crear un anticuerpo de baja afinidad dirigido contra una fracción que sea capaz de atravesar la hematoencefálica (por ejemplo, un receptor de transferrina), en el que el anticuerpo de baja afinidad es biespecífico y comprende una segunda especificidad de unión contra una diana deseada (es decir, el anticuerpo se une a la fracción que atraviesa, y también se une a una diana diferente de la fracción que atraviesa).

En el presente documento se describen métodos para crear y utilizar ratones y ratas modificados genéticamente. Se describen métodos para colocar una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada, en unión operativa con una secuencia de ácido nucleico de región constante de cadena ligera o cadena pesada de inmunoglobulina, en el genoma de un ratón o una rata. En diversos aspectos, la secuencia de ácido nucleico de región constante es humana o no humana. En diversos aspectos, los métodos comprenden la creación de un ratón o una rata que adicionalmente comprende un locus de cadena ligera de inmunoglobulina que comprende uno o más segmentos génicos de región variable de cadena ligera humana, pero menos que el número de segmentos de tipo silvestre, por ejemplo, dos segmentos génicos  $V_k$  humanos y uno o más segmentos génicos  $J_k$  humanos, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera humana o no humana. En diversos aspectos, los métodos comprenden colocar las secuencias mencionadas anteriormente, en la línea germinal de un ratón o una rata, empleando, por ejemplo, tecnología transgénica, incluyendo, por ejemplo, células donantes pluripotentes o totipotentes modificadas (por ejemplo, células ES o células iPS) con embriones hospedadores, células germinativas (por ejemplo ovocitos), etc. Por tanto, los aspectos incluyen un locus de cadena pesada de inmunoglobulina no humana en un genoma de una célula germinativa no humana que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia génica de región constante de cadena pesada, en la que la secuencia génica de región constante comprende una secuencia no humana, una secuencia humana, o una combinación de las mismas. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, está unida operativamente a una secuencia génica de región constante de inmunoglobulina no humana endógena. En algunos aspectos, la secuencia génica de región constante de inmunoglobulina no humana endógena, es una secuencia génica de región constante de cadena pesada de rata o de ratón.

En diversos aspectos, se describe un método para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente, en el que el método comprende: (a) modificar un genoma de una rata o un ratón para delecionar y hacer no funcionales, segmentos génicos V, D y J de cadena pesada de inmunoglobulina funcional endógena; y (b) colocar en el genoma, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. En un aspecto de este tipo, se describe un método para crear un ratón o una rata que exprese una sola cadena pesada de inmunoglobulina de una secuencia génica de cadena pesada reordenada en la línea germinal del ratón o rata, comprendiendo el método una etapa de modificar genéticamente un ratón o una rata de tal manera que toda la población de linfocitos B maduros, que expresa el anticuerpo, exprese una cadena pesada procedente de (i) un solo segmento génico  $V_H$ ; (ii) un espaciador de aminoácidos de uno, dos, tres, cuatro, cinco o seis aminoácidos y (iii) un solo segmento génico  $J_H$ . En algunos aspectos, el método comprende inactivar o reemplazar un locus variable de inmunoglobulina de cadena pesada endógena con un solo gen de cadena pesada reordenada como se describe en el presente documento.

En otros aspectos, se describen métodos para crear un ratón o una rata que comprenda un locus de cadena pesada de inmunoglobulina genéticamente modificada, comprendiendo dichos métodos: (a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcionales, segmentos génicos V, D y J de cadena pesada de inmunoglobulina funcional endógena y (b) colocar en el genoma, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada. Todos los segmentos génicos  $V_H$ , D y  $J_H$  funcionales endógenos se delecionan del locus de cadena pesada de inmunoglobulina del ratón o de la rata o se vuelven no funcionales (por ejemplo, mediante la inserción de una secuencia de nucleótidos (por ejemplo, una secuencia de nucleótidos exógena en el locus de inmunoglobulina o mediante una reordenación no funcional, o una inversión, de segmentos  $V_H$ , D,  $J_H$  endógenos). En algunos aspectos, el método comprende la inserción de una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, una secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada) en una localización endógena (es decir dirigida a donde se localiza la secuencia de nucleótidos en un animal no humano de tipo silvestre). En algunos aspectos, por ejemplo, aproximadamente el 80 % o más, aproximadamente el 85 % o más, aproximadamente el 90 % o más, aproximadamente el 95 % o más, aproximadamente el 96 % o más, aproximadamente el 97 % o más, aproximadamente el 98 % o más o aproximadamente el 99 % o más de todos los segmentos génicos V, D o J funcionales endógenos están delecionados o se han hecho no funcionales. En algunos aspectos, por ejemplo, al menos el 95 %, 96 %, 97 %, 98 % o 99 % de los segmentos génicos V, D o J de cadena pesada funcional endógena están delecionados o se han hecho no funcionales.

En otro aspecto, se describen métodos para crear un ratón o una rata que comprenda un locus de inmunoglobulina modificado genéticamente, que comprenden: (a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcionales, segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena y (b) colocar en un locus de cadena ligera de inmunoglobulina endógena, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, una secuencia de nucleótidos que codifica un

dominio variable de cadena pesada reordenada), en la que la secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera. En algunos aspectos, el locus de inmunoglobulina modificado genéticamente está presente en el genoma de la línea germinal del ratón o de la rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$ . En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$  de ratón o rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$ . En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$  de ratón o rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$  humana.

En otro aspecto, se describen métodos para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente, que comprenden: (a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcionales: (i) segmentos génicos V, D y J de cadena pesada de inmunoglobulina funcional endógena y (ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena y (b) colocar en el genoma: (i) una primera secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera y (ii) una segunda secuencia de nucleótidos que codifica un dominio variable de cadena ligera de inmunoglobulina humana (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada. En algunos aspectos, el locus de inmunoglobulina modificado genéticamente está presente en el genoma de la línea germinal del ratón o de la rata. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$ . En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$  de ratón o de rata. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$  humana. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$ . En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$  de ratón o de rata. En algunos aspectos, la primera secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera  $\lambda$  humana. En algunos aspectos, el dominio variable de cadena ligera de inmunoglobulina humana es un dominio variable de cadena ligera  $\kappa$ . Por tanto, en algunos aspectos, la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera kappa humana. En algunos aspectos, el dominio variable de cadena ligera de inmunoglobulina humana es un dominio variable de cadena ligera  $\lambda$ . Por tanto, en algunos aspectos, la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera lambda humana. En algunos aspectos, la secuencia génica de región constante de cadena pesada es una secuencia génica de región constante de cadena pesada de inmunoglobulina no humana. En algunos aspectos, la secuencia génica de región constante de cadena pesada de inmunoglobulina no humana es una secuencia génica de región constante de cadena pesada de ratón o de rata.

En otro aspecto, se describen métodos para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente, que comprenden: (a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcionales: (i) segmentos génicos V, D y J de cadena pesada de inmunoglobulina funcional endógena y (ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena y (b) colocar en el genoma: (i) una primera secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera la secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada, y (ii) una segunda secuencia de nucleótidos que codifica un dominio variable de cadena ligera de inmunoglobulina humana (es decir, en la que la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera. En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera  $\kappa$ . En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera  $\kappa$  de ratón o de rata. En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera  $\kappa$ . En algunos aspectos, la secuencia génica de región constante de cadena ligera

es una secuencia génica de región constante de cadena ligera  $\lambda$ . En algunos aspectos, la secuencia génica de la región constante de cadena ligera es una secuencia génica de región constante de cadena ligera  $\lambda$  de ratón o de rata. En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera  $\lambda$  humana. En algunos aspectos, el dominio variable de cadena ligera de inmunoglobulina humana es un dominio variable de cadena ligera  $\kappa$ . En algunos aspectos, el dominio variable de cadena ligera de inmunoglobulina humana es un dominio variable de cadena ligera  $\lambda$ . En algunos aspectos, la secuencia génica de la región constante de cadena pesada es una secuencia génica de región constante de cadena pesada de inmunoglobulina no humana. En algunos aspectos, la secuencia génica de región constante de cadena pesada de inmunoglobulina no humana es una secuencia génica de región constante de cadena pesada de ratón o de rata.

En otro aspecto, se describe un método para crear un ratón o una rata que comprende un locus de cadena pesada de inmunoglobulina modificada genéticamente que comprenden: (a) modificar un genoma de un ratón o una rata para deletionar o hacer no funcional, segmentos génicos V, D y J de cadena pesada de inmunoglobulina funcional endógena y (b) colocar en el genoma una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, en la que la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada comprende una secuencia de segmento génico V de cadena pesada ( $V_H$ ) que está unida operativamente, mediante un espaciador, a una secuencia de segmento génico J de cadena pesada ( $J_H$ ), en la que el espaciador comprende al menos un resto de aminoácido. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada de inmunoglobulina no humana. En algunos aspectos, la secuencia génica de región constante de cadena pesada de inmunoglobulina no humana es una secuencia génica de región constante de cadena pesada de inmunoglobulina de ratón o de rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera de inmunoglobulina no humana. En algunos aspectos, la secuencia génica de región constante de cadena ligera de inmunoglobulina no humana es una secuencia génica de región constante de cadena ligera de inmunoglobulina de ratón o de rata. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está presente en una localización endógena (es decir, en la que la secuencia de nucleótidos se localiza en un ratón o una rata de tipo silvestre). En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está presente dentro de su locus endógeno, por ejemplo, dentro de un locus variable de inmunoglobulina, en el que el locus endógeno se coloca o traslada a una localización diferente en el genoma). En algunos aspectos, los espaciadores tienen un solo resto de aminoácido. En algunos aspectos, los espaciadores tienen dos restos de aminoácido. En algunos aspectos, los espaciadores tienen tres restos de aminoácido. En algunos aspectos, los espaciadores tienen cuatro restos de aminoácido. En algunos aspectos, los espaciadores tienen cinco restos de aminoácido. En algunos aspectos, los espaciadores tienen seis restos de aminoácido. En algunos aspectos, las secuencias de nucleótidos codifican dos copias, tres copias, cuatro copias o más de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia génica de dominio constante de cadena pesada. En algunos aspectos, la secuencia de nucleótidos codifica una pluralidad de copias de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia génica de dominio constante de cadena pesada.

Se describen métodos para crear un ratón o una rata que comprenden: (a) modificar un genoma de un ratón o una rata para deletionar o hacer no funcionales (i) segmentos génicos  $V_H$ ,  $D_H$  y/o  $J_H$ , de cadena pesada de inmunoglobulina funcional endógena y (ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena; y (b) colocar (i) una secuencia de ácido nucleico de región variable de cadena pesada reordenada en un locus de cadena pesada, en el que la secuencia de ácido nucleico de región variable de cadena pesada reordenada comprende una secuencia de segmento génico V de cadena pesada ( $V_H$ ) que está unida operativamente, mediante un espaciador, a una secuencia de segmento génico J de cadena pesada ( $J_H$ ), en la que el espaciador comprende al menos un resto de aminoácido; y (ii) uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, (por ejemplo, dos segmentos génicos  $V_k$  humanos, y al menos un segmento génico  $J_k$  humano), unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera humana o no humana. En algunos aspectos, al menos uno de los segmentos génicos de región variable de cadena ligera codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente.

En algunos aspectos, se describe un método para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente que comprende:

(a) modificar un genoma de un ratón o una rata para deletionar o hacer no funcionales, segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena; y

(b) colocar en el genoma del ratón o de la rata una secuencia de nucleótidos de región variable de cadena pesada humana reordenada, en unión operativa con una secuencia de nucleótidos de región constante de cadena ligera.

En algunos aspectos, la región constante de cadena ligera es una región constante de rata o ratón, por ejemplo, una región constante  $C_k$  de rata o ratón.

5 En otro aspecto, se describe un método para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente, que comprende:

(a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcionales:

10 (i) segmentos génicos V, D y/o J de cadena pesada de inmunoglobulina funcional endógena; y

(ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena; y

(b) colocar en el genoma del ratón o de la rata:

15 (i) una primera secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada (es decir, en la que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera, y

20 (ii) una segunda secuencia de nucleótidos que codifica un dominio variable de cadena ligera no humana (es decir, en la que la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada.

25 En algunos aspectos, el ratón o la rata es un ratón. En algunos aspectos, la región constante de cadena ligera es una región constante de una rata o de un ratón, por ejemplo, una región constante  $C_k$  de una rata o de un ratón. En algunos aspectos, la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada de ratón o de rata seleccionada de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y una combinación de los mismos. En algunos aspectos, la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada humana seleccionada de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  o una combinación de los mismos.

30 En otro aspecto, se describe un método para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente, que comprende:

(a) modificar un genoma de una rata o un ratón para delecionar o hacer no funcional:

40 (i) segmentos génicos V, D y/o J de cadena pesada de inmunoglobulina funcional endógena, y

(ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena; y

(b) colocar en el genoma del ratón o de la rata:

45 (i) una primera secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada (es decir, en la que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada; y

50 (ii) una segunda secuencia de nucleótidos que codifica un dominio variable de cadena ligera (es decir, en la que la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera.

55 En otro aspecto, se describe un método para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente, que comprende:

(a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcional:

60 (i) segmentos génicos V, D y/o J de cadena pesada de inmunoglobulina funcional endógena, y

(ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena; y

(b) colocar en el genoma del ratón o de la rata:

65 (i) un primer alelo que comprende:

(1) una primera secuencia de nucleótidos que codifica un dominio variable de cadena pesada reordenada (es decir, en la que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), unida operativamente a una secuencia génica de región constante de cadena pesada; y

(2) una segunda secuencia de nucleótidos que codifica un dominio variable de cadena ligera (es decir, en la que la segunda secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada), unida operativamente a una secuencia génica de región constante de cadena ligera.

(ii) un segundo alelo que comprende

(1) una tercera secuencia de nucleótidos que codifica un dominio variable de cadena ligera (es decir, en la que la tercera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada), unida operativamente a una secuencia génica de región constante de cadena pesada; y

(2) una cuarta secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada (es decir, en la que la cuarta secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), unida operativamente a una secuencia génica de región constante de cadena ligera.

En otro aspecto, se describe un método para crear un ratón o una rata que comprende un locus de cadena pesada de inmunoglobulina modificado genéticamente que comprende: (a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcionales, segmentos génicos V, D y/o J de cadena pesada de inmunoglobulina funcional endógena y (b) colocar en el genoma la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, en la que la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada comprende una secuencia del segmento génico V de cadena pesada ( $V_H$ ) que está unida operativamente, mediante un espaciador, a una secuencia de segmento génico J de cadena pesada ( $J_H$ ), en la que el espaciador comprende al menos un resto de aminoácido.

En otro aspecto, se describe un método para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente que comprende:

(a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcional:

(i) segmentos génicos V, D y/o J de cadena pesada de inmunoglobulina funcional endógena, y

(ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena; y

(b) colocar en el genoma del ratón o de la rata:

(i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada en unión operativa con una secuencia de nucleótidos de región constante de cadena pesada; y

(ii) uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, en unión operativa con una secuencia de ácido nucleico de región constante de cadena ligera.

En algunos aspectos, la rata o el ratón es un ratón. En algunos aspectos, la región constante de cadena ligera es una región constante de rata o ratón, por ejemplo, una región constante  $C_k$  de una rata o un ratón. En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena pesada humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada de ratón o de rata de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de región variable de cadena pesada reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada humana seleccionada de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y una combinación de los mismos.

En otro aspecto, se describe un método para crear un ratón o una rata que comprende un locus de inmunoglobulina modificado genéticamente, que comprende:

(a) modificar un genoma de un ratón o una rata para delecionar o hacer no funcional:

(i) segmentos génicos V, D y/o J de cadena pesada de inmunoglobulina funcional endógena, y

(ii) segmentos génicos V y J de cadena ligera de inmunoglobulina funcional endógena; y

(b) colocar en el genoma del ratón o de la rata:

(i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada en un unión operativa con una secuencia de nucleótidos de región constante de cadena ligera; y

(ii) uno o más segmentos génicos  $V_L$  y  $J_L$  de región variable de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, en unión operativa con una secuencia de ácido nucleico de región constante de cadena pesada.

En algunos aspectos, la rata o el ratón es un ratón. En algunos aspectos, la región constante de cadena ligera es una región constante de una rata o un ratón, por ejemplo, una región constante  $C_k$  de una rata o un ratón.

En otro aspecto, se describen secuencias de ácido nucleico que codifican un dominio variable de cadena pesada reordenada (es decir, secuencias de nucleótidos que son secuencias de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada; es decir una secuencia de nucleótidos de VDJ de cadena pesada de región variable pre-reordenada). En algunos aspectos, la secuencia de ácido nucleico procede de un segmento o secuencia génica V, D y J humana. En algunos aspectos, la secuencia de ácido nucleico procede de un segmento V de la línea germinal humana, de un segmento D de la línea germinal humana y de un segmento J de la línea germinal humana. En algunos aspectos, el segmento  $V_H$  humano corresponde a variantes observadas en la población humana. El segmento V humano es  $V_H3-23$  o una variante polimórfica del mismo. En algunos aspectos, la secuencia de ácido nucleico comprende un segmento D humano que no es autorreactivo (no inmunógeno) en el ratón o rata. En algunos aspectos, la secuencia de ácido nucleico comprende un segmento D humano que puede expresarse en una secuencia variable de cadena pesada de un linfocito B maduro de un ratón. En algunos aspectos, la secuencia de ácido nucleico comprende adicionalmente una secuencia génica de región constante de cadena pesada de un animal humano o no humano, seleccionada de un  $C_H1$ , una bisagra, un  $C_H2$ , un  $C_H3$  o una combinación de los mismos. En aspectos específicos, el ácido nucleico comprende una secuencia génica de región constante que comprende un  $C_H1$ , una bisagra, un  $C_H2$  y un  $C_H3$ . En diversos aspectos, la secuencia de ácido nucleico comprende un segmento génico J humano seleccionado del grupo que consiste en  $J_H1$ ,  $J_H2$ ,  $J_H3$ ,  $J_H4$ ,  $J_H5$ ,  $J_H6$ , y una variante polimórfica de los mismos. La secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada codifica la secuencia de  $V_H3-23/GY/J_H4-4$  (SEQ ID NO: 137) humana. En algunos aspectos, la secuencia de ácido nucleico codifica un dominio variable de cadena pesada reordenada que comprende la secuencia de  $V_H3-23/X_1X_2/J$  humana (en la que  $X_1$  y  $X_2$  es cualquier aminoácido). En algunos aspectos,  $X_1$  es Gly y  $X_2$  es Tyr. En algunos aspectos, la secuencia de ácido nucleico codifica un dominio variable de cadena pesada reordenada que comprende la secuencia de  $V_H3-23/X_1X_2/J_H4-4$  humana (en la que  $X_1$  y  $X_2$  es cualquier aminoácido). En algunos aspectos,  $X_2$  es un aminoácido que comprende un grupo fenilo. En aspectos específicos,  $X_2$  se selecciona de Tyr y Phe. En algunos aspectos, la secuencia de ácido nucleico comprende adicionalmente una secuencia génica de región constante de cadena ligera de un animal humano o no humano

En otro aspecto, se describe una construcción de ácido nucleico que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, una secuencia VDJ de cadena pesada pre-reordenada) como se describe en el presente documento. En algunos aspectos, la construcción de ácido nucleico se diseña de tal manera que la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada de un animal humano o no humano. En algunos aspectos, la construcción de ácido nucleico contiene dos copias, tres copias, cuatro copias o más de la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia génica de región constante de cadena pesada. En algunos aspectos, la construcción de ácido nucleico es un vector diana. En algunos aspectos, el vector diana comprende un gen Adam6a, un gen Adam6b o ambos, para impedir que se produzcan problemas de fertilidad asociados a la delección de los genes Adam6a/6b (véase, por ejemplo, el documento US 2012-0322108A1). En algunos aspectos, los genes Adam6a y Adam6b se colocan cadena arriba en dirección 5' de la unidad transcripcional de la secuencia de cadena pesada universal. En algunos aspectos, el vector diana comprende un casete de selección flanqueado por sitios de recombinación. En algunos aspectos, el vector diana comprende uno o más sitios de recombinación específicos de sitio (por ejemplo un sitio loxP o FRT).

En otro aspecto, se describen métodos para obtener una secuencia de aminoácidos de región variable de cadena ligera ( $V_L$ ) capaz de unirse a un antígeno independientemente de una secuencia de aminoácidos de región variable de cadena pesada, que comprende: (a) inmunizar a un ratón o a una rata modificado genéticamente como se describe en el presente documento (por ejemplo, un ratón o una rata modificado genéticamente que comprende una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada en unión operativa con una secuencia de ácido nucleico de región constante de cadena pesada o ligera) con un antígeno de interés, en el que el ratón o la rata genera una respuesta inmunitaria contra el antígeno; y (b) obtener una secuencia de ácido nucleico de cadena ligera (VJ) reordenada de un dominio variable de cadena ligera que se une específicamente al antígeno de una célula (por ejemplo, un linfocito B maduro) del ratón o de la rata modificado genéticamente. En diversos aspectos, se describen las regiones variables de cadena ligera producidas por dichos métodos.

En algunos aspectos, se describen métodos para obtener una secuencia de ácido nucleico que codifique un dominio de región variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de una región variable de cadena pesada, que comprende: (a) inmunizar a una rata o a un ratón con un antígeno de interés o con un inmunógeno del mismo, en el que la rata o el ratón comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada, (b) permitir que el ratón o la rata genere una respuesta inmunitaria, (c) aislar del ratón o de la rata inmunizados, una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno, y (d) obtener de la célula, una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno. En algunos aspectos, la secuencia génica de región constante de cadena pesada es una secuencia génica de región constante de cadena pesada de ratón o de rata. En algunos aspectos, la secuencia génica de región constante de cadena pesada es una secuencia génica de región constante de cadena pesada humana. En algunos aspectos, el dominio variable de cadena pesada reordenada expresado por el locus modificado genéticamente no es autorreactivo, es decir, no inmunógeno al ratón o la rata. En algunos aspectos, el ratón o la rata comprende, en su genoma, dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre. En algunos aspectos, los segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana están unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera. En algunos aspectos, la etapa de aislamiento (c) se realiza mediante separación de células activadas por fluorescencia (FACS) o citometría de flujo. En algunos aspectos, la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno es un linfocito. En algunos aspectos, el linfocito comprende linfocitos citolíticos naturales, linfocitos T o linfocitos B. En algunos aspectos, el método comprende adicionalmente una etapa (c)' de fusionar el linfocito con una célula cancerosa. En determinados aspectos, la célula cancerosa es una célula de mieloma.

Por tanto, en diversos aspectos, se describen métodos para obtener una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprenden:

(a) inmunizar a un ratón o a una rata con un antígeno de interés o con un inmunógeno del mismo, en el que el ratón o la rata comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y (ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana no reordenada unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera;

(b) permitir que el ratón o la rata genere una respuesta inmunitaria;

(c) aislar del ratón o de la rata inmunizado una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

(d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

En algunos aspectos, la etapa de aislamiento (c) se realiza mediante separación de células activadas por fluorescencia (FACS) o citometría de flujo. En algunos aspectos, la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno es un linfocito. En aspectos particulares, el linfocito comprende linfocitos citolíticos naturales, linfocitos T o linfocitos B. En algunos aspectos, los métodos comprenden adicionalmente una etapa (c)' de fusionar el linfocito con una célula cancerosa. En aspectos particulares, la célula cancerosa es una célula de mieloma. En algunos aspectos, la secuencia de ácido nucleico de (d) se fusiona con una secuencia de ácido nucleico que codifica una secuencia de ácido nucleico de región constante de inmunoglobulina. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa humana o una secuencia lambda humana. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa de ratón o una secuencia lambda de ratón. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa de rata o una secuencia lambda de rata. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia humana seleccionada de un  $C_H1$ , una bisagra, un  $C_H2$ , un  $C_H3$  y una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia de ratón o de rata seleccionada de un  $C_H1$ , una bisagra, un  $C_H2$ , un  $C_H3$  y una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de (d) comprende una o más sustituciones o inserciones de codones de histidina que proceden del segmento génico  $V_L$  no reordenado en el genoma del ratón o de la rata.

En diversos aspectos, se describen métodos para obtener una secuencia de ácido nucleico que codifica un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprenden:

(a) inmunizar a un ratón o a una rata con un antígeno de interés o con un inmunógeno del mismo, en el que el ratón o la rata comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera; y (ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana no reordenada, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena pesada;

(b) permitir que el ratón o la rata genere una respuesta inmunitaria;

(c) aislar del ratón o de la rata inmunizado una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

(d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

En algunos aspectos, la etapa de aislamiento (c) se realiza mediante separación de células activadas por fluorescencia (FACS) o citometría de flujo. En algunos aspectos, la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno es un linfocito. En aspectos particulares, el linfocito comprende linfocitos citolíticos naturales, linfocitos T o linfocitos B. En algunos aspectos, el método comprende adicionalmente una etapa (c)' de fusionar el linfocito con una célula cancerosa. En aspectos particulares, la célula cancerosa es una célula de mieloma. En algunos aspectos, la secuencia de ácido nucleico de (d) se fusiona con una secuencia de ácido nucleico que codifica una secuencia de ácido nucleico de región constante de inmunoglobulina. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia lambda humana o una secuencia kappa humana. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia lambda de ratón o una secuencia kappa de ratón. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa de rata o una secuencia lambda de rata. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia humana seleccionada de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y de una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia de rata o de ratón seleccionada de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y de una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de (d) comprende una o más sustituciones o inserciones de codones de histidina que proceden del segmento génico  $V_L$  no reordenado en el genoma de la rata o ratón.

En algunos aspectos, se describen métodos para obtener una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprenden:

(a) inmunizar a un ratón o a una rata con un antígeno de interés o con un inmunógeno del mismo, en el que el ratón o la rata comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y (ii) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera;

(b) permitir que el ratón o la rata genere una respuesta inmunitaria;

(c) aislar del ratón o de la rata inmunizados, una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

(d) obtener de la célula una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

En algunos aspectos, la etapa de aislamiento (c) se realiza mediante separación de células activadas por fluorescencia (FACS) o citometría de flujo. En algunos aspectos, la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno es un linfocito. En aspectos particulares, el linfocito comprende linfocitos citolíticos naturales, linfocitos T o linfocitos B. En algunos aspectos, los métodos comprenden adicionalmente una etapa (c)' de fusionar el linfocito con una célula cancerosa. En aspectos particulares, la célula cancerosa es una célula de mieloma. En algunos aspectos, la secuencia de ácido nucleico de (d) se fusiona con una secuencia de ácido nucleico que codifica una secuencia de ácido nucleico de región constante de inmunoglobulina. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa humana o una secuencia lambda humana. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa de ratón o una secuencia lambda de ratón. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa de rata o una secuencia lambda de rata. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia humana seleccionada de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y de

una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de región constante de cadena pesada es una secuencia de ratón o de rata seleccionada de un C<sub>H</sub>1, una bisagra, un C<sub>H</sub>2, un C<sub>H</sub>3 o de una combinación de los mismos. En algunos aspectos, la secuencia de ácido nucleico de (d) comprende una o más sustituciones o inserciones de codones de histidina que proceden del segmento génico V<sub>L</sub> no reordenado en el genoma del ratón o de la rata.

En algunos aspectos, se describen métodos para obtener una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera (V<sub>L</sub>) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprenden:

- (a) inmunizar a un ratón o a una rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el ratón o la rata comprende en su genoma una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada;
- (b) permitir que el ratón o la rata genere una respuesta inmunitaria;
- (c) recoger, de la rata o del ratón inmunizado, un linfocito (por ejemplo, un linfocito B);
- (d) fusionar el linfocito con una célula de mieloma para formar una célula de hibridoma; y
- e) obtener de la célula de hibridoma, una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera (dominio V<sub>L</sub>) que pueda unirse al antígeno.

En otro aspecto, se describen métodos para obtener una secuencia de ácido nucleico que codifique una secuencia de ácido nucleico de dominio variable de cadena ligera (V<sub>L</sub>) de una inmunoglobulina de una cadena ligera de inmunoglobulina, capaz de unirse a un antígeno independientemente de una región variable de cadena pesada, que comprenden:

- (a) inmunizar a un ratón o a una rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el ratón o la rata comprende en su genoma una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada;
- (b) permitir que el ratón o la rata genere una respuesta inmunitaria;
- (c) identificar un linfocito (por ejemplo, un linfocito B) de la rata o del ratón inmunizado, que exprese una secuencia de aminoácidos de V<sub>L</sub> que se una al antígeno independientemente de una región variable de cadena pesada; y,
- (d) clonar una secuencia de ácido nucleico que codifique la secuencia de aminoácidos de V<sub>L</sub> de (c) del linfocito de (c).

En otro aspecto, se describen métodos para obtener una secuencia de aminoácidos de región variable de cadena ligera (V<sub>L</sub>) de inmunoglobulina capaz de unirse a un antígeno independientemente de una región variable de cadena pesada, que comprenden:

- (a) inmunizar a un ratón o a una rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el ratón o la rata comprende en su genoma (i) una primera secuencia de nucleótidos que codifique un dominio variable de cadena pesada reordenada (es decir, en el que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera; y (ii) una segunda secuencia de nucleótidos que codifique un dominio variable de cadena ligera humana o no humana (es decir, en la que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada;
- (b) permitir que el ratón o la rata genere una respuesta inmunitaria;
- (c) recoger, de la rata o del ratón inmunizado, un linfocito (por ejemplo, un linfocito B);
- (d) fusionar el linfocito con una célula de mieloma para formar una célula de hibridoma; y

e) obtener de la célula de hibridoma, una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

5 En otro aspecto, se describen métodos para obtener una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprenden:

10 (a) inmunizar a un ratón o a una rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el ratón o la rata comprende en su genoma (i) una primera secuencia de nucleótidos que codifique un dominio variable de cadena pesada reordenada (es decir, en la que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera; y (ii) una segunda secuencia de nucleótidos que codifique un dominio variable de cadena ligera humana o no humana (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena pesada;

20 (b) permitir que el ratón o la rata genere una respuesta inmunitaria;

(c) identificar un linfocito (por ejemplo, un linfocito B) de la rata o del ratón inmunizado, que exprese una secuencia de aminoácidos de  $V_L$  que se una al antígeno independientemente de una región variable de cadena pesada; y,

25 (d) clonar una secuencia de ácido nucleico que codifique la secuencia de aminoácidos de  $V_L$  de (c) del linfocito de (c).

30 En otro aspecto, se describen métodos para obtener una secuencia de aminoácidos de región variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de una región variable de cadena pesada, que comprenden:

35 (a) inmunizar a un ratón o a una rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el ratón o la rata comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y (ii) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre;

40 (b) permitir que el ratón o la rata genere una respuesta inmunitaria;

(c) recoger, de la rata o del ratón inmunizado, un linfocito (por ejemplo, un linfocito B);

(d) fusionar el linfocito con una célula de mieloma para formar una célula de hibridoma; y

45 e) obtener de la célula de hibridoma, una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera (dominio  $V_L$ ) que pueda unirse al antígeno.

50 En otro aspecto, se describen métodos para obtener una secuencia de ácido nucleico de región variable de cadena ligera ( $V_L$ ) de inmunoglobulina de una cadena ligera de inmunoglobulina capaz de unirse a un antígeno independientemente de una región variable de cadena pesada, que comprenden:

55 (a) inmunizar a un ratón o a una rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el ratón o la rata comprende en su genoma (i) una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y (ii) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre;

60 (b) permitir que el ratón o la rata genere una respuesta inmunitaria;

(c) identificar un linfocito (por ejemplo, un linfocito B) de la rata o del ratón inmunizado que exprese una secuencia de aminoácidos  $V_L$  que se una al antígeno independientemente de una región variable de cadena pesada; y,

65 (d) clonar una secuencia de ácido nucleico que codifique la secuencia de aminoácidos de  $V_L$  de (c) del linfocito de (c).

En diversos aspectos, el dominio variable de cadena ligera descrito en el presente documento, es un dominio efector variable de cadena ligera. En algunos aspectos, el dominio variable efector de cadena ligera se une específicamente al FcRn para mejorar la semivida de los anticuerpos multiespecíficos. Por ejemplo, un anticuerpo biespecífico que comprende un dominio variable de cadena pesada que se une a un antígeno y un dominio variable de cadena ligera que se une al FcRn. En algunos aspectos, los ratones modificados genéticamente como se describe en el presente documento, están inmunizados con FcRn, para obtener anticuerpos que se unen exclusivamente al FcRn a través de las cadenas ligeras. Estas cadenas ligeras producidas por la rata o el ratón modificado genéticamente, se utilizan como cadenas ligeras universales o comunes que ayudan al anticuerpo biespecífico a asociarse a un FcRn, ayudando de esta manera a aumentar la semivida. El resto del anticuerpo (por ejemplo, una segunda cadena ligera diferente, o una cadena pesada que se une a un antígeno diferente de FcRn) se selecciona para realizar una segunda función.

En aspectos adicionales, se describe un locus de inmunoglobulina modificado genéticamente que puede obtenerse mediante cualquiera de los métodos descritos en el presente documento. En diversos aspectos, también se describen regiones variables de cadena ligera producidas por los métodos descritos en el presente documento y la secuencia de ácido nucleico que codifica dichas regiones variables de cadena ligera.

En algunos aspectos, se describe un locus de inmunoglobulina en un genoma de línea germinal de una rata o un ratón que comprende (1) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que está unida operativamente a una secuencia génica de región constante de cadena pesada, y (2) una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada que está unida operativamente a una secuencia génica de región constante de cadena ligera. En algunos aspectos, se describe un locus de inmunoglobulina en un genoma de línea germinal de una rata o un ratón que comprende (1) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que está unida operativamente a una secuencia génica de región constante de cadena ligera, y (2) una secuencia de nucleótidos de región variable de cadena ligera de inmunoglobulina humana no reordenada que está unida operativamente a una secuencia génica de región constante de cadena pesada. En algunos aspectos, se describe un locus de inmunoglobulina en un genoma de línea germinal de una rata o un ratón que comprende (1) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que está unida operativamente a una secuencia génica de región constante de cadena pesada, y (2) una secuencia de nucleótidos que codifica dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre. En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera  $\kappa$ . En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera  $\lambda$ . En algunos aspectos, la secuencia génica de región constante de cadena ligera es una secuencia génica de región constante de cadena ligera de rata o ratón. En algunos aspectos, la secuencia de nucleótidos de región variable de cadena ligera es una secuencia génica de región variable de cadena ligera  $\kappa$ . En algunos aspectos, la secuencia de nucleótidos de región variable de cadena ligera es una secuencia génica de región variable de cadena ligera  $\lambda$ . En algunos aspectos, la secuencia de nucleótidos de región variable de cadena ligera es una secuencia génica de región variable de cadena ligera de rata o ratón.

Aspectos adicionales incluyen proteínas de unión a antígeno (por ejemplo, anticuerpos) creadas por los ratones y ratas modificados genéticamente descritos en el presente documento. Asimismo, también se describen proteínas de unión a antígeno (por ejemplo anticuerpos recombinantes) con secuencias de región variable de cadena ligera ( $V_L$ ) procedentes o producidas por (es decir expresadas a partir de los segmentos génicos de región variable de cadena ligera de inmunoglobulina humana no reordenada) por los ratones y ratas modificados genéticamente descritos en el presente documento. En algunos aspectos, las proteínas de unión a antígeno producidas por los métodos descritos en el presente documento comprenden una cadena pesada y una cadena ligera, en la que la cadena pesada no interfiere con la unión de la cadena ligera con el antígeno, y/o la cadena pesada no se une al antígeno en ausencia de la cadena ligera. En algunos aspectos, el dominio variable de cadena ligera se une a un antígeno de interés con una  $K_D$  que no es más de un orden de magnitud mayor en ausencia de cadena pesada que en presencia de la cadena pesada (por ejemplo,  $K_D \sim 10^{-10}$  en presencia de cadena pesada o  $K_D \sim 10^{-9}$  en ausencia de cadena pesada). En algunos aspectos, las proteínas de unión a antígeno como se describen en el presente documento, incluyen una cadena ligera de inmunoglobulina que puede unirse específicamente a un antígeno de interés con una afinidad ( $K_D$ ) menor de  $10^{-6}$ ,  $10^{-7}$ ,  $10^{-8}$ ,  $10^{-9}$  o  $10^{-10}$ . En algunos aspectos, la cadena ligera de inmunoglobulina producida por los métodos puede unirse específicamente a un antígeno de interés en ausencia de una región variable de cadena pesada con una afinidad ( $K_D$ ) menor de  $10^{-6}$ ,  $10^{-7}$ ,  $10^{-8}$ ,  $10^{-9}$  o  $10^{-10}$ .

En diversos aspectos, los dominios variables de cadena ligera generados como se describe en el presente documento, se unen específicamente a una molécula diana ("D"). Una molécula diana es cualquier proteína, polipéptido u otra macromolécula cuya actividad o concentración extracelular se desee atenuar, reducir o eliminar. En muchos casos, la molécula diana a la cual se une una región variable de cadena ligera es una proteína o un polipéptido (es decir, una "proteína diana"); sin embargo, también se describen aspectos en los que la molécula diana ("D") es un hidrato de carbono, una glucoproteína, un lípido, una lipoproteína, un lipopolisacárido u otro polímero no proteico o molécula a la cual se une la región variable de cadena ligera. En diversos aspectos, D puede ser una proteína diana expresada en la superficie celular o una proteína diana soluble. La unión a la diana por parte

- de la molécula de unión a antígeno puede tener lugar en un contexto extracelular o en la superficie celular. Sin embargo, en determinados aspectos, la molécula de unión a antígeno se une a una molécula diana dentro de la célula, por ejemplo, dentro de un componente intracelular tal como el retículo endoplasmático, el aparato de Golgi, un endosoma, lisosoma, etc. Como ejemplos de moléculas diana expresadas en la superficie celular se incluyen
- 5 receptores expresados en la superficie celular, ligandos unidos a membrana, canales iónicos y cualquier otro componente polipeptídico monomérico o multimérico con una parte extracelular que esté conectada o asociada a una membrana celular. Como ejemplos no limitantes de moléculas diana expresadas en la superficie celular que pueden dirigirse mediante las moléculas de unión a antígenos multiespecíficos descritas en el presente documento se incluyen, por ejemplo, receptores de citocina (por ejemplo, receptores de IL-1, IL-4, IL-6, IL-13, IL-22, IL-25, IL-33,
- 10 etc.) así como dianas de superficie celular incluyendo otros receptores transmembrana de tipo 1 tales como PRLR, receptores acoplados a proteína G tales como GCGR, canales iónicos tales como Nav1.7, ASIC1 o ASIC2, proteínas de superficie no receptoras tales como MHC-I (por ejemplo, HLA-B\*27), etc. En aspectos en los que D es una proteína diana expresada en la superficie celular, el componente D1 de la molécula de unión a antígeno multiespecífico puede ser, por ejemplo, un anticuerpo o fragmento de unión a antígeno de un anticuerpo que se une especialmente a D, o un ligando o parte de un ligando que interaccione específicamente con la proteína diana expresada en la superficie celular. Por ejemplo, si D es IL-4R, el componente D1 puede comprender o consistir en IL-4 o una parte de unión al receptor del mismo. Como ejemplos de moléculas diana solubles se incluyen citocinas, factores de crecimiento y otros ligandos y proteínas de señalización. Una proteína diana soluble ejemplar no limitante que puede dirigirse por las moléculas de unión a antígeno multiespecífico descritas en el presente documento incluyen, por ejemplo, IL-1, IL-4, IL-6, IL-13, IL-22, IL-25, IL-33, SOST, DKK1, etc. Las moléculas diana solubles también incluyen, por ejemplo, moléculas diana no humanas tales como alérgenos (por ejemplo, Fel D1, Betv1, CryJ1), patógenos (por ejemplo, *Candida albicans*, *S. aureus*, etc.), y moléculas patógenas (por ejemplo lipopolisacárido (LPS), ácido lipoteicoico (LTA), Proteína A, toxinas, etc.). En aspectos en los que D es una molécula diana soluble, el componente D1 de la molécula de unión a antígeno multiespecífico puede ser, por ejemplo, un anticuerpo o fragmento de unión a antígeno de un anticuerpo que se une específicamente a D, o un receptor o parte de un receptor que interacciona específicamente con la molécula diana soluble. Por ejemplo, si D es IL-4, el componente D1 puede comprender o consistir en IL-4R o una parte de unión al ligando del mismo. Las moléculas diana también incluyen antígenos asociados a tumor.
- 30 En otro aspecto, las proteínas de unión a antígeno (por ejemplo, anticuerpos biespecíficos o trispecíficos) pueden prepararse utilizando dominios variables de cadena ligera específicos de antígeno procedentes de (es decir, con secuencias de región variable de cadena ligera humana ( $V_L$ ) generadas por) un ratón o una rata que comprenden un locus de inmunoglobulina con una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada (es decir, un ratón o una rata que comprende una secuencia VDJ de cadena pesada reordenada, prediseñada). Dichas cadenas ligeras específicas de antígeno, quiméricas inversas (por ejemplo, variable humana/constante de ratón) pueden utilizarse para obtener secuencias de región variable de cadena ligera específicas de antígeno que pueden clonarse en fase en un vector de expresión con una secuencia de región constante de cadena ligera humana adecuada. Una región (o regiones) variable de cadena pesada humana específica de antígeno (específica para un epítipo diferente en el mismo o diferente antígeno que la cadena ligera específica de antígeno) de un animal que comprende un locus de inmunoglobulina con una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada (es decir, un ratón que comprende una secuencia VDJ de cadena pesada reordenada, prediseñada), puede clonarse en fase en un vector de expresión que comprende una secuencia de región constante de cadena pesada humana, y las cadenas ligeras y pesadas humanas específicas de antígeno pueden coexpresarse en una célula adecuada para obtener una proteína de unión a antígeno (por ejemplo, un anticuerpo humano biespecífico o trispecífico). Como alternativa, una cadena pesada específica de antígeno previamente seleccionada, por ejemplo, una cadena pesada de un anticuerpo que comprende una cadena ligera procedente del mismo segmento génico de región variable que el utilizado en la secuencia de ácido nucleico de región variable de cadena pesada humana reordenada, puede clonarse en fase en un vector de expresión que comprende una secuencia de región constante de cadena pesada humana, y las cadenas ligera y pesada humanas específicas de antígeno pueden coexpresarse en una célula adecuada para obtener una proteína de unión a antígeno (por ejemplo, un anticuerpo humano biespecífico o trispecífico). En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena pesada no humana (por ejemplo, una secuencia kappa o lambda de rata o ratón). En algunos aspectos, la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada está unida operativamente a una secuencia génica de región constante de cadena ligera no humana (por ejemplo, una secuencia kappa o lambda de ratón o rata,). En algunos aspectos, las secuencias de región variable de cadena ligera ( $V_L$ ) humana son secuencias génicas kappa.
- 60 En otro aspecto, se describe un método para crear una proteína de unión a antígeno multiespecífico que comprende:
- (a) inmunizar a un primer ratón o rata, que contenga un primer locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el primer ratón o rata comprende en su genoma (i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada;
- 65

(b) permitir que el primer ratón o rata genere una respuesta inmunitaria;

(c) recoger un primer linfocito (por ejemplo, un linfocito B) del primer ratón o rata inmunizado, en el que el primer linfocito expresa anticuerpos madurados por afinidad, en el que los anticuerpos madurados por afinidad comprenden un dominio variable humano fusionado a un dominio constante no humano;

(d) identificar una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera humana de los anticuerpos madurados por afinidad;

(e) clonar la secuencia de ácido nucleico de (d) en una primera construcción de expresión en fase con una secuencia de ácido nucleico de región constante humana adecuada (por ejemplo, una secuencia kappa o lambda humana) para formar un primer gen polipeptídico;

(f) inmunizar a un segundo ratón o rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un segundo antígeno de interés, en el que el segundo ratón o rata comprende en su genoma (i) segmentos génicos V, D y J humanos no reordenados unidos a una secuencia de ácido nucleico de región constante de cadena pesada no humana; y (ii) una sola secuencia de región variable de cadena ligera humana reordenada;

(g) permitir que el segundo ratón o rata genere una respuesta inmunitaria;

(h) recoger un segundo linfocito del segundo ratón o rata inmunizado, en el que el segundo linfocito expresa anticuerpos madurados por afinidad, en el que los anticuerpos madurados por afinidad comprenden un dominio variable de cadena pesada humana fusionado a un dominio constante no humano;

(i) identificar una secuencia de ácido nucleico que codifique un dominio variable de cadena pesada humana de los anticuerpos madurados por afinidad que se unen específicamente al segundo antígeno;

(j) clonar la secuencia de ácido nucleico de (i) en una segunda construcción de expresión en fase con una secuencia de ácido nucleico de región constante humana adecuada (por ejemplo, una secuencia constante de IgG1 humana) para formar un segundo gen polipeptídico; y

(k) introducir la primera construcción de expresión y la segunda construcción de expresión en una célula adecuada para expresar el primer gen polipeptídico y el segundo gen polipeptídico para formar una proteína de unión a antígeno que comprenda un dímero del segundo polipéptido, en el que cada monómero del segundo polipéptido está asociado a un monómero del primer polipéptido.

En diversos aspectos, la primera construcción de expresión y la segunda construcción de expresión están en vectores distintos. En diversos aspectos, la primera construcción de expresión y la segunda construcción de expresión están en el mismo vector. En diversos aspectos, el primer antígeno y el segundo antígeno son diferentes. En un aspecto, el primer antígeno y el segundo antígeno son iguales. En diversos aspectos, el primer antígeno es un receptor de superficie celular y el segundo antígeno se selecciona de un antígeno soluble y de un antígeno unido a una superficie celular. En aspectos específicos, el primer antígeno es un receptor Fc (por ejemplo, un FcRn), el segundo antígeno es una proteína soluble y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico V<sub>L</sub> en el genoma del ratón o de la rata.

En otro aspecto, se describe un método para crear una proteína de unión a antígeno multiespecífico que comprende:

(a) inmunizar a un primer ratón o rata, que contenga un primer locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el primer ratón o rata comprende en su genoma (i) una primera secuencia de nucleótidos que codifique un dominio variable de cadena pesada reordenada (es decir, en la que la primera secuencia de nucleótidos es una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada), en la que la primera secuencia de nucleótidos está unida operativamente a una secuencia génica de región constante de cadena ligera; y (ii) una segunda secuencia de nucleótidos que codifique un dominio variable de cadena ligera humana o no humana (es decir, en el que la segunda secuencia de nucleótidos es una secuencia de nucleótidos variable de cadena ligera de inmunoglobulina humana no reordenada), en la que la segunda secuencia de nucleótidos está opcionalmente unida a una secuencia génica de región constante de cadena pesada;

(b) permitir que el primer ratón o rata genere una respuesta inmunitaria;

(c) recoger un primer linfocito (por ejemplo, un linfocito B) del primer ratón o rata inmunizado, en el que el primer linfocito expresa anticuerpos madurados por afinidad, en el que los anticuerpos madurados por afinidad comprenden un dominio variable humano fusionado a un dominio constante no humano;

(d) identificar una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera humana de los anticuerpos madurados por afinidad;

5 (e) clonar la secuencia de ácido nucleico de (d) en una primera construcción de expresión en fase con una secuencia de ácido nucleico de región constante humana adecuada (por ejemplo, una secuencia lambda o kappa humana) para formar un primer gen polipeptídico;

10 (f) inmunizar a un segundo ratón o rata que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un segundo antígeno de interés, en el que el segundo ratón o rata comprende en su genoma (i) segmentos génicos V, D y J humanos no reordenados unidos a una secuencia de ácido nucleico de región constante de cadena pesada no humana y (ii) una sola secuencia de región variable de cadena ligera humana reordenada;

15 (g) permitir que el segundo ratón o rata genere una respuesta inmunitaria;

(h) recoger un segundo linfocito del segundo ratón o rata inmunizado, en el que el segundo linfocito expresa anticuerpos madurados por afinidad, en el que los anticuerpos madurados por afinidad comprenden un dominio variable de cadena pesada humana fusionado a un dominio constante no humano;

20 (i) identificar una secuencia de ácido nucleico que codifique un dominio variable de cadena pesada humana de los anticuerpos madurados por afinidad que se unen específicamente al segundo antígeno;

25 (j) clonar la secuencia de ácido nucleico de (i) en una segunda construcción de expresión en fase con una secuencia de ácido nucleico de región constante humana adecuada (por ejemplo, una región constante IgG1 humana) para formar un segundo gen polipeptídico; y

30 (k) introducir la primera construcción de expresión y la segunda construcción de expresión en una célula adecuada para expresar el primer gen polipeptídico y el segundo gen polipeptídico para formar una proteína de unión a antígeno que comprenda un dímero del segundo polipéptido, en el que cada monómero del segundo polipéptido está asociado a un monómero del primer polipéptido.

35 En diversos aspectos, la primera construcción de expresión y la segunda construcción de expresión están en vectores distintos. En diversos aspectos, la primera construcción de expresión y la segunda construcción de expresión están en el mismo vector. En algunos aspectos, el primer antígeno y el segundo antígeno son diferentes. En algunos aspectos, el primer antígeno y el segundo antígeno son iguales. En diversos aspectos, el primer antígeno es un receptor de superficie celular, y el segundo antígeno se selecciona de un antígeno soluble y un antígeno unido a una superficie celular. En aspectos específicos, el primer antígeno es un receptor Fc (por ejemplo, un FcRn), el segundo antígeno es una proteína soluble y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico V<sub>L</sub> en el genoma del ratón o de la rata.

40 En otro aspecto, se describe un método para crear una proteína de unión a antígeno multiespecífico que comprende:

45 (a) inmunizar a un primer ratón o rata, que contenga un primer locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un antígeno de interés, en el que el primer ratón o rata comprende en su genoma (i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada y (ii) dos o más segmentos génicos de región variable de cadena ligera (V<sub>L</sub> y J<sub>L</sub>) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre;

50 (b) permitir que el primer ratón o rata genere una respuesta inmunitaria;

(c) recoger un primer linfocito (por ejemplo, un linfocito B) del primer ratón o rata inmunizado, en el que el primer linfocito expresa anticuerpos madurados por afinidad, en el que los anticuerpos madurados por afinidad comprenden un dominio variable humano fusionado a un dominio constante no humano;

55 (d) identificar una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera humana de los anticuerpos madurados por afinidad;

60 (e) clonar la secuencia de ácido nucleico de (d) en una primera construcción de expresión en fase con una secuencia de ácido nucleico de región constante humana adecuada (por ejemplo, una secuencia lambda o kappa humana) para formar un primer gen polipeptídico;

65 (f) inmunizar a un segundo ratón o rata, que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un segundo antígeno de interés, en el que el segundo ratón o rata comprende en su genoma (i) segmentos génicos V, D y J humanos no reordenados unidos a una secuencia de ácido nucleico de región constante de cadena pesada no humana; y (ii) una sola secuencia de región variable

de cadena ligera humana reordenada, en la que la sola secuencia de región variable de cadena ligera humana reordenada procede del mismo segmento génico  $V_L$  que el segmento génico  $V_L$  que codifica el dominio variable de cadena ligera de la etapa (c);

5 (g) permitir que el segundo ratón o rata genere una respuesta inmunitaria;

(h) recoger un segundo linfocito del segundo ratón o rata inmunizado, en el que el segundo linfocito expresa anticuerpos madurados por afinidad, en el que los anticuerpos madurados por afinidad comprenden un dominio variable de cadena pesada humana fusionado a un dominio constante no humano;

10 (i) identificar una secuencia de ácido nucleico que codifique un dominio variable de cadena pesada humana de los anticuerpos madurados por afinidad que se una específicamente al segundo antígeno;

15 (j) clonar la secuencia de ácido nucleico de (i) en una segunda construcción de expresión en fase con una secuencia de ácido nucleico de región constante humana adecuada (por ejemplo, una secuencia constante de IgG1 humana) para formar un segundo gen polipeptídico; y

20 (k) introducir la primera construcción de expresión y la segunda construcción de expresión en una célula adecuada para expresar el primer gen polipeptídico y el segundo gen polipeptídico para formar una proteína de unión a antígeno que comprenda un dímero del segundo polipéptido, en el que cada monómero del segundo polipéptido está asociado a un monómero del primer polipéptido.

25 En diversos aspectos, tanto la primera como la segunda construcción de expresión están en vectores distintos. En diversos aspectos, tanto la primera como la segunda construcción de expresión están en el mismo vector. En diversos aspectos, el primer y segundo antígeno son diferentes. En diversos aspectos, el primer y segundo antígeno son iguales. En diversos aspectos, el primer antígeno es un receptor de superficie celular, y el segundo antígeno se selecciona de un antígeno soluble y un antígeno unido a una superficie celular. En diversos aspectos, el primer antígeno es un receptor Fc (por ejemplo, un FcRn), el segundo antígeno es una proteína soluble, y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma del ratón o la rata.

30 En otro aspecto, se describe un método para la creación de una proteína de unión a antígeno multiespecífico que comprende:

35 (a) inmunizar a un ratón o a una rata que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un primer antígeno, en el que el ratón o la rata comprende en su genoma: (i) una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y (ii) uno o más segmentos génicos  $V_L$  y  $J_L$  de inmunoglobulina humana;

40 (b) permitir que el ratón o la rata genere una respuesta inmunitaria;

45 (c) recoger un linfocito (por ejemplo, un linfocito B) del ratón o la rata inmunizado, en el que el linfocito expresa anticuerpos madurados por afinidad que comprenden un dominio variable de cadena ligera de inmunoglobulina humana fusionado a un dominio constante de inmunoglobulina de ratón;

(d) identificar una primera secuencia de ácido nucleico que codifique el dominio variable de cadena ligera humana de los anticuerpos madurados por afinidad;

50 (e) clonar la primera secuencia de ácido nucleico de (d) en un primer vector de expresión en fase con una secuencia de ácido nucleico de región constante de cadena ligera humana;

55 (f) introducir en una célula hospedadora: (i) el primer vector de expresión que comprende la primera secuencia de ácido nucleico en fase con la secuencia de ácido nucleico de región constante de cadena ligera humana; y (ii) un segundo vector de expresión que comprende una segunda secuencia de ácido nucleico que codifica un primer dominio variable de cadena pesada específico de antígeno, fusionado a una región constante de cadena pesada humana;

60 (g) cultivar la célula hospedadora para permitir la expresión de anticuerpos multiespecíficos; y

(h) aislar los anticuerpos multiespecíficos, en el que los anticuerpos multiespecíficos comprenden el primer dominio variable de cadena pesada y de cadena ligera específico de antígeno, en el que el dominio variable de cadena pesada de los anticuerpos multiespecíficos presenta una especificidad de unión a antígeno distinta de la del dominio variable de cadena ligera.

65

En diversos aspectos, los anticuerpos multiespecíficos son anticuerpos biespecíficos. En algunos aspectos, los anticuerpos biespecíficos son anticuerpos trispecíficos, y la etapa (f) comprende adicionalmente la introducción de un tercer vector de expresión que comprende una tercera secuencia de ácido nucleico que codifica un segundo dominio variable de cadena pesada específico de antígeno, fusionado a la secuencia de región constante de cadena pesada humana.

En otro aspecto, se describe un método para crear una proteína de unión a antígeno multiespecífico que comprende:

(a) inmunizar a un ratón o una rata que contenga un locus de inmunoglobulina modificado genéticamente como se describe en el presente documento, con un primer antígeno, en el que el ratón o la rata comprende en su genoma: (i) un dominio variable de cadena pesada reordenada unida operativamente a un secuencia de ácido nucleico de región constante de cadena pesada; y (ii) uno o más segmentos génicos  $V_L$  y  $J_L$  de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre;

(b) permitir que el ratón o la rata genere una respuesta inmunitaria;

(c) recoger un linfocito (por ejemplo, un linfocito B) del ratón o rata inmunizado, en el que el linfocito expresa anticuerpos madurados por afinidad que comprenden un dominio variable de cadena ligera de inmunoglobulina humana fusionado a un dominio constante de inmunoglobulina de ratón;

(d) identificar una primera secuencia de ácido nucleico que codifique el dominio variable de cadena ligera humana de los anticuerpos madurados por afinidad;

(e) clonar la primera secuencia de ácido nucleico de (d) en un primer vector de expresión en fase con una secuencia de ácido nucleico de región constante de cadena ligera humana;

(f) introducir en una célula hospedadora: (i) el primer vector de expresión que comprende la primera secuencia de ácido nucleico en fase con la secuencia de ácido nucleico de región constante de cadena ligera humana; y (ii) un segundo vector de expresión que comprende una segunda secuencia de ácido nucleico que codifica un primer dominio variable de cadena pesada específico de antígeno fusionado a una región constante de cadena pesada humana;

(g) cultivar la célula hospedadora para permitir la expresión de anticuerpos multiespecíficos; y

(h) aislar los anticuerpos multiespecíficos, en el que los anticuerpos multiespecíficos comprenden el primer dominio variable de cadena ligera y de cadena pesada específico de antígeno, en el que el dominio variable de cadena pesada de los anticuerpos multiespecíficos presenta una especificidad de unión a antígeno distinta de la del dominio variable de cadena ligera.

En diversos aspectos, los anticuerpos multiespecíficos son anticuerpos biespecíficos. En algunos aspectos, los anticuerpos multiespecíficos son anticuerpos trispecíficos y la etapa (f) comprende adicionalmente la introducción de un tercer vector de expresión que comprende una tercera secuencia de ácido nucleico que codifica un segundo dominio variable de cadena pesada específico de antígeno fusionado a la secuencia de región constante de cadena pesada humana.

En otro aspecto, se describen métodos para crear una proteína de unión a antígeno que comprende un dominio variable de cadena ligera de inmunoglobulina que puede unirse a un antígeno independientemente de un dominio variable de cadena pesada. Dichos métodos comprenden

(a) inmunizar a un ratón o a una rata modificado genéticamente, con un primer antígeno que comprenda un primer epítipo o una parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma:

(i) una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada, unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y

(ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana no reordenada unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera de inmunoglobulina;

(b) permitir que el ratón o la rata genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

(c) aislar del ratón o de la rata una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique el dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

5 (e) emplear la secuencia de ácido nucleico de (d) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

10 (f) expresar la secuencia de ácido nucleico de (d) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo para formar una proteína de unión a antígeno cuya cadena ligera esté codificada por el ácido nucleico de (d) y que se una al primer epítipo o parte inmunógena del mismo, independientemente de la cadena pesada y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

15 En algunos aspectos al menos uno de los segmentos génicos de cadena ligera  $V_L$  o  $J_L$  humana no reordenada codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos, el primer epítipo procede de un receptor de superficie celular. En algunos aspectos, el receptor de superficie celular es un receptor Fc. En aspectos particulares, el receptor Fc es FcRn. En algunos aspectos, el segundo antígeno o epítipo procede de un antígeno soluble. En algunos aspectos, el segundo antígeno o epítipo procede de un receptor de superficie celular. En algunos aspectos, el primer antígeno es un receptor Fc, el segundo antígeno es una proteína soluble y la proteína de unión a antígeno comprende una o más sustituciones o inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma del ratón o de la rata.

20 En otro aspecto, se describen métodos para crear una proteína de unión a antígeno que comprenda un dominio variable de cadena ligera de inmunoglobulina que se pueda unir a un antígeno independientemente de un dominio variable de cadena pesada. Dicho método comprende

(a) inmunizar a un ratón o a una rata modificado genéticamente, con un primer antígeno que comprenda un primer epítipo o parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma:

30 (i) una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera; y

35 (ii) segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana no reordenada unidos operativamente a una secuencia de ácido nucleico de región constante de cadena pesada de inmunoglobulina;

(b) permitir que el ratón o la rata genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

40 (c) aislar del ratón o de la rata una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

45 (d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique el dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;

(e) emplear la secuencia de ácido nucleico de (d) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

50 (f) expresar la secuencia de ácido nucleico de (d) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente al segundo antígeno o epítipo para formar una proteína de unión a antígeno cuya cadena ligera esté codificada por el ácido nucleico de (d) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

55 En algunos aspectos al menos uno de los segmentos génicos de cadena ligera  $V_L$  o  $J_L$  humana no reordenada codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos, el primer epítipo procede de un receptor de superficie celular. En algunos aspectos, el receptor de superficie celular es un receptor Fc. En aspectos particulares, el receptor Fc es FcRn. En algunos aspectos, el segundo antígeno o epítipo procede de un antígeno soluble. En algunos aspectos, el segundo antígeno o epítipo procede de un receptor de superficie celular. En algunos aspectos, el primer antígeno es un receptor Fc, el segundo antígeno es una proteína soluble, y la proteína de unión a antígeno comprende una o más sustituciones o inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma del ratón o de la rata.

En otro aspecto, se describen métodos para crear una proteína de unión a antígeno que comprende un dominio variable de cadena ligera de inmunoglobulina que pueda unirse a un antígeno independientemente de un dominio variable de cadena pesada. Dichos métodos comprenden

- 5 (a) inmunizar a una rata o a un ratón modificado genéticamente, con un primer antígeno que comprende un primer epítipo o parte inmunógena del mismo, en el que el ratón o la rata comprende en su genoma:
- 10 (i) una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y
- 15 (ii) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera de inmunoglobulina;
- 20 (b) permitir que el ratón o la rata genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;
- (c) aislar del ratón o de la rata una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;
- 25 (d) obtener de la célula de (c) la secuencia de ácido nucleico que codifique el dominio variable de cadena ligera que se una específicamente al primer epítipo o parte inmunógena del mismo;
- (e) emplear secuencia de ácido nucleico de (d) en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y
- 30 (f) expresar la secuencia de ácido nucleico de (d) en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo para formar una proteína de unión a antígeno cuya cadena ligera esté codificada por el ácido nucleico de (d) y que se una al primer epítipo o parte inmunógena del mismo independientemente de la cadena pesada y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

35 En algunos aspectos, al menos uno de los segmentos génicos de cadena ligera  $V_L$  o  $J_L$  humana, codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente. En algunos aspectos el primer epítipo procede de un receptor de superficie celular. En algunos aspectos, el receptor de superficie celular es un receptor Fc. En aspectos particulares, el receptor Fc es FcRn. En algunos aspectos, el segundo antígeno o epítipo procede de un antígeno soluble. En algunos aspectos, el antígeno soluble o epítipo procede de un receptor de superficie celular. En algunos aspectos, el primer antígeno es un receptor Fc, el segundo antígeno es una proteína soluble y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma del ratón o de la rata.

45 En otro aspecto, para facilitar la separación de las proteínas de unión a antígeno descritas en el presente documento, se modifica una de las cadenas pesadas para excluir un determinante de unión a la Proteína A, dando como resultado una afinidad de unión a la Proteína A diferencial de una proteína de unión homodimérica a partir de una proteína de unión heterodimérica. Las composiciones y métodos que abordan este tema se describen en la Patente de Estados Unidos n.º 8.586.713, otorgada el 19 de noviembre de 2013, titulada "Readily Isolated Bispecific Antibodies with Native Immunoglobulin Format". Una vez seleccionada la especie que comprende la cadena pesada heterodimérica con una cadena ligera idéntica, esta proteína de unión a antígeno biespecífico puede explorarse para confirmar la conservación de su propiedad de unión a antígeno dependiente de pH.

50 En diversos aspectos, en el presente documento se describe una célula pluripotente, células madre pluripotentes o totipotentes inducidas, procedentes de un ratón o una rata, que comprenden las diversas modificaciones genómicas. En algunos aspectos, la célula pluripotente o totipotente procede de un animal no humano. En algunos aspectos, el ratón o la rata es un ratón. En aspectos específicos, la célula pluripotente es una célula madre embrionaria (ES, *embryonic stem*). En algunos aspectos, la célula pluripotente comprende en su genoma: (i) un locus de cadena pesada de inmunoglobulina que comprende una secuencia de ácido nucleico de región variable de cadena pesada humana reordenada unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada; y (ii) un locus de cadena ligera de inmunoglobulina que comprende uno o más segmentos génicos variables de cadena ligera  $V_L$  y  $J_L$  de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera. En aspectos específicos, la célula pluripotente, células madre pluripotentes o totipotentes inducidas, son células madre embrionarias (ES) de ratón o de rata. En algunos aspectos, la célula pluripotente, células madre pluripotentes o totipotentes inducidas, tienen un cariotipo XX o un cariotipo XY.

65 También se describen células que comprenden un núcleo que contiene una modificación genética como se describe en el presente documento, por ejemplo, una modificación introducida en una célula mediante inyección pronuclear.

En otro aspecto, se describe un hibridoma o cuadro, procedente de una célula del ratón o de la rata como se describe en el presente documento. En algunos aspectos, el ratón o la rata es un ratón.

5 En otro aspecto, se describe un linfocito aislado de una rata o un ratón modificado genéticamente como se describe en el presente documento. En algunos aspectos, el linfocito es un linfocito B, en el que el linfocito B comprende un locus genómico de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, unida operativamente a una secuencia génica de región constante de cadena ligera o de cadena pesada de rata o de ratón. En algunos aspectos, el linfocito B es capaz de producir anticuerpos en los que el dominio variable de cadena pesada reordenada, como se describe en el presente documento, está unido operativamente a un dominio constante de cadena ligera o de cadena pesada.

15 En otro aspecto, se describe un embrión de una rata o ratón que comprende una célula cuyo genoma comprende: (i) un locus de cadena pesada de inmunoglobulina que comprende una secuencia de ácido nucleico de región variable de cadena pesada de inmunoglobulina reordenada, unida operativamente a una secuencia de ácido nucleico de región constante; y (ii) un locus de cadena ligera de inmunoglobulina que comprende dos o más segmentos génicos de región variable de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera.

20 En diversos aspectos, la rata o el ratón modificado genéticamente, expresa un repertorio de anticuerpos (por ejemplo, un repertorio de IgG) que procede de la secuencia de nucleótidos que codifica el dominio variable de cadena pesada reordenada, y una pluralidad de segmentos V de cadena ligera (y una pluralidad de segmentos J de cadena ligera). En algunos aspectos, el locus modificado genéticamente produce una población de anticuerpos que comprende una cadena ligera de inmunoglobulina que es capaz de unirse específicamente a un antígeno de interés con una afinidad ( $K_D$ ) menor que  $10^{-6}$ ,  $10^{-7}$ ,  $10^{-8}$ ,  $10^{-9}$  o  $10^{-10}$ . En algunos aspectos, la cadena ligera de inmunoglobulina expresada por el locus modificado genéticamente es capaz de unirse específicamente a un antígeno de interés en ausencia de una región variable de cadena pesada con una afinidad ( $K_D$ ) menor que  $10^{-6}$ ,  $10^{-7}$ ,  $10^{-8}$ ,  $10^{-9}$  o  $10^{-10}$ .

30 En diversos aspectos, las modificaciones genéticas descritas en el presente documento no afectan a la fertilidad de los ratones o de las ratas (véase, por ejemplo, el documento US 2012-0322108A1). En algunos aspectos, el locus de cadena pesada comprende un gen Adam6a endógeno, un gen Adam6b, o ambos, y la modificación genética no afecta a la expresión y/o función del gen Adam6a endógeno, gen Adam6b o ambos. En algunos aspectos, el genoma de la rata o ratón modificado genéticamente comprende un gen Adam6a, un gen Adam6b, o ambos, localizado ectópicamente. En algunos aspectos, un gen Adam6a y/o Adam6b se coloca en el extremo 5' cadena arriba de la unidad transcripcional de la secuencia de ácido nucleico de región variable de cadena pesada reordenada. En algunos aspectos, el gen Adam6a y/o Adam6b se coloca en el extremo 3' cadena abajo de la unidad transcripcional de la secuencia de ácido nucleico de región variable de cadena pesada reordenada.

40 En algunos aspectos, el locus de cadena pesada modificado genéticamente no comprende una secuencia de ácido nucleico de la Región de Control Intergénica 1 (IGCR1). En algunos aspectos, el locus de cadena pesada modificado genéticamente comprende una secuencia de IGCR1 cadena abajo de la secuencia de ácido nucleico de región variable de cadena pesada reordenada. En algunos aspectos, la secuencia de ácido nucleico de IGCR1 está presente entre la secuencia de ácido nucleico de región variable de cadena pesada reordenada y el segmento génico  $D_H$  más próximo a V.

45 En algunos aspectos, como se ha indicado anteriormente, el locus de cadena ligera de inmunoglobulina de los ratones o ratas descritos en el presente documento, comprende un repertorio limitado de segmentos génicos variables de cadena ligera, por ejemplo, (i) uno, dos o más segmentos génicos  $V_L$  humanos, pero menos que el número de segmentos de tipo silvestre. En algunos aspectos, el ratón o la rata es un ratón; y el dominio variable de cadena ligera de inmunoglobulina se genera a partir de una reordenación de uno de dos segmentos génicos Vk humanos y de uno de 1, 2, 3, 4 o 5 segmentos génicos Jk humanos. En algunos aspectos, el ratón presenta una proporción de (a) linfocitos B en la médula ósea que expresan una inmunoglobulina que tiene una cadena ligera  $\lambda$  con respecto a (b) linfocitos B en la médula ósea que expresan una inmunoglobulina que tiene una cadena ligera  $\kappa$ , de aproximadamente 1 a aproximadamente 15. En algunos aspectos, la reordenación incluye un segmento génico Vk1-39 humano. En algunos aspectos, la reordenación incluye un segmento génico Vk3-20 humano. En algunos aspectos, los dos segmentos génicos Vk humanos están en un locus Vk de inmunoglobulina endógena, y, en algunos aspectos los dos segmentos génicos Vk humanos reemplazan a todos o sustancialmente a todos los segmentos génicos Vk de inmunoglobulina de ratón. En algunos aspectos, los dos segmentos génicos Vk humanos están en un locus Vk de inmunoglobulina endógena, y, en algunos aspectos, los dos segmentos génicos Vk humanos reemplazan a todos o sustancialmente a todos los segmentos génicos Vk y Jk de inmunoglobulina de ratón. En algunos aspectos, los dos segmentos génicos Vk humanos están unidos operativamente a dos o más (por ejemplo, 2, 3, 4, 5) segmentos génicos Jk humanos. En algunos otros aspectos, el dominio variable de cadena ligera del ratón se genera a través de una reordenación de un segmento génico Vk1-39 humano o un segmento génico Vk3-20 humano y uno de dos o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos Jk humanos. En algunos de dichos aspectos, la proporción de linfocitos B inmaduros en la médula ósea que expresan una inmunoglobulina que tiene una cadena ligera  $\lambda$  con respecto a los linfocitos B inmaduros que expresan una inmunoglobulina que tiene una cadena

ligera  $\kappa$ , es de aproximadamente 1 a aproximadamente 13. En algunos otros aspectos, el dominio variable de cadena ligera del ratón se genera a través de una reordenación de un segmento génico Vk1-39 humano o un segmento génico Vk3-20 humano y uno de dos o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos Jk humanos y la proporción de linfocitos B maduros en la médula ósea que expresan una inmunoglobulina que tiene una cadena ligera  $\lambda$  con respecto a linfocitos B inmaduros que expresan una inmunoglobulina que tiene una cadena ligera  $\kappa$ , es de aproximadamente 1 a aproximadamente 7.

En aspectos particulares, el dominio variable de cadena ligera de un ratón modificado genéticamente como se describe en el presente documento, se genera a través de una reordenación de un segmento génico Vk1-39 humano o un segmento génico Vk3-20 humano y uno de dos o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos Jk humanos, y tiene una población de linfocitos pro B en la médula ósea en el intervalo de aproximadamente  $2,5 \times 10^4$  a aproximadamente  $1,5 \times 10^5$  linfocitos, inclusive, por ejemplo, de aproximadamente  $2,5 \times 10^4$ ,  $3,0 \times 10^4$ ,  $3,5 \times 10^4$ ,  $4,0 \times 10^4$ ,  $4,5 \times 10^4$ ,  $5,0 \times 10^4$ ,  $5,5 \times 10^4$ ,  $6,0 \times 10^4$ ,  $6,5 \times 10^4$ ,  $7,0 \times 10^4$ ,  $7,5 \times 10^4$ ,  $8,0 \times 10^4$ ,  $8,5 \times 10^4$ ,  $9,0 \times 10^4$ ,  $9,5 \times 10^4$ ,  $1,0 \times 10^5$ , o  $1,5 \times 10^5$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos pro B en la médula ósea de aproximadamente  $2,88 \times 10^4$  células; en algunos aspectos, un ratón o una rata (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos pro B en la médula ósea de aproximadamente  $6,42 \times 10^4$  células; en algunos aspectos, un ratón o una rata (por ejemplo, un ratón) modificado descrito en el presente documento comprende una población de linfocitos pro B en la médula ósea de aproximadamente  $9,16 \times 10^4$  células; en algunos aspectos, un ratón o una rata (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos pro B en la médula ósea de aproximadamente  $1,19 \times 10^5$  células. Los linfocitos pro B ejemplares en la médula ósea de ratas y ratones (por ejemplo, ratones) modificados genéticamente, como se describe en el presente documento, se caracterizan por la expresión de CD19, CD43, c-kit y/o una combinación de los mismos (por ejemplo, CD19<sup>+</sup>, CD43<sup>+</sup>, c-kit<sup>+</sup>). En algunos aspectos, una rata o ratón (por ejemplo, ratón), como se describe en el presente documento, expresa una cadena ligera generada a través de una reordenación de un segmento génico Vk1-39 humano o un segmento génico Vk3-20 humano y uno de dos o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos Jk humanos, y tiene una población de linfocitos pre B en la médula ósea en el intervalo de aproximadamente  $1 \times 10^6$  a aproximadamente  $2 \times 10^6$  linfocitos, inclusive, por ejemplo, de aproximadamente  $1,0 \times 10^6$ ,  $1,1 \times 10^6$ ,  $1,2 \times 10^6$ ,  $1,3 \times 10^6$ ,  $1,4 \times 10^6$ ,  $1,5 \times 10^6$ ,  $1,6 \times 10^6$ ,  $1,7 \times 10^6$ ,  $1,8 \times 10^6$ ,  $1,9 \times 10^6$ , o  $2,0 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos pre B en la médula ósea de aproximadamente  $1,25 \times 10^6$  células; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos pre B en la médula ósea de aproximadamente  $1,46 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos pre B en la médula ósea de aproximadamente  $1,64 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos pre B en la médula ósea de aproximadamente  $2,03 \times 10^6$  linfocitos. Los linfocitos pre B ejemplares en la médula ósea de las ratas y ratones (por ejemplo, ratones) modificados genéticamente, como se describe en el presente documento, se caracterizan por la expresión de CD19, CD43, c-kit y/o una combinación de los mismos (por ejemplo, CD19<sup>+</sup>, CD43<sup>+</sup>, c-kit<sup>+</sup>).

En diversos aspectos, un ratón modificado genéticamente como se describe en el presente documento, expresa una cadena ligera generada a través de una reordenación de un segmento génico Vk1-39 humano o un segmento génico Vk3-20 humano y uno de dos o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos Jk humanos, y tiene una población de linfocitos B inmaduros en la médula ósea en el intervalo de aproximadamente  $5 \times 10^5$  a aproximadamente  $7 \times 10^5$  linfocitos, inclusive, por ejemplo, de aproximadamente  $5,0 \times 10^5$ ,  $5,1 \times 10^5$ ,  $5,2 \times 10^5$ ,  $5,3 \times 10^5$ ,  $5,4 \times 10^5$ ,  $5,5 \times 10^5$ ,  $5,6 \times 10^5$ ,  $5,7 \times 10^5$ ,  $5,8 \times 10^5$ ,  $5,9 \times 10^5$ ,  $6,0 \times 10^5$ ,  $6,1 \times 10^5$ ,  $6,2 \times 10^5$ ,  $6,3 \times 10^5$ ,  $6,4 \times 10^5$ ,  $6,5 \times 10^5$ ,  $6,6 \times 10^5$ ,  $6,7 \times 10^5$ ,  $6,8 \times 10^5$ ,  $6,9 \times 10^5$  o  $7,0 \times 10^5$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B inmaduros en la médula ósea de aproximadamente  $5,33 \times 10^5$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B inmaduros en la médula ósea de aproximadamente  $5,80 \times 10^5$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B inmaduros en la médula ósea de aproximadamente  $5,92 \times 10^5$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B inmaduros en la médula ósea de aproximadamente  $6,67 \times 10^5$  linfocitos. Los linfocitos B inmaduros ejemplares en la médula ósea de una rata o un ratón (por ejemplo, ratones) modificado genéticamente como se describe en el presente documento, se caracterizan por la expresión de IgM, B220 y/o una combinación de los mismos (por ejemplo, IgM<sup>+</sup>, B220<sup>int</sup>).

En diversos aspectos, un ratón modificado genéticamente como se describe en el presente documento, expresa una cadena ligera generada a través de una reordenación de un segmento génico Vk1-39 humano o un segmento génico Vk3-20 humano y uno de dos o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos Jk humanos, y tiene una población de linfocitos B maduros en la médula ósea en el intervalo de aproximadamente  $3 \times 10^4$  a aproximadamente  $1,5 \times 10^5$  linfocitos, inclusive, por ejemplo, de aproximadamente  $3,0 \times 10^4$ ,  $3,5 \times 10^4$ ,  $4,0 \times 10^4$ ,  $4,5 \times 10^4$ ,  $5,0 \times 10^4$ ,  $5,5 \times 10^4$ ,  $6,0 \times 10^4$ ,  $6,5 \times 10^4$ ,  $7,0 \times 10^4$ ,  $7,5 \times 10^4$ ,  $8,0 \times 10^4$ ,  $8,5 \times 10^4$ ,  $9,0 \times 10^4$ ,  $9,5 \times 10^4$ ,  $1,0 \times 10^5$  o  $1,5 \times 10^5$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B maduros en la médula ósea de aproximadamente  $3,11 \times 10^4$  linfocitos; en algunos aspectos,

una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B maduros en la médula ósea de aproximadamente  $1,09 \times 10^5$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B maduros en la médula ósea de aproximadamente  $1,16 \times 10^5$  células; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B maduros en la médula ósea de aproximadamente  $1,44 \times 10^5$  linfocitos. Los linfocitos B maduros ejemplares en la médula ósea de ratas o ratones (por ejemplo, ratones) modificados genéticamente, como se describe en el presente documento, se caracterizan por la expresión de IgM, B220 y/o una combinación de los mismos (por ejemplo, IgM<sup>+</sup>, B220<sup>hi</sup>).

En diversos aspectos, una rata o un ratón (por ejemplo, ratón) modificado genéticamente como se describe en el presente documento, expresa una cadena ligera generada a través de una reordenación de un segmento génico Vκ1-39 humano o un segmento génico Vκ3-20 humano y uno de dos o más (por ejemplo, 2, 3, 4 o 5) segmentos génicos Jκ humanos y tiene una población total de linfocitos B en la médula ósea en el intervalo de aproximadamente  $1 \times 10^6$  a aproximadamente  $3 \times 10^6$  linfocitos, inclusive, por ejemplo, de aproximadamente  $1,0 \times 10^6$ ,  $1,1 \times 10^6$ ,  $1,2 \times 10^6$ ,  $1,3 \times 10^6$ ,  $1,4 \times 10^6$ ,  $1,5 \times 10^6$ ,  $1,6 \times 10^6$ ,  $1,7 \times 10^6$ ,  $1,8 \times 10^6$ ,  $1,9 \times 10^6$ ,  $2,0 \times 10^6$ ,  $2,1 \times 10^6$ ,  $2,2 \times 10^6$ ,  $2,3 \times 10^6$ ,  $2,4 \times 10^6$ ,  $2,5 \times 10^6$ ,  $2,6 \times 10^6$ ,  $2,7 \times 10^6$ ,  $2,8 \times 10^6$ ,  $2,9 \times 10^6$  o  $2,0 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población total de linfocitos B en la médula ósea de aproximadamente  $1,59 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, ratón) modificado, descrito en el presente documento, comprende una población total de linfocitos B en la médula ósea de aproximadamente  $1,75 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, ratón) modificado, descrito en el presente documento, comprende una población total de linfocitos B en la médula ósea de aproximadamente  $2,13 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, ratón) modificado, descrito en el presente documento, comprende una población total de linfocitos B en la médula ósea de aproximadamente  $2,55 \times 10^6$  linfocitos. El total de linfocitos B ejemplares en la médula ósea de ratas y ratones (por ejemplo, ratones) modificados genéticamente, como se describe en el presente documento, se caracteriza por la expresión de CD19, CD20 y/o una combinación de los mismos (por ejemplo, CD19<sup>+</sup>).

En diversos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado genéticamente como se describe en el presente documento, comprende un locus de cadena ligera κ de inmunoglobulina que comprende dos segmentos génicos Vκ de inmunoglobulina humana no reordenada y dos o más (por ejemplo, 2, 3, 4, o 5) segmentos génicos Jκ humanos no reordenados, en el que el ratón o la rata (por ejemplo, ratón) comprende una población de linfocitos B esplénicos periféricos, que comprende poblaciones de linfocitos B transicionales (por ejemplo, T1, T2 y T3) que son aproximadamente iguales a las de una rata o un ratón (por ejemplo, un ratón) que comprende un complemento de tipo silvestre de segmentos génicos V y J de cadena ligera κ de inmunoglobulina. Las poblaciones de linfocitos B transicionales (por ejemplo, T1, T2 y T3) ejemplares en el bazo de un ratón o una rata (por ejemplo, un ratón) modificado genéticamente como se describe en el presente documento, se caracterizan por la expresión de IgM, CD23, CD93, B220 y/o una combinación de los mismos.

En diversos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado genéticamente como se describe en el presente documento, comprende una población de linfocitos B T1 en el bazo (por ejemplo, CD93<sup>+</sup>, B220<sup>+</sup>, IgM<sup>hi</sup>, CD23<sup>+</sup>) en el intervalo de aproximadamente  $2 \times 10^6$  a aproximadamente  $7 \times 10^6$  linfocitos, inclusive, por ejemplo de aproximadamente  $2,0 \times 10^6$ ,  $2,5 \times 10^6$ ,  $3,0 \times 10^6$ ,  $3,5 \times 10^6$ ,  $4,0 \times 10^6$ ,  $4,5 \times 10^6$ ,  $5,0 \times 10^6$ ,  $5,5 \times 10^6$ ,  $6,0 \times 10^6$ ,  $6,5 \times 10^6$  o  $7,0 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón), como se describe en el presente documento, comprende una población de linfocitos B T1 en el bazo de aproximadamente  $2,16 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, como se describe en el presente documento, comprende una población de linfocitos B T1 en el bazo de aproximadamente  $3,63 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, como se describe en el presente documento, comprende una población de linfocitos B T1 en el bazo de aproximadamente  $3,91 \times 10^6$ ; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, como se describe en el presente documento, comprende una población de linfocitos B T1 en el bazo de aproximadamente  $6,83 \times 10^6$  linfocitos.

En diversos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado genéticamente como se describe en el presente documento, comprende una población de linfocitos B T2 en el bazo (por ejemplo, CD93<sup>+</sup>, B220<sup>+</sup>, IgM<sup>hi</sup>, CD23<sup>+</sup>) en el intervalo de aproximadamente  $1 \times 10^6$  a aproximadamente  $7 \times 10^6$  linfocitos, inclusive, por ejemplo, de aproximadamente  $1,0 \times 10^6$ ,  $1,5 \times 10^6$ ,  $2,0 \times 10^6$ ,  $2,5 \times 10^6$ ,  $3,0 \times 10^6$ ,  $3,5 \times 10^6$ ,  $4,0 \times 10^6$ ,  $4,5 \times 10^6$ ,  $5,0 \times 10^6$ ,  $5,5 \times 10^6$ ,  $6,0 \times 10^6$ ,  $6,5 \times 10^6$  o  $7,0 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B T2 en el bazo de aproximadamente  $1,30 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B T2 en el bazo de aproximadamente  $2,46 \times 10^6$  linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento comprende una población de linfocitos B T2 en el bazo de aproximadamente  $3,24 \times 10^6$ ; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B T2 en el bazo de aproximadamente  $6,52 \times 10^6$  linfocitos.

En diversos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado genéticamente como se describe en el presente documento, comprende una población de linfocitos B T3 en el bazo (por ejemplo, CD93<sup>+</sup>, B220<sup>+</sup>, IgM<sup>lo</sup>, CD23<sup>+</sup>) en el intervalo de aproximadamente 1x10<sup>6</sup> a aproximadamente 4x10<sup>6</sup> linfocitos, inclusive, por ejemplo, de aproximadamente 1,0x10<sup>6</sup>, 1,5x10<sup>6</sup>, 2,0x10<sup>6</sup>, 2,5x10<sup>6</sup>, 3,0x10<sup>6</sup>, 3,5x10<sup>6</sup> o 4,0x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B T3 en el bazo de aproximadamente 1,08x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento comprende una población de linfocitos B T3 en el bazo de aproximadamente 1,35x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B T3 en el bazo de aproximadamente 3,37x10<sup>6</sup>; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B T1 en el bazo de aproximadamente 3,63x10<sup>6</sup> linfocitos.

Los linfocitos B de la zona marginal, son linfocitos B maduros no circulantes que se segregan anatómicamente en la zona marginal (ZM) del bazo. En los roedores, los linfocitos B de la ZM son sésiles y residen en la pulpa blanca externa del bazo, entre el seno marginal y la pulpa roja. Esta región contiene múltiples subtipos de macrófagos, células dendríticas y linfocitos B de la ZM; no se forma por completo hasta las 2 a 3 semanas después del nacimiento en roedores y de 1 a 2 años en seres humanos. En esta región, los linfocitos B de la ZM expresan típicamente altos niveles de sIgM, CD21, CD1, CD9 con niveles de bajos a insignificantes de sIgD, CD23, CD5 y CD11 b que ayudan a distinguirlos fenotípicamente de los linfocitos B foliculares (FO) y los linfocitos B B1. Similares a los linfocitos B B1, los linfocitos B ZM pueden acumularse rápidamente en las respuestas inmunoadaptativas tempranas de una manera independiente de linfocitos T. Los linfocitos B ZM están específicamente bien posicionados como una primera línea de defensa contra antígenos transmitidos por la enfermedades sistémicas en sangre que entran en la circulación y que quedan atrapados en el bazo. Se piensa que son especialmente reactivos contra componentes de la pared celular bacteriana y que son una fuente importante de anticuerpos específicos de lípidos. Los linfocitos B ZM también presentan un umbral de activación más bajo que sus linfocitos B FO homólogos con una propensidad ponderada para la diferenciación celular plasmática que contribuye adicionalmente a la respuesta de anticuerpos primaria acelerada.

En diversos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado genéticamente como se describe en el presente documento, que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (por ejemplo, V<sub>H</sub>3-23/D/J<sub>H</sub>4), tiene niveles aumentados de linfocitos B de la zona marginal con respecto a los ratones o ratas de tipo silvestre (por ejemplo, ratones de tipo silvestre). En algunos aspectos, los linfocitos B de la zona marginal en una rata o un ratón (por ejemplo, ratón) modificado genéticamente, comprenden una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada que está aumentada en un 10 %, 20 %, 30 %, 40 %, 50 % o más, con respecto a los roedores de tipo silvestre (por ejemplo, ratones de tipo silvestre).

En diversos aspectos, una rata o un ratón (por ejemplo, ratón) modificado genéticamente como se describe en el presente documento, comprende un locus de cadena ligera κ de inmunoglobulina que comprende dos segmentos génicos Vκ de inmunoglobulina humana no reordenada y 1, 2, 3, 4 o 5 segmentos génicos Jκ de inmunoglobulina humana no reordenada, y en el que el ratón o la rata (por ejemplo, ratón) comprende una población de linfocitos B esplénicos periféricos que comprende poblaciones de linfocitos B de la zona marginal y precursores de la zona marginal, que son aproximadamente iguales a las del ratón o la rata (por ejemplo, ratón) que comprende un complemento de tipo silvestre de segmentos génicos Vκ y Jκ de inmunoglobulina. Las poblaciones de linfocitos B de la zona marginal en el bazo de una rata o un ratón (por ejemplo, ratón) modificado genéticamente como se describe en el presente documento, se caracterizan por la expresión de IgM, CD21/35, CD23, CD93, B220 y/o una combinación de los mismos.

En diversos aspectos, una rata o un ratón (por ejemplo, ratón) modificado genéticamente como se describe en el presente documento, comprende una población de linfocitos B de la zona marginal en el bazo (por ejemplo, CD93<sup>+</sup>, B220<sup>+</sup>, IgM<sup>hi</sup>, CD21/35<sup>hi</sup>, CD23<sup>+</sup>) en el intervalo de aproximadamente 1x10<sup>6</sup> a aproximadamente 3x10<sup>6</sup> linfocitos, inclusive, por ejemplo, aproximadamente 1,0x10<sup>6</sup>, 1,5x10<sup>6</sup>, 2,0x10<sup>6</sup>, 2,5x10<sup>6</sup> o 3,0x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B de la zona marginal en el bazo de aproximadamente 1,47x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B de la zona marginal en el bazo de aproximadamente 1,49x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B de la zona marginal en el bazo de aproximadamente 2,26x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado, descrito en el presente documento, comprende una población de linfocitos B de la zona marginal en el bazo de aproximadamente 2,33x10<sup>6</sup> linfocitos.

En diversos aspectos, se describe un ratón o una rata (por ejemplo, ratón) modificado genéticamente, en el que el ratón o la rata (por ejemplo, ratón) comprende un locus de cadena ligera κ de inmunoglobulina que comprende dos segmentos génicos Vκ de inmunoglobulina humana no reordenada y 1, 2, 3, 4 o 5 segmentos génicos Jκ de inmunoglobulina humana no reordenada y en el que el ratón o la rata (por ejemplo, ratón) comprende una población

de linfocitos B esplénicos periféricos que comprende una o más poblaciones de linfocitos B foliculares (por ejemplo, FO-I y FO-II) que son aproximadamente iguales a las de una rata o un ratón (por ejemplo, un ratón) que comprende un complemento de tipo silvestre de segmentos génicos V<sub>k</sub> y J<sub>k</sub> de inmunoglobulina. Las poblaciones de linfocitos B foliculares ejemplares (por ejemplo, FO-I y FO-II) en el bazo de una rata o un ratón (por ejemplo, ratón) modificado genéticamente como se describe en el presente documento, se caracterizan por la expresión de IgM, IgD, CD21/35, CD93, B220 y/o una combinación de los mismos.

En diversos aspectos, una rata o un ratón (por ejemplo, ratón) modificado genéticamente como se describe en el presente documento, comprende una población de linfocitos B folicular de tipo 1 en el bazo (por ejemplo, CD93<sup>-</sup>, B220<sup>+</sup>, Cod21/35<sup>int</sup>, IgM<sup>lo</sup>, IgD<sup>hi</sup>) en el intervalo de aproximadamente 3x10<sup>6</sup> a aproximadamente 1,5x10<sup>7</sup> linfocitos, inclusive, por ejemplo, de aproximadamente 3,0x10<sup>6</sup>, 3,5x10<sup>6</sup>, 4,0x10<sup>6</sup>, 4,5x10<sup>6</sup>, 5,0x10<sup>6</sup>, 5,5x10<sup>6</sup>, 6,0x10<sup>6</sup>, 6,5x10<sup>6</sup>, 7,0x10<sup>6</sup>, 7,5x10<sup>6</sup>, 8,0x10<sup>6</sup>, 8,5x10<sup>6</sup>, 9,0x10<sup>6</sup>, 9,5x10<sup>6</sup>, 1,0x10<sup>7</sup> o 1,5x10<sup>7</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, ratón) modificado, como se describe en el presente documento, comprende una población de linfocitos B folicular de tipo 1 en el bazo de aproximadamente 3,57x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, ratón) modificado, como se describe en el presente documento, comprende una población de linfocitos B folicular de tipo 1 en el bazo de aproximadamente 6,31x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, ratón) modificado como se describe en el presente documento, comprende una población de linfocitos B folicular de tipo 1 en el bazo de aproximadamente 9,42x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, ratón) modificado como se describe en el presente documento, comprende una población de linfocitos B folicular de tipo 1 en el bazo de aproximadamente 1,14x10<sup>7</sup> linfocitos.

En diversos aspectos, una rata o un ratón (por ejemplo, ratón) modificado genéticamente como se describe en el presente documento, comprende una población de linfocitos B folicular de tipo 2 en el bazo (por ejemplo, CD93<sup>-</sup>, B220<sup>+</sup>, CD21/35<sup>int</sup>, IgM<sup>int</sup>, IgD<sup>hi</sup>) en el intervalo de aproximadamente 1x10<sup>6</sup> a aproximadamente 2x10<sup>6</sup> linfocitos, inclusive, por ejemplo, 1,0x10<sup>6</sup>, 1,25x10<sup>6</sup>, 1,5x10<sup>6</sup>, 1,75x10<sup>6</sup> o 2,0x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado descrito en el presente documento comprende una población de linfocitos B folicular de tipo 2 en el bazo de aproximadamente 1,14x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado descrito en el presente documento, comprende una población de linfocitos B folicular de tipo 2 en el bazo de aproximadamente 1,45x10<sup>6</sup> linfocitos; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado descrito en el presente documento, comprende una población de linfocitos B folicular de tipo 2 en el bazo de aproximadamente 1,80x10<sup>6</sup>; en algunos aspectos, una rata o un ratón (por ejemplo, un ratón) modificado descrito en el presente documento, comprende una población de linfocitos B folicular de tipo 2 en el bazo de aproximadamente 2,06x10<sup>6</sup> linfocitos

Las aptitudes de los ratones y ratas modificados genéticamente descritos en el presente documento, para aplicar presión selectiva a los genes o polinucleótidos que codifican regiones o dominios variables de cadena ligera (por ejemplo, las CDR3 de cadena ligera), pueden aplicarse a una variedad de secuencias génicas de cadena ligera variable. En otras palabras, las secuencias génicas de dominio variable de cadena pesada reordenada desveladas en el presente documento pueden emparejarse con una o más modificaciones genéticas de un locus de cadena ligera y/o la inserción de secuencias de nucleótidos que codifican dominios variables de cadena ligera en un locus de cadena pesada. Esto puede realizarse, por ejemplo, apareando (es decir, cruzando o entrecruzando animales con una sola modificación) los ratones y las ratas descritos en el presente documento (restringido a un dominio variable de cadena pesada común o universal) con ratones y ratas que comprenden modificaciones genéticas en uno o más locus codificantes de cadena ligera. Los ratones y ratas modificados genéticamente que comprenden locus de inmunoglobulina con un dominio variable de cadena pesada reordenada y con una o más modificaciones de cadena ligera también pueden generarse por reemplazo génico dirigido de locus múltiples, simultánea o secuencialmente (por ejemplo, mediante recombinación secuencial en células madre embrionarias). Ni el tipo ni el método de modificación en los locus de cadena ligera limitan los aspectos descritos en el presente documento a menos que se indique específicamente. En cambio, la presión selectiva facilitada por los aspectos descritos en el presente documento, puede aplicarse prácticamente a cualquier secuencia polinucleotídica capaz de expresarse y de funcionar como una secuencia de unión a antígeno de cadena ligera, dirigiendo por lo tanto la evolución de regiones variables de anticuerpos adecuadas.

Por ejemplo, como se describe en el presente documento, los ratones y las ratas modificados genéticamente que comprenden un locus de inmunoglobulina con una secuencia génica de dominio variable de cadena pesada reordenada, pueden comprender adicionalmente (por ejemplo, mediante cruzamiento o estrategias múltiples de direccionamiento génico) una o más modificaciones como se describe en los documentos WO 2011/072204, WO 2011/163311, WO 2011/163314, WO 2012/018764, WO 2012/141798, US 2013/0185821, WO 2013/022782, WO 2013/096142, WO 2013/116609. En aspectos particulares, un ratón modificado genéticamente, que comprende una secuencia de ácido nucleico de región variable de cadena pesada reordenada en un locus de cadena ligera (es decir, una secuencia génica de dominio variable de cadena pesada reordenada unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$  humana o no humana), se cruza con un ratón modificado genéticamente que comprende un locus de cadena pesada de inmunoglobulina que comprende segmentos génicos de región variable de cadena ligera humana (por ejemplo, 40 genes V<sub>k</sub> humanos y todos los genes J<sub>k</sub> humanos insertados en un locus de cadena pesada de ratón; véase, por ejemplo, la publicación previa a la concesión de patente de Estados Unidos 2012/0096572). En aspectos específicos, un ratón modificado genéticamente que

comprende una secuencia de ácido nucleico de región variable de cadena pesada reordenada en un locus de cadena ligera (es decir, una secuencia génica de dominio variable de cadena pesada reordenada, unida operativamente a una secuencia génica de región constante de cadena ligera  $\kappa$  humana o no humana) se cruza con un ratón modificado genéticamente que comprende un locus de cadena pesada de inmunoglobulina que comprende uno o más (por ejemplo, dos) segmentos génicos de región variable de cadena ligera humana, pero menos que el número de segmentos de tipo silvestre. Los ratones resultantes pueden producir linfocitos B  $\kappa$ + con cadenas pesadas variables procedentes de una secuencia variable de cadena ligera genómica, facilitando de este modo la identificación de secuencias VJ  $\kappa$  que se unen a dianas específicas, que después pueden reformarse de nuevo a una cadena ligera y emparejarse con una diversidad de cadenas pesadas para producir anticuerpos bi o tri específicos.

### Ejemplos

Los siguientes ejemplos se proponen para proporcionar a los expertos habituales en la técnica una completa divulgación y descripción de cómo crear y utilizar los ratones y las ratas que se describen en el presente documento, y no pretenden limitar el alcance de lo que los inventores consideran como su invención ni tampoco pretenden representar que los siguientes experimentos sean todos o los únicos experimentos realizados. Se han hecho esfuerzos para garantizar la precisión con respecto a los números utilizados (por ejemplo, cantidades, temperatura, etc.), pero puede haber algunos errores y desviaciones experimentales. A menos que se indique de otro modo, las partes son partes en peso, el peso molecular es peso molecular promedio en peso, la temperatura es en grados centígrados y la presión es la atmosférica o casi la atmosférica.

#### Ejemplo 1. Clonación y análisis de expresión de secuencias de cadena pesada universal candidatas

Estudios previos han demostrado que  $hV_{H3-23}$  es un segmento génico de cadena pesada variable humana termoestable y es también uno de los segmentos variables más utilizados habitualmente en el repertorio humano. Por tanto, los segmentos génicos humanos  $V_{H3-23}$ , D4-4 (fase de lectura 2 o 3) y  $J_{H4}$  (o  $J_{H6}$ ) se seleccionaron para diseñar una secuencia variable de cadena pesada reordenada (en lo sucesivo en el presente documento "Cadena Pesada Universal" o "UHC, *Universal Heavy Chain*").

Resumiendo, las siguientes cuatro secuencias de VDJ reordenadas candidatas, se sintetizaron *de novo* (mediante IDT) y se clonaron en vectores de expresión de CMV (por ejemplo, pRG1301 o pRG1368): (1)  $hV_{H3-23}(D_{H4-4\_RF2})J_{H4}$  (SEQ ID NO: 148); (2)  $hV_{H3-23}(D_{H4-4\_RF2})J_{H6}$  (SEQ ID NO: 146); (3)  $hV_{H3-23}(D_{H4-4\_RF3})J_{H4}$  (SEQ ID NO: 147); (4)  $hV_{H3-23}(D_{H4-4\_RF3})J_{H6}$  (SEQ ID NO: 145). Todas estas construcciones se diseñaron de una manera que los genes UHC sintetizados puede ligarse en los vectores pRG1301 (hlgG1) o pRG1368 (mlgG1) después de digestión con *Xho I/Sap I*. Para los análisis de expresión, los cuatro genes UHC (1-4) en pIDTSMART se subclonaron en los sitios *Xho I/Sap I* del vector pRG 1301 (hlgG1) o pRG1368 (mlgG1), y cada construcción de expresión se transfectó por separado en células CHO. Después de la transfección, las cuatro secuencias de VDJ candidatas se expresaron a un nivel suficiente y los péptidos expresados pudieron emparejarse con diferentes cadenas ligeras  $\kappa$  y  $\lambda$ .

Para evitar posibles anticuerpos autorreactivos que pudiesen conducir a un empobrecimiento de linfocitos B en los ratones modificados genéticamente, la base de datos de anticuerpos ASAP de Regeneron Pharmaceuticals, que se generó de los anticuerpos producidos por ratones humanizados VELOCIMMUNE®, se investigó para anticuerpos que contenían una secuencia de aminoácidos que es similar a  $hV_{H3-23}$  (D4-4)  $J_{H4}$  (Fig. 5). Más específicamente, los criterios que se utilizaron para identificar anticuerpos no reactivos incluían la secuencia de aminoácidos de DYSNY (SEQ ID NO: 144) o secuencias similares a DYSNY (SEQ ID NO: 144). Sin embargo, estudios de expresión en células CHO, revelaron que las secuencias UHC que contenían DYSNY (SEQ ID NO: 144) no se expresaban bien en células de mamífero. Por lo tanto, otra secuencia que carecía de D pero que tenía la secuencia YSNY (es decir, el anticuerpo H1H2002B; AKGYFDY (SEQ ID NO: 143); en la que AK es de 3-23; GY es de una adición de D o N o de adiciones N y P; e YFDY es  $J_{H4}$ ) se seleccionó en su lugar y su expresión se ensayó en células CHO. La secuencia UHC modificada se expresó a un nivel suficiente en células CHO. Estos resultados sugieren que, para la correcta expresión de la secuencia VDJ reordenada en células de mamífero, entre la secuencia codificada por un segmento génico V de cadena pesada y la secuencia codificada por un segmento génico J de cadena pesada se requieren algunos restos de aminoácido (es decir, un espaciador).

Además, los niveles de expresión del péptido AKGYFDY, procedente de la secuencia VDJ reordenada ( $V_{H3-23}/GY/J_{H4}$ ; HIH2002B) en células CHO, se compararon con los del péptido procedente de  $V_{H3-23}/D_{H4-4}$  (fase de lectura 2)/ $J_{H4}$  (SEQ ID NO: 148), con respecto a la expresión con cinco cadenas  $\kappa$  humanas, tres cadenas  $\lambda$  humanas y otras secuencias VDJ reordenadas (es decir,  $V_{H3-20}$  y  $V_{H1-39}$ ). La secuencia VDJ reordenada seleccionada ( $V_{H3-23}/GY/J_{H4}$ ) mostró niveles de expresión equivalentes a los de los controles.

Basándose en estos datos, para la creación de un ratón modificado genéticamente, la secuencia  $V_{H3-23}/GY/J_{H4}$  (SEQ ID NO: 137; HIH2002B) se seleccionó como una secuencia de dominio variable de cadena pesada reordenada. En las Figs. 1-9 se ilustran estrategias de direccionamiento detalladas para la generación de un ratón que contiene un locus de inmunoglobulina modificado genéticamente que codifica un dominio variable de cadena

pesada reordenada (es decir, un ratón que comprende un locus de inmunoglobulina que comprende una región variable de cadena pesada de inmunoglobulina humana reordenada) y se describen más adelante.

## 5 Ejemplo 2. Construcción de locus de cadena pesada de inmunoglobulina que contienen una secuencia VDJ reordenada

La construcción de locus de cadena pesada de inmunoglobulina que contienen una secuencia VDJ humana reordenada se realizó mediante series de reacciones de recombinación homóloga en células bacterianas (BHR) utilizando ADN de Cromosoma Artificial Bacteriano (BAC, *Bacterial Artificial Chromosome*). Para la creación de un ratón modificado genéticamente que expresase el dominio variable de cadena pesada reordenada, se generaron diversas construcciones diana utilizando tecnología de ingeniería genética VELOCIGENE® (véase, por ejemplo, la Patente de Estados Unidos n.º 6.586.251 y Valenzuela, D. M *et al* (2003), High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, *Nature Biotechnology* 21 (6): 652-659).

15 En resumen, para introducir una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada (es decir, hV<sub>H</sub>3-23(D)J<sub>H</sub>4; SEQ ID NO: 136) en un ratón modificado genéticamente en el que todos o sustancialmente todos los segmentos génicos V, D, J de cadena pesada de inmunoglobulina funcional endógena se habían delecionado, se diseñaron vectores diana. Además, los vectores diana incluían una región genómica que comprendía los genes Adam6a y Adam6b para evitar problemas de fertilidad asociados con la delección de la región genómica que comprende los genes Adam6a/6b en los ratones (véase, por ejemplo, el documento US 2012-03221 08A1).

25 Inicialmente, se construyó un donante BHR para la modificación de un clon BAC de ratón que comprendía una secuencia líder (que conduce la cadena pesada a través del retículo endoplasmático), una secuencia de nucleótidos de región variable de cadena pesada reordenada (V<sub>H</sub>3-23(D)<sub>H</sub>J<sub>H</sub>4; SEQ ID NO : 136) y un intrón de hJ<sub>H</sub>4 (SEQ ID NO: 140) que están unidos operativamente a un promotor V<sub>H</sub>3-23 de 2239 pb (SEQ ID NO: 139). Adicionalmente, el locus genómico se flanqueó en 5' y 3' mediante cajas de homología de IgH de ratón para la recombinación homóloga con el clon MAAC1115 BAC (**Fig. 1**)

30 Además, para crear una construcción diana que contuviese una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, se realizaron las cinco modificaciones siguientes.

35 En primer lugar, se introdujo un casete de selección de espectinomicina cadena arriba del promotor V<sub>H</sub>3-23 (entre los sitios I-CeuI y SpeI) para generar pJSh0038 (minilocus UHC, SEQ ID NO: 142) (**Fig. 1**, Ligamiento 1. I-CeuI/SpeI (Amp + Espec)). El minilocus UHC contiene: (1) un casete de espectinomicina (Espec) con sitios I-CeuI/Ascl para el ligamiento; (2) un promotor hV<sub>H</sub>3-23 de 2239 pb (SEQ ID NO: 139); (3) una secuencia de nucleótidos hV<sub>H</sub>3-23(D)J<sub>H</sub>4 reordenada (SEQ ID NO: 136); (4) un intrón hJ<sub>H</sub>4 (SEQ ID NO: 140); y (5) cajas de homología de ratón para BHR (MAID 1115).

40 En segundo lugar, un casete de selección de higromicina (EM7-HIG) se dirigió al extremo 5' de la región genómica del clon MAID 1115 BAC, que contiene un casete de neomicina flanqueado por sitios loxP (P<sub>gk</sub>-Neo) cadena arriba de la región genómica IgM. La inserción del casete de higromicina delecionó el sitio loxO localizado en el extremo 5' del clon MAID 1115. Mediante selección con higromicina/kanamicina, se seleccionaron células bacterianas que contenían el clon BAC modificado genéticamente (VI432) (**Fig. 2**, 2. BHR (Hig+Kan)).

45 En tercer lugar, el mini locus UHC que se construyó en la Etapa 1, se direccionó cadena arriba del locus de IgM del clon VI432 BAC. La introducción del minilocus UHC reemplazó el casete de selección de neomicina floxeado con un nuevo casete de espectinomicina (VI443). Las células bacterianas que contenían los clones BAC modificados genéticamente (VI443) se seleccionaron mediante selección con espectinomicina e higromicina (**Fig. 2**; 3. BHR (Spec + Hig)).

50 En cuarto lugar, el clon VI421 BAC, que comprendía, de 5' a 3', (1) un gen Adam6a (presente en una dirección 3' a 5'); (2) un casete de neomicina (presente en una dirección 3' a 5') flanqueado por sitios FRT; (3) un gen Adam6b (presente en una dirección 3' a 5'); (4) la Región 1 de Control Intergénica (IGCR1; es decir, una región reguladora clave de recombinación V(D)J); y (5) un casete de espectinomicina (presente en una dirección 5' a 3'), se direccionó con la construcción pDBa0049, que contenía un casete de cloranfenicol (Cm); un sitio de restricción Ascl cadena arriba del gen de cloranfenicol y brazos de homología 5' y 3'. El direccionamiento de la construcción pDBa0049 eliminó la IGCR1 y el casete de espectinomicina del clon VI421, y en el gen Adam6b cadena abajo se introdujo un nuevo sitio de restricción Ascl y un casete de cloranfenicol. Las células bacterianas que contenían el clon direccionado satisfactoriamente (VI444) se seleccionaron mediante selección con cloranfenicol y kanamicina (**Fig. 3**; 4. BHR (Cm+Kan)).

65 En quinto lugar, la región genómica del clon VI444 BAC que contenía los genes Adam6a y/o 6b se introdujo cadena arriba en el locus genómico de la cadena pesada universal en el clon VI443 BAC entre los sitios I-CeuI y Ascl mediante digestión con enzimas de restricción y ligamiento (**Fig. 3**). Esta modificación introduce los genes Adam6a y/o 6b en el clon y reemplaza el casete de espectinomicina con un casete de neomicina, dando lugar a una

construcción diana final (MAID6031; VI445). Las células bacterianas (BHR) que contenían la construcción diana final (MAID 6031; VI445) se seleccionaron basándose en selección con higromicina y kanamicina (**Fig. 3**, ligamiento 5. I-Ceu/Ascl (Hig+Kan)).

- 5 La construcción diana final (MAID6031), para la creación de un locus genómico que contiene una secuencia de dominio variable de cadena pesada humana reordenada, contiene, de 5' a 3', (1) un brazo de homología en 5' que contiene aproximadamente 20000 pb de una secuencia genómica de ratón, cadena arriba del locus de la cadena pesada de Ig endógena; (2) un gen Adam6a; (3) un sitio FRT en 5'; (4) un casete de neomicina; (5) un sitio FRT en 3', (6) un gen Adam6b; (7) el promotor hV<sub>H</sub>3-23 de 2239 pb (SEQ ID NO: 139); (8) una secuencia de nucleótidos de cadena pesada de inmunoglobulina humana reordenada (hV<sub>H</sub>3-23(D)<sub>J<sub>H</sub>4</sub>; SEQ ID NO: 136); (9) un intrón h<sub>J<sub>H</sub>4</sub> (SEQ ID NO: 140); y (10) un brazo de homología en 3' que contiene aproximadamente 7400 pb de una secuencia genómica de ratón cadena abajo de los segmentos génicos J<sub>H</sub> del ratón.

- 15 La construcción diana final, ADN MAID6031 BAC, se linealizó y se sometió a electroporación en células ES aisladas del ratón heterocigoto 1661 (**Fig. 4**), que contenía locus genómicos VDJ de cadena pesada de Ig de tipo silvestre y un locus genómicos VDJ mutados en el que todos los genes V<sub>H</sub>, D, J<sub>H</sub> se habían delecionado. Las células ES de ratón direccionadas satisfactoriamente se exploraron utilizando las sondas y los cebadores expuestos en las **Figuras 6-8**. Las células ES de ratón direccionadas satisfactoriamente se introdujeron en embriones de ratón hospedador utilizando tecnología VELOCIMOUSE® para producir un ratón F0 heterocigoto modificado genéticamente. Para generar ratones (MAID 6032 het) sin el casete de selección (es decir, FRT-Ub-Neo-FRT), las células ES direccionadas satisfactoriamente se sometieron a electroporación con un plásmido que expresa Flp recombinasa antes de su introducción en los embriones de ratón hospedador. Como alternativa, ratones macho heterocigotos MAID6031 que llevaban el casete de selección se cruzaron con ratones hembra que expresaban Flp recombinasa para eliminar el casete. Los ratones heterocigotos portadores de la modificación se cruzaron entre sí para generar homocigotos (MAID 6032 HO) que tienen la capacidad de crear cadenas pesadas de inmunoglobulina solo a partir del locus modificado genéticamente.

### **Ejemplo 3. Caracterización de ratones modificados genéticamente que expresan un dominio variable de cadena pesada reordenada**

- 30 Todos los ratones se instalaron en jaulas y se reprodujeron en Regeneron Pharmaceuticals, en condiciones específicas libres de patógenos. Se sacrificaron tres ratones de control de la misma camada de tipo silvestre (TS) (16 semanas de vida, machos, n = 2; Fondo: 75 % C57/BL6 y 25 % 129) y de dos a cuatro ratones MAID 6032 F0 HET (Fig. 9; 9 semanas de vida, machos, n = 2; Fondo: 50 % C57/BL6 y 50 % 129) y se extrajo la sangre, los bazo y la médula ósea de los animales. Adicionalmente, se sacrificaron cuatro ratones de control de la misma camada de tipo silvestre (TS) (10 semanas de vida; 2 machos y 2 hembras) y cuatro ratones MAID 6032 F2 homocigotos ("HO") (10 semanas de vida, 3 machos; 1 hembra) y se extrajo la sangre, los bazo y la médula ósea de los animales. La sangre se recogió en tubos BD microtainer con EDTA (Cat n.º 365973). La médula ósea se extrajo de los fémures lavando abundantemente con medio RPMI completo (medio RPMI complementado con suero bovino fetal, piruvato sódico, Hepes, 2-mercaptoetanol, aminoácidos no esenciales y gentamicina). Las preparaciones de eritrocitos de sangre periférica, bazo y médula ósea se lisaron con tampón de lisis ACK y se lavaron con medio RPMI completo.

#### **Citometría de flujo**

- 45 Para examinar la capacidad de los ratones F0 heterocigotos modificados genéticamente (MAID 6032 HET) descritos en el presente documento para producir anticuerpos procedentes del alelo modificado genéticamente (es decir, del alelo que contiene una sola copia del V-3-23/D/J<sub>H</sub>4 reordenado), se realizó un análisis de separación de células activadas por fluorescencia (FACS) utilizando células sanguíneas, de bazo, y de médula ósea aislada de un ratón de tipo silvestre o heterocigoto 6032.

- 50 En resumen, durante 10 minutos  $1 \times 10^6$  células se incubaron en hielo con anticuerpos anti-CD16/CD32 de ratón (2.4G2, BD), seguido de marcaje con el siguiente cóctel de anticuerpos durante 30 minutos en hielo: anti IgM<sup>a</sup> de ratón marcado con FITC (DS-1, BD Biosciences), anti CD3 conjugado con Pacific blue (17A2, BioLegend), anti CD19 marcado con APC-H7 (1 D3, BD Biosciences) y anti IgM<sup>b</sup> marcado con PE (AF6-78, BioLegend). Las células teñidas se lavaron y se fijaron en formaldehído al 2%. La adquisición de los datos se realizó en el citómetro de flujo LSRFortessa de BD y se analizaron con FlowJo. Las células de médula ósea, bazo y sanguíneas aisladas de un ratón de tipo silvestre o heterocigoto F0 6032 se seleccionaron en singletes y se separaron basándose en la expresión de CD19 (un marcador de linfocitos B) o la expresión de CD3 (un marcador de linfocitos T). Además, los linfocitos B seleccionados CD19+ se separaron basándose en la presencia de anticuerpos de IgM<sup>b</sup> (anticuerpos de IgM producidos a partir de un alelo de tipo silvestre (alelo B6)) o anticuerpos de IgM<sup>a</sup> (anticuerpos producidos a partir de un alelo modificado genéticamente (alelo 129) que comprende una secuencia de nucleótidos de región variable de cadena pesada reordenada (hV<sub>H</sub>3-23(D)<sub>J<sub>H</sub>4</sub>). El análisis FACS (**Fig. 11-12**) sugirió que los ratones heterocigotos, con respecto al alelo direccionado (es decir que contenía una copia de la secuencia variable de cadena pesada reordenada; MAID 6032 het), podían producir anticuerpos de IgM principalmente procedentes del alelo 129 (IgM<sup>a</sup>) modificado genéticamente.

Para examinar la capacidad de los ratones F2 homocigotos modificados genéticamente (MAID 6032 HO) descritos en el presente documento para producir anticuerpos procedentes del alelo modificado genéticamente (es decir, del alelo que contiene una sola copia de la V<sub>H</sub>3-23/D/J<sub>H</sub>4 reordenada), se realizó el análisis de separación de células activadas por fluorescencia (FACS) como se ha descrito anteriormente, utilizando células de bazo y médula ósea aisladas de un ratón de tipo silvestre u homocigoto 6032.

Únicamente los linfocitos B maduros pueden entrar en los folículos linfoides del bazo y ganglios linfáticos y por tanto participar de un modo eficaz en la respuesta inmunitaria. Los linfocitos B maduros, de vida larga, proceden de precursores de vida corta generados en la médula ósea. La selección en el grupo de maduros es un proceso activo y tiene lugar en el bazo. Se han identificado dos poblaciones de linfocitos B esplénicos como precursores de linfocitos B maduros. Los linfocitos B transicionales de tipo 1 (T1) son inmigrantes recientes de la médula ósea. Estos se desarrollan en los linfocitos B transicionales de tipo 2 (T2), que son cíclicos y se encuentran exclusivamente en los folículos primarios del bazo. Los linfocitos B maduros pueden generarse a partir de linfocitos B T1 o T2. Loder, F. *et al.*, J. Exp. Med., 190(1): 75-89, 1999.

El análisis FACS (**Figs. 13A y 13B**) sugirió que los ratones homocigotos, con respecto al alelo direccionado (es decir, que contiene dos copias de la secuencia variable de cadena pesada reordenada: MAID 6032 HO) pudieron producir poblaciones de linfocitos B inmaduros y maduros esplénicos normales, aunque con una ligera disminución en las secuencias lambda con respecto al tipo silvestre (**Figs. 13C y 13D**). Además en el bazo, los ratones MAID 6032HO demostraron una ligera disminución en los linfocitos B de la población T1 y un aumento en los linfocitos B de la zona marginal (**Fig. 13E**).

En la médula ósea, los ratones MAID6032 HO produjeron poblaciones de linfocitos B casi normales (**Figs. 14A-14E**) con un uso de secuencias lambda que era la mitad del de tipo silvestre (**Fig. 14F**).

### Estudios de inmunización

Cinco ratones TS (fondo 75 % C57BL6/25 % 129) y de tres a cuatro ratones MAID 6032 HET se inmunizaron en la almohadilla plantar con 0,025 ml de una mezcla que contenía 2,35 µg de un antígeno X, 10 µg de oligonucleótido CpG (ODN) 1826, InvivoGen, cat n.º tlr-1826) y 25 µg de adyuvante de gel fosfato de aluminio (Brenntag cat n.º 7784-30-7). Los ratones recibieron un refuerzo seis veces con la misma dosificación. Los días 15 y 24 posteriores a la primera inmunización, se extrajo sangre de ratones anestesiados utilizando un sangrado retroorbital en tubos separadores con suero BD (BD, cat. n.º 365956), y el suero se recogió siguiendo las instrucciones del fabricante.

Para medir los niveles de los anticuerpos de IgG específicos de antígeno y explorar con contraste la etiqueta mmh (myc-myc-his), se revistieron placas de ELISA (Nunc) con 1 µg/ml de un antígeno X incubado durante una noche a 4 grados C. El exceso de antígeno se eliminó con lavado antes de bloquear con PBS+BSA al 1 % durante 1 hora a TA. Se aplicaron diluciones en serie de suero y las placas se incubaron durante 1 hora a TA antes del lavado. Las placas se incubaron con anticuerpo anti IgG (cat. n.º 1030-05, Southern Biotech) conjugado con peroxidasa de rábano picante (HRP, *horseradish peroxidase*) durante 1 hora a TA. Después del lavado, las placas se revelaron con sustrato TMB (cat n.º 555214, BD). Las reacciones se detuvieron con ácido sulfúrico 1 N y la D.O. se leyó a 450 nm, utilizando un Lector Victor X5 de Perkin Elmer. Los datos se analizaron con el programa GraphPad Prism para calcular la dilución de suero que se encontraba dos veces por encima del fondo. Todos los experimentos con los animales fueron aprobados por la IACUC y por Regeneron Pharmaceuticals.

Como se muestra en la **Fig. 15**, los ratones F0 y F1 modificados genéticamente (MAID 6032 het), que son heterocigotos con respecto al alelo diana (es decir, que contienen una copia de la secuencia de nucleótidos V<sub>H</sub>3-23/D/J<sub>H</sub>4 reordenada), pudieron producir anticuerpos anti IgG específicos a niveles comparables a los producidos por los ratones de tipo silvestre los Días 15 y 24 posteriores a la primera inmunización.

### Ejemplo 3. Generación y análisis de ratones que comprenden dos segmentos V de cadena ligera humana

*Ejemplo 3.1: construcción del vector diana para la generación de ratones que comprenden dos segmentos V de cadena ligera humana*

Se construyeron (**Fig. 20**) dos locus de cadena ligera modificada por ingeniería genética, que contenían dos segmentos génicos Vk humanos (por ejemplo, un segmento génico Vk1-39 humano y otro Vk3-20 humano; es decir, una doble cadena ligera ("DLC")). Un locus de cadena ligera modificado por ingeniería genética contenía dos segmentos génicos Vk humanos y cinco segmentos génicos Jk humanos en una configuración no reordenada (DLC-5J). El segundo locus de cadena ligera modificado por ingeniería genética contenía dos segmentos génicos Vk humanos y un segmento génico Jk humano en configuración no reordenada (DLC-1J). Para cada uno de los dos locus de cadena ligera adicionales modificados por ingeniería genética, los segmentos génicos humanos se flanquearon en 3' con secuencias señal de recombinación para permitir la reordenación *in vivo* de los segmentos génicos humanos en linfocitos B.

- 5 *Modificación por ingeniería genética y Generación de Ratones DLC-1J.* En la **Fig. 21** se representan las etapas de modificación por ingeniería genética que dieron como resultado la generación de un locus de cadena ligera que comprendía dos segmentos génicos Vk humanos (Vk1-39 y Vk3-20) y un segmento génico Jk humano (Jk5), denominado de otra manera DLC-1J. Específicamente, las secuencias Vk1-39 y Vk3-20 humanas se amplificaron por PCR a partir de moldes BAC (Invitrogen), y conjuntamente con una secuencia amplificada que contenía una secuencia señal recombinante (rss, siglas de *recombination signal sequence*) y un segmento Jk5 humano, se clonaron mediante un ligamiento de cuatro vías en un plásmido que contenía un casete de selección de UB-higromicina (**Fig. 21A**). Los brazos 5' y 3' se conectaron como se representa en las **Figs. 21B y 21C**.
- 10 La construcción diana resultante se representa en la **FIG. 21C** (diagrama inferior; DLC-1 J), con secuencias señal de recombinación (RSS) en óvalos blancos. El clon de ADN BAC modificado del locus de cadena ligera DLC-1J modificado por ingeniería genética unido operativamente a secuencias de ratón (es decir, secuencias cadena arriba y cadena abajo del locus de cadena ligera  $\kappa$  de inmunoglobulina endógena) se confirmó por PCR, utilizando cebadores localizados en las secuencias dentro del locus de cadena ligera modificado por ingeniería genética, que
- 15 contenía los dos segmentos génicos Vk humanos, seguido de electroporación en células ES que comprendían la delección del locus variable Igk de ratón (que comprende los segmentos génicos de unión y variable  $\kappa$ ) (**Fig. 21D**) para crear un ratón que exprese cualquiera de los dos segmentos génicos Vk humanos. Los clones positivos de células ES que contenían el locus de cadena ligera DLC-1J modificado por ingeniería genética, se confirmaron mediante exploración Taqman™ y cariotipado utilizando sondas específicas para el locus de cadena ligera DLC-1J
- 20 modificado por ingeniería genética. En la siguiente Tabla 8 se representan las secuencias de los cebadores y sondas utilizados para la exploración de células ES de células ES con DLC-1J y se incluyen en el Listado de Secuencias.

**Tabla 8: cebadores y sondas utilizados para la exploración de células ES**

Nombre de la Sonda	Ensayo/tipo de sonda	Localización detectada	Secuencia de la Sonda	Cebador Directo	Cebador Inverso
1633h2	GOA/TAQMAN™	Vk1-39	ATCAGCAGAA ACCAGGGAAA GCCCCT (SEQ ID NO:44)	GGGCAAG TCAGAGC ATTAGCA (SEQ ID NO:45)	TGCAAACCTGG ATGCAGCATA G (SEQ ID NO:46)
1635h2	GOA/TAQMAN™	Vk3-20	AAAGAGCCAC CCTCTCCTGC AGGG (SEQ ID NO:65)	TCCAGGC ACCCTGTC TTTG (SEQ ID NO:66)	AAGTAGCTGC TGCTAACACT CTGACT (SEQ ID NO:67)
Neo	GOA	neo	TGGGCACAAC AGACAATCGG CTG (SEQ ID NO:38)	GGTGGAG AGGCTATT CGGC (SEQ ID NO:39)	GAACACGGC GGCATCAG (SEQ ID NO:40)
Jxn 1-39/3-20	GOA/BHQ1	1-39/3-20 unión BamHI	TCTTTTGCCCC <b>GGATCC</b> GATC AG (SEQ ID NO:84; sitio de restricción en negrita)	GGGAGGC TCCTCTGA ACTCTAAG (SEQ ID NO:85)	GTCCAGTCAC TCGGTTGCTA T (SEQ ID NO:86)

25

Después, se utilizaron clones confirmados de células ES para implantar en ratones hembra para producir una camada de crías que comprendían un locus de cadena ligera DLC-1J y que expresaban un dominio variable de

cadena ligera humana fusionado a un dominio Ck de ratón. Las secuencias de las sondas y cebadores utilizados para el genotipado de las crías se indican en la Tabla 8 anterior. La secuencia completa del locus DLC-1J modificado por ingeniería genética, incluyendo 100 nucleótidos de la secuencia de ratón cadena arriba y cadena abajo de la secuencia insertada modificada por ingeniería genética, se presenta en las **Figs. 22A-22D** y se expone en la SEQ ID NO: 82.

Las células ES portadoras del locus de cadena ligera modificado por ingeniería genética pueden transfectarse con una construcción que expresa FLP para eliminar el casete de neomicina flanqueado con FRT introducido mediante la construcción diana (véase la **Fig. 21E**). Opcionalmente, el casete de neomicina se elimina cruzando a los ratones que expresan FLP recombinasa FLP (por ejemplo, documento US 6.774.279). Opcionalmente, el casete de neomicina se conserva en los ratones.

*Modificación por ingeniería genética y generación de ratones DLC-5J.* Para generar un locus de cadena ligera que comprenda dos segmentos génicos Vk humanos (Vk1-39 y Vk3-20) y cinco segmentos génicos Jk humanos (Jk1, Jk2, Jk3, Jk4 y Jk5), denominado de otra manera DLC-5J, una secuencia amplificada de 2000 pares de bases que comprendía los 5 segmentos génicos Jk humanos, se ligó en un vector que comprendía dos segmentos génicos Vk humanos y un segmento génico Jk humano, representado en la **Fig. 21B** (central) (véase la **Fig. 23A**). Las etapas de modificación genética posteriores implican la conexión de brazos 3' y 5' como se representa en la **Fig. 23B**.

La construcción diana resultante se representa en la **Fig. 23B** (diagrama inferior; DLC-5J), con secuencias señal de recombinación (RSS) en óvalos blancos. El clon de ADN BAC modificado del locus de cadena ligera DLC-5J modificado por ingeniería genética unido operativamente a secuencias de ratón (es decir, secuencias cadena arriba y cadena abajo del locus de cadena ligera κ de inmunoglobulina endógena) se confirmó por PCR utilizando los cebadores localizados en las secuencias dentro del locus de cadena ligera modificado por ingeniería genética que contenía los dos segmentos génicos Vk humanos, seguido de electroporación en células ES que comprende la delección del locus variable de Igκ de ratón (que comprende los segmentos génicos de unión y variable κ) (**Fig. 23C**) para crear un ratón que exprese cualquiera de los dos segmentos génicos Vk humanos. Los clones positivos de células ES que contenían el locus de cadena ligera DLC-5J modificado por ingeniería genética se confirmaron mediante exploración Taqman™ y cariotipado utilizando sondas específicas para el locus de cadena ligera DLC-5J modificado por ingeniería genética. Las secuencias de cebadores y sondas utilizados para la exploración de células ES de células ES de células ES con DLC-5J se representan en la Tabla 9 a continuación y se incluyen en el Listado de Secuencias.

Tabla 9: cebadores y sondas utilizados para la exploración de células ES

Nombre de la Sonda	Ensayo/tipo de sonda	Localización detectada	Secuencia de la Sonda	Cebador Directo	Cebador Inverso
1633h2	GOA/TAQMAN™	Vk1-39	ATCAGCAGAA ACCAGGGAAA GCCCT (SEQ ID NO:44)	GGCAAG TCAGAGC ATTAGCA (SEQ ID NO:45)	TGCAAACTGG ATGCAGCATA G (SEQ ID NO:46)
1635h2	GOA/TAQMAN™	Vk3-20	AAAGAGCCAC CCTCTCCTGC AGGG (SEQ ID NO:65)	TCCAGGC ACCCTGTC TTTG (SEQ ID NO:66)	AAGTAGCTGC TGCTAACACT CTGACT (SEQ ID NO:67)
Neo	GOA	neo	TGGGCACAAC AGACAATCGG CTG (SEQ ID NO:38)	GGTGGAG AGGCTATT CGGC (SEQ ID NO:39)	GAACACGGC GGCATCAG (SEQ ID NO:40)
Jxn 1-39/3-20	GOA/BHQ1	1-39/3-20 unión BamHI	TCTTTTGCCCC <b>GGATCCGATC</b> AG (SEQ ID NO:84; sitio de restricción en negrita)	GGGAGGC TCCTCTGA ACTCTAAG (SEQ ID NO:85)	GTCCAGTCAC TCGGTTGCTA T (SEQ ID NO:86)

Nombre de la Sonda	Ensayo/tipo de sonda	Localización detectada	Secuencia de la Sonda	Cebador Directo	Cebador Inverso
Jxn3-20/Jk1-5	GOA/BHQ1	3-20/Jk1-5 unión BsiWI	CTTCAACTGTG <b>GCGTACGCAC</b> C (SEQ ID NO:87, sitio de restricción en negrita)	ACGCAGA TGTAGCCA AACCCCT (SEQ ID NO:88)	CAGCTGCTGA AGCTCAACTC (SEQ ID NO:89)

Después, se utilizaron clones confirmados de células ES para implantar en ratones hembra para producir una camada de crías que comprendían un locus de cadena ligera DLC-5J y que expresaban un dominio variable de cadena ligera humana fusionado a un dominio C $\kappa$  de ratón. Las secuencias de las sondas y cebadores utilizados para el genotipado de las crías se indican en la Tabla 9 anterior. La secuencia completa del locus DLC-5J modificado por ingeniería genética, incluyendo 100 nucleótidos de la secuencia de ratón cadena arriba y cadena abajo de la secuencia insertada modificada por ingeniería genética, se presenta en las **Figs. 24A-24D** y se expone en la SEQ ID NO: 83.

Las células ES portadoras del locus de cadena ligera modificado por ingeniería genética pueden transfectarse con una construcción que expresa FLP para eliminar el casete de neomicina flanqueado con FRT introducido mediante la construcción diana (véase la **Fig. 23D**). Opcionalmente, el casete de neomicina se elimina cruzando a los ratones que expresan FLP recombinasa FLP (por ejemplo, documento US 6.774.279). Opcionalmente, el casete de neomicina se conserva en los ratones.

### 15 *Ejemplo 3.2: caracterización de ratones que comprenden dos segmentos V humanos*

*Citometría de Flujo.* Las poblaciones de linfocitos B y el desarrollo de linfocitos B en ratones con DLC se validaron mediante análisis de citometría de flujo de preparaciones de esplenocitos y médula ósea. Se realizaron suspensiones celulares de ratones homocigotos para dos segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos (n=4), de ratones homocigotos para dos segmentos génicos V $\kappa$  humanos y un segmento génico J $\kappa$  humano (n=4), y de ratones de tipo silvestre (n=4) utilizando métodos estándar y se tiñeron con anticuerpos marcados con fluorescencia.

En resumen, durante 10 minutos,  $1 \times 10^6$  células se incubaron en hielo con anti CD16/CD32 de ratón (clon 2.4G2, BD Pharmigen), seguido de marcaje con el siguiente cóctel de anticuerpos durante 30 minutos en hielo: anti CD19 de ratón conjugado con APC-H7 (clon 1 D3, BD Pharmigen), anti CD3 de ratón conjugado con Pacific Blue (clon 17A2, BioLegend), anti Ig $\kappa$  de ratón conjugado con FITC (clon 187.1, BD Pharmigen) o anti CD43 de ratón (clon 1 B11, BioLegend), anti Ig $\lambda$  de ratón conjugado con PE (clon RML-42, BioLegend) o anti c-kit de ratón (clon 2B8, BioLegend), anti IgD de ratón conjugado con PerCP-Cy5.5 (BioLegend), anti IgM de ratón conjugado con PE-Cy7 (clon II/41, eBioscience), anti B220 de ratón conjugado con APC (clon RA3-6B2, eBioscience). Después de la tinción, las células se lavaron y se fijaron en formaldehído al 2 %. La adquisición de los datos se realizó en un citómetro de flujo LSRII y se analizaron con FlowJo (Tree Star, Inc.). Selección: linfocitos B totales (CD19+CD3-), linfocitos B Ig $\kappa$ + (Ig $\kappa$ +Ig $\lambda$ -CD19+CD3-), linfocitos B Ig $\lambda$ + (Ig $\kappa$ -Ig $\lambda$ +CD19+CD3-). Los resultados del compartimento de médula ósea se muestran en las **Figs. 25A-27B**. Los resultados del compartimento esplénico se muestran en las **Figs. 28A-31**.

Como se muestra en este ejemplo, los ratones con DLC-5J demuestran poblaciones de linfocitos B normales en los compartimentos esplénico y de médula ósea (**Figs. 25A-31**). Los ratones con DLC-5J demuestran poblaciones de linfocitos B inmaduros, maduros y pre/pro, en el compartimento de médula ósea que son sustancialmente iguales a las observadas en camadas de tipo silvestre. De hecho, el locus DLC-5J fue capaz de competir con el locus de cadena ligera lambda endógeno para dar una proporción kappa:lambda que es sustancialmente igual a la observada en los ratones de tipo silvestre (**Fig. 27B**). Además, los ratones DLC-5J demuestran un desarrollo normal de linfocitos B periféricos a medida que los linfocitos B progresan a través de diversos estadios en el compartimento esplénico (por ejemplo, inmaduros, maduros, T1, T2 T3, precursores de zona marginal, zona marginal, folicular-I, folicular- II, etc.) de una manera sustancialmente igual a la observada en los ratones de tipo silvestre (**Figs. 30A-31**). En cambio, los ratones DLC-1J demostraron un número global inferior de linfocitos B y un uso de cadena ligera lambda aumentado en comparación con el de la cadena ligera kappa modificada por ingeniería genética (datos no mostrados).

*Expresión de Doble de Cadena Ligera.* La expresión de ambos segmentos génicos V $\kappa$  humanos se analizó en ratones homocigotos utilizando un ensayo de PCR cuantitativa. En resumen, linfocitos B CD19+ se purificaron de médula ósea y de bazo intactos de ratones de tipo silvestre, de ratones homocigotos para un reemplazo de la cadena pesada de ratón y locus variable de cadena ligera  $\kappa$  con locus de región variable de cadena ligera  $\kappa$  y cadena pesada humana correspondientes (H $\kappa$ ), así como de ratones homocigotos para un locus de cadena ligera  $\kappa$  modificada por ingeniería genética que contiene dos segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos (DLC-5J) o un segmento génico J $\kappa$  humano (DLC-1 J). La expresión relativa se normalizó a la expresión de la región C $\kappa$  de ratón (n=3 a 5 ratones por grupo). En las **Figs. 32** y **33** se muestran los resultados.

La expresión de las cadenas ligeras que contienen un segmento génico V $\kappa$ 3-20 humano o V $\kappa$ 1-39 humano reordenado se detectaron tanto en la médula ósea como en el bazo de ratones con DLC-5J y DLC-1J (**Figs. 32** y **33**). En el compartimento de médula ósea, la expresión de las cadenas ligeras procedentes tanto de V $\kappa$ 3-20 humana como V $\kappa$ 1-39 humana en ambas cepas de ratones con DLC, era significativamente superior en comparación con la de los ratones que comprendían un reemplazo del segmento génico V $\kappa$  y J $\kappa$  de ratón con los segmentos génicos V $\kappa$  y J $\kappa$  humanos correspondientes (H $\kappa$ , **Fig. 32**). Se observó que expresión de la cadena ligera procedente de V $\kappa$ 3-20 humana era de aproximadamente seis veces (DLC-5J) a quince veces (DLC-1J) mayor que en ratones con H $\kappa$ . En el compartimento de la médula ósea, los ratones con DLC-1J demostraron una expresión

aproximadamente dos veces mayor de cadenas ligeras procedentes de la Vk3-20 humana sobre los ratones con DLC-5J. Se observó que la expresión de la cadena ligera procedente de la Vk1-39 humana era de aproximadamente seis veces (DLC-5J) a trece veces (DLC-1J) mayor que la de los ratones con Hk. En el compartimento de la médula ósea, los ratones con DLC-1J demostraron una expresión de aproximadamente dos veces mayor de las cadenas ligeras procedentes de Vk1-39 humana sobre la de los ratones con DLC-5J.

En el compartimiento esplénico, la expresión de las cadenas ligeras procedentes de Vk3-20 humana y Vk1-39 humana en ambas cepas de ratones con DLC, fue significativamente mayor en comparación con la de ratones con Hk (**Fig. 33**). Se observó que la expresión de la cadena ligera procedente de Vk3-20 humana era aproximadamente cuatro veces (DLC-5J) y ocho veces (DLC-1J) mayor que la de ratones con Hk. En el compartimento esplénico, los ratones con DLC-1J demostraron una expresión de aproximadamente dos veces mayor de cadenas ligeras procedentes de Vk3-20 humana sobre la de los ratones con DLC-5J. Se observó una expresión de cadena ligera procedente de Vk1-39 humana de aproximadamente cuatro veces (DLC-5J) a cinco veces (DLC-1J) mayor que en ratones con Hk. En el compartimento esplénico, los ratones con DLC-1J demostraron una expresión similar de cadenas ligeras procedentes de Vk1-39 humana en comparación con los ratones con DLC-5J.

*Uso de los segmentos Vk/Jk Humano en Ratones con DLC-5J.* Ratones homocigotos para dos segmentos génicos Vk humanos no reordenados y cinco segmentos génicos Jk humanos no reordenados (DLC-5J) se analizaron con respecto al uso de segmentos génicos Vk/Jk humanos en linfocitos B esplénicos mediante reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR).

En resumen, se extrajeron los bazo de ratones homocigotos DLC-5J (n=3) y de tipo silvestre (n=2) y para crear suspensiones unicelulares se mezclaron en 10 ml de RPMI 1640 (Sigma) que contenía suero bovino fetal termoinactivado al 10 % utilizando portaobjetos de vidrio esmerilados. Los esplenocitos se sedimentaron con una centrífuga (1200 rpm durante cinco minutos), y los eritrocitos se lisaron en 5 ml de tampón de lisis ACK (GIBCO) durante tres minutos. Los esplenocitos se diluyeron con PBS (Irvine Scientific), se filtraron con un filtrador celular de 0,7 µm y se centrifugaron de nuevo dando células sedimentadas, que fue seguido de resuspensión de 1 ml de PBS.

Utilizando el mini kit de ADN/ARN AllPrep (Qiagen) se aisló ARN de los esplenocitos sedimentados de acuerdo con las especificaciones del fabricante. Se realizó RT-PCR en el ARN de los esplenocitos utilizando el sistema 5' RACE (Amplificación Rápida de extremos de ADNc) con cebadores específicos para el gen Ck de ratón de acuerdo con las especificaciones del fabricante (Invitrogen). Los cebadores específicos para el gen Ck de ratón fueron 3' mlgKC RACE1 (AAGAAGCACA CGACTGAGGC AC; SEQ ID NO: 90) y mlgKC3'-1 (CTCACTGGAT GGTGGGAAGA TGGA; SEQ ID NO: 91). Los productos de la PCR se purificaron en gel y se clonaron en el vector pCR@2.1-TOPO® (Kit TOPO® TA Cloning®, Invitrogen) y se secuenciaron con el cebador Directo M13 (GTAAAACGAC GGCCAG; SEQ ID NO: 92) y el cebador Inverso M13 (CAGGAAACAG CTATGAC; SEQ ID NO: 93) localizados dentro del vector en lugares flanqueantes del sitio de clonación. Se secuenciaron diez clones de cada muestra de bazo. Las secuencias se compararon con las de los conjuntos de inmunoglobulina de ratón y ser humano de los conjuntos del directorio de referencia IMGT/V-QUEST para determinar el uso de Vk/Jk. En la Tabla 10 se exponen combinaciones de Vk/Jk para clones seleccionados observados en clones de RT-PCR de cada muestra de esplenocitos. En la Tabla 11 se expone la secuencia de aminoácidos de las uniones de Vk humano/Jk humano y Jk humano/Ck de ratón de los clones de RT-PCR seleccionados de ratones homocigotos para DLC-5J. Las letras minúsculas indican mutaciones en la secuencia de aminoácidos de la región variable o adiciones sin molde resultantes de adiciones de N y/o P durante la recombinación.

Como se muestra en este ejemplo, los ratones homocigotos para dos segmentos génicos Vk humanos no reordenados y cinco segmentos génicos Jk humanos no reordenados (DLC-5J) unidos operativamente al gen Ck de ratón, pudieron recombinar productivamente ambos segmentos génicos Vk humanos con segmentos génicos Jk humanos múltiples para producir un repertorio limitado de cadenas ligeras de inmunoglobulina. Entre los reordenamientos en los ratones homocigotos DLC-5J mostrados en la Tabla 10, únicamente se observaron reordenamientos Vk/Jk humanos para Vk1-39/Jk2 (1), Vk1-39/Jk3 (1), Vk3-20/Jk1 (7), Vk3-20/Jk2 (4) y Vk3-20/Jk3 (1). Además, dichos reordenamientos únicos demostraron diversidad de unión por la presencia de aminoácidos únicos dentro de la región CDR3 de la cadena ligera (Tabla 11), dando como resultado cualquier mutación y/o la recombinación de los segmentos génicos Vk y Jk humanos durante el desarrollo. Todas las reordenaciones mostraron lectura integral funcional en el segmento Ck de ratón (Tabla 11).

En su conjunto, estos datos demuestran que los ratones modificados por ingeniería genética presentan una elección de a lo sumo dos segmentos génicos Vk humanos, ambos de los cuales son capaces de reordenarse (por ejemplo, con uno o más y, en algunos aspectos, hasta cinco segmentos génicos Jk humanos) y codificar un dominio V<sub>L</sub> humano de una cadena ligera de inmunoglobulina que tiene cantidades de linfocitos B y desarrollo que es casi de tipo silvestre en todos los aspectos. Dichos ratones producen una colección de anticuerpos que tienen cadenas ligeras de inmunoglobulina que, en la colección, tienen uno de dos posibles segmentos génicos V<sub>L</sub> humanos presentes. El ratón produce esta colección de anticuerpos en respuesta a la exposición antigénica y, la colección de anticuerpos está asociada con una diversidad de cadenas pesadas (variable humana/constante de ratón) quiméricas inversas.

**Tabla 10: Combinaciones Vk/Jk observadas en muestras de esplenocitos**

Ratón ID n.º	Combinación Vk/Jk del Clon	Genotipo	
1089451	DLC-5J	1-2	1-39/3
		1-4	3-20/2
		1-7	3-20/1
		1-8	3-20/2
1089452	DLC-5J	2-2	3-20/1
		2-3	3-20/1
		2-6	3-20/2
		2-8	3-20/2
		2-9	3-20/1
		2-10	1-39/2
1092594	DLC-5J	3-1	3-20/1
		3-2	3-20/1
		3-4	3-20/1
		3-6	3-20/3
		3-9	3-20/2
1092587	TS	1-1	19-93/1
		1-2	6-25/1
		1-3	4-91/5
		1-5	3-10/4
		1-6	4-86/4
		1-8	19-93/1
		1-10	19-93/2
1092591	TS	2-1	19-93/1
		2-3	6-20/5
		2-4	6-25/5
		2-5	1-117/1
		2-6	8-30/1
		2-7	8-19/2
		2-8	8-30/1
		2-10	1-117/1

**Tabla 11: Secuencias de aminoácidos de las uniones Vk humano/Jk humano y Jk humano/Ck de ratón de ratones homocigotos para DLC 5J**

Clon	Vk/Jk	Secuencia de Unión hVk/hJk/mCk (la CDR3 se indica subrayada, la mlgkC en cursiva)	SEQ ID NO:
2-10	1-39/2	QPEDFATYYC <u>QQSYSTPYTF</u> FGQG <b>TKLEIKRADAAPT</b> <i>VS</i> I	94
1-2	1-39/3	QPEDFATYYC <u>QQSYSTPFT</u> FGPG <b>TKVDIKRADAAPT</b> <i>VS</i> I	95
1-7	3-20/1	EPEDFAVYYC <u>QQYGSSPr</u> TFGQG <b>TKVEIKRADAAPT</b> <i>VS</i> I	96
2-2	3-20/1	EPEDFAVYYC <u>QQYGSSr</u> TFGQG <b>TKVEIKRADAAPT</b> <i>VS</i> I	97
2-3	3-20/1	EPEDFAVYYC <u>QQYGSSPWTF</u> FGQG <b>TKVEIKRADAAPT</b> <i>VS</i> I	98
2-9	3-20/1	dPEDFAVYYC <u>QQYGSSPr</u> TFGQG <b>TKVEIKRADAAPT</b> <i>VS</i> I	99
3-1	3-20/1	EPEDFAVYYC <u>QQYGSSPr</u> TFGQG <b>TKVEIKRADAAPT</b> <i>VS</i> I	100
3-2	3-20/1	EPEDFAVYYC <u>QQYGSSPWTF</u> FGQG <b>TKVEIKRADAAPT</b> <i>VS</i> I	101
3-4	3-20/1	EPEDFAVYYC <u>QQYGSSPPT</u> FGQG <b>TKVEIKRADAAPT</b> <i>VS</i> I	102
3-9	3-20/2	EPEDFAVYYC <u>QQYGSSPYTF</u> FGQG <b>TKLEIKRADAAPT</b> <i>VS</i> I	103
3-6	3-20/3	EPEDFAVYYC <u>QQYGSSi</u> FTFGPG <b>TKVDIKRADAAPT</b> <i>VS</i> I	104

5

**Ejemplo 4: generación y caracterización de ratones que comprenden dos cadenas ligeras humanas sustituidas con dos histidinas**

10

*Ejemplo 4.1: modificación por ingeniería genética y generación de ratones que comprenden dos segmentos kappa V conteniendo cada uno de ellos cuatro sustituciones de histidina*

En los locus de doble cadena ligera de ratones ULC V $\kappa$ 1-39 y V $\kappa$ 3-20 se introdujeron sustituciones de histidina como se ha descrito anteriormente. En resumen, la secuencia DLC representada en la **Fig. 23A** (inferior), se sometió a mutagénesis dirigida, modificando primero la secuencia V $\kappa$ 1-39 y después modificando la secuencia V $\kappa$ 3-20, utilizando los cebadores representados en la **Fig. 34**. La secuencia de doble cadena ligera resultante contenía el segmento V $\kappa$ 1-39 con histidinas introducidas en la secuencia de la línea germinal en las posiciones 105, 106, 108 y 111, el segmento V $\kappa$ 3-20 con histidinas introducidas en la secuencia de la línea germinal en las posiciones 105, 106, 107 y 109 así como los cinco segmentos J $\kappa$  (J $\kappa$ 1, J $\kappa$ 2, J $\kappa$ 3, J $\kappa$ 4 y J $\kappa$ 5). Una etapa posterior de modificación por ingeniería genética implicó la conexión de un brazo 5' que lleva un casete FRT-UB-NEO-FRT y de un brazo 3' que lleva potenciadores de I $\mu$ k de ratón y una región constante. Este vector diana se sometió a electroporación en células ES que comprendían la delección del locus variable de I $\mu$ k de ratón (que comprende los segmentos génicos de unión y variables  $\kappa$ ), como se representa en la **Fig. 35A** (en esta figura no se indican las secuencias señal de recombinación, RSS). Las células ES diana se exploraron mediante una modificación de ensayo de alelo como se ha descrito anteriormente, utilizando cebadores y sondas que detectan las regiones descritas anteriormente en las Tablas 1, 5, 8 y 9 (específicamente, 1633h2, 1635h2, neo, J $\kappa$ n 1-39/3 -20, mlgk2 y mlgk15), así como dos conjuntos adicionales de cebadores y sondas indicados en la Tabla 12 más adelante. Las secuencias de estos dos conjuntos de cebadores y sondas adicionales se incluyen en el Listado de Secuencias.

**Tabla 12: cebadores y sondas utilizados para la exploración de células ES**

Nombre de la Sonda	Ensayo/tipo de sonda	Localización detectada	Secuencia de la Sonda	Cebador Directo	Cebador Inverso
hVI492 1-39	GOA/FAM-BHQ+	MAID 6185 (específico de 4 HIS-1-39)	AACTTACTACT GTCACCA (SEQ ID NO:111)	CAGCAGT CTGCAAC CTGAA (SEQ ID NO:112)	GGCTCGTCCT CACACATC (SEQ ID NO:113)
hVI492 3-20	GOA/FAM-BHQ+	MAID 6185 (específico de 4 HIS-1-39)	TTACTGTCAC CATCATG (SEQ ID NO:114)	GCAGACT GGAGCCT GAAGA (SEQ ID NO:115)	AAGCTGAATC ACTGTGGGAG GTG (SEQ ID NO:116)

Después, se utilizó un clon de células ES confirmado para implantar en ratones hembra para producir una camada de crías que comprenden el locus de cadena ligera DLC-5J con cuatro modificaciones de histidina en cada una de las secuencias del segmento V $\kappa$ L presente y que expresan un dominio variable de cadena ligera humana fusionado con un dominio C $\kappa$  de ratón. Algunas de las mismas secuencias que se utilizan para la exploración de células ES también se utilizan para el genotipado de las crías.

Las células ES portadoras del locus de cadena ligera modificado por ingeniería genética, pueden transfectarse con una construcción que expresa FLP (por ejemplo, FLPo) para eliminar el casete de neomicina flanqueado con FRT, introducido por la construcción diana (véase la **Fig. 35B**, en esta figura se omiten las RSS). Opcionalmente, el casete de neomicina se elimina cruzando a los ratones que expresan FLP recombinasa (por ejemplo, documento US 6.774.279). Opcionalmente, el casete de neomicina se conserva en los ratones.

*Ejemplo 4.2: modificación por ingeniería genética y generación de ratones que comprenden dos segmentos kappa V conteniendo cada uno de ellos tres sustituciones de histidina*

En cada uno de los segmentos V $\kappa$ 1-39 y V $\kappa$ 3-20 de los ratones con doble cadena ligera se introdujeron tres sustituciones de histidina. En resumen, la secuencia DLC representada en la **Fig. 23A** (inferior), se sometió a mutagénesis dirigida, modificando primero la secuencia V $\kappa$ 1-39 y después modificando la secuencia V $\kappa$ 3-20, utilizando los cebadores representados en la **Fig. 36**. La secuencia resultante de doble cadena ligera contenía el segmento V $\kappa$ 1-39 con histidinas introducidas en las secuencias de la línea germinal en las posiciones 106, 108 y 111, el segmento V $\kappa$ 3-20 con histidinas introducidas en la secuencia de la línea germinal en las posiciones 105, 106 y 109, así como los cinco segmentos J $\kappa$  (J $\kappa$ 1, J $\kappa$ 2, J $\kappa$ 3, J $\kappa$ 4 y J $\kappa$ 5). Una etapa posterior de modificación por ingeniería genética implicó la conexión de un brazo 5' portador de un casete FRT-UB-NEO-FRT y de un brazo 3' portador de potenciadores de I $\mu$ k y región constante de ratón. Este vector diana se sometió a electroporación en

células ES que comprendían la delección del locus variable I $\mu$ k de ratón (que comprende los segmentos génicos variable  $\kappa$  y de unión), como se representa en la **Fig. 37A** (en esta figura se omite la RSS). Las células ES diana se exploraron mediante una modificación de ensayo de alelo como se ha descrito anteriormente, utilizando sondas y cebadores que detectan las regiones descritas anteriormente en las Tablas 1, 5, 8 y 9 (específicamente, 1633h2, 1635h2, neo, Jxn 1-39/3 -20, mlgkd2 y mlgkp15), así como dos conjuntos de cebadores y sondas adicionales indicados en la siguiente Tabla 13. Las secuencias de estos dos conjuntos de cebadores y sondas adicionales se incluyen en el Listado de Secuencias.

**Tabla 13: cebadores y sondas utilizados para la exploración de células ES**

Nombre de la Sonda	Ensayo/tipo de sonda	Localización detectada	Secuencia de la Sonda	Cebador Directo	Cebador Inverso
hVI493 1-39	GOA/FAM-BHQ+	MAID 6187 (específico de 3 HIS-1-39)	CTTACTACTGT CAACATAG (SEQ ID NO:123)	CAGCAGT CTGCAAC CTGAA (SEQ ID NO:124)	GGCTCGTCCT CACACATC (SEQ ID NO:125)
hVI493 3-20	GOA/FAM-BHQ+	MAID 6187 (específico de 3 HIS-3-20)	TACTGTCAC CATTATGG (SEQ ID NO:126)	GCAGACT GGAGCCT GAAGA (SEQ ID NO:127)	AAGCTGAATC ACTGTGGGAG GTG (SEQ ID NO:128)

Después, se utilizó un clon de células ES confirmado para implantar en hembras de ratón para producir una camada de crías que comprende el locus de cadena ligera DLC-5J con cuatro modificaciones de histidina en cada una de las secuencias de segmento V<sub>L</sub> presentes, y que expresa un dominio variable de cadena ligera humano fusionado a un dominio C $\kappa$  de ratón. Algunas de las mismas secuencias que se utilizan para la exploración de células ES también se utilizan para el genotipado de las crías.

Las células ES portadoras del locus de cadena ligera modificado por ingeniería genética pueden transfectarse con una construcción que expresa FLP (por ejemplo, FLPo) para eliminar el casete de neomicina flanqueado con FRT introducido por la construcción diana (véase la **Fig. 37B**, en esta figura se omite la RSS). Opcionalmente el casete de neomicina se elimina cruzando a los ratones que expresan FLP recombinasa (por ejemplo el documento US 6.774.279). Opcionalmente el casete de neomicina se conserva en los ratones.

*Ejemplo 4.3: cría de ratones que comprenden dobles cadenas ligeras sustituidas con histidina humana*

Ratones portadores de un locus de doble cadena ligera sustituida con histidina humana se cruzaron con ratones que contenían una delección del locus de cadena ligera  $\lambda$  endógena para generar una descendencia que exprese, solo en sus cadenas ligeras, las cadenas ligeras sustituidas con histidina modificadas por ingeniería genética procedentes del locus de doble cadena ligera.

Ratones portadores de un locus de doble cadena ligera sustituida con histidina humana se cruzaron con ratones que contenían un reemplazo del locus variable de cadena pesada de ratón endógeno con locus variable de cadena pesada humana (véanse los documentos US 6.596.541 y US 8.502.018; el ratón VELOCIMMUNE®, Regeneron Pharmaceuticals, Inc.).

*Ejemplo 4.4: detección de modificaciones de histidina en cadenas ligeras de inmunoglobulina obtenidas de ratones que comprenden dos segmentos kappa V cada uno de ellos conteniendo tres sustituciones de histidina*

Se prepararon amplicones kappa V de ARNm de linfocitos B esplénicos utilizando PCR con transcriptasa inversa (RT-PCR) y exploración de alto rendimiento.

En resumen, se extrajeron bazo de cinco ratones heterocigotos que comprendían dos segmentos kappa V (Vk1-39 y Vk3-20) cada uno de ellos conteniendo tres sustituciones de histidina (ratones cuyo locus kappa se representa en la **Fig. 35**) y cadenas pesadas de ratón endógenas y se homogeneizaron en 1xPBS (Gibco) utilizando portaobjetos

de vidrio. Las células se sedimentaron en una centrífuga (500xg durante 5 minutos) y los eritrocitos se lisaron en tampón de lisis ACK (Gibco) durante 3 minutos. Las células se lavaron con 1xPBS y se filtraron utilizando un filtrador celular de 0,7 µm. Los linfocitos B se aislaron de las células de bazo utilizando selección positiva magnética MACS para CD19 (Miltenyi Biotec). Utilizando el kit RNeasy Plus (Qiagen) se extrajo el ARN total de los linfocitos B sedimentados. El ARNm PolyA+ se aisló del ARN total utilizando el mini kit de ARNm Oligotex Direct (Qiagen).

A partir de ARNm de linfocitos B esplénicos se preparó ADNc bicatenario mediante 5' RACE utilizando el Kit de Síntesis de ADNc SMARTer Pico (Clontech). La transcriptasa inversa de Clontech y los dNTP se sustituyeron con Superscript II y los dNTP de Invitrogen. Los repertorios de cadena ligera de inmunoglobulina se amplificaron del ADNc utilizando cebadores específicos para la región constante de K $\kappa$  y el cebador 5' RACE SMARTer (Tabla 14). Los productos de la PCR se lavaron utilizando el Kit de Purificación de PCR QIAquick (Qiagen). Se realizó una segunda ronda de PCR utilizando el mismo cebador 5' RACE y un cebador 3' anidado específico para la región constante de IgK (Tabla 15). Los productos de la segunda ronda de PCR se purificaron utilizando un sistema E-gel SizeSelect (Invitrogen). Se realizó una tercera PCR con cebadores que añadían 454 aptámeros y códigos de barra. Los productos de la tercera ronda de PCR se purificaron utilizando Perlas Agencourt AMPure XP. Los productos de PCR purificados se cuantificaron mediante SYBR-qPCR utilizando un Kit de Cuantificación de Biblioteca KAPA (KAPA Biosystems). Las bibliotecas agrupadas se sometieron a PCR con emulsión (emPCR) utilizando el Kit 454 GS Junior Titanium Series Lib-A emPCR (Roche Diagnostics) y a secuenciación bidireccional utilizando el instrumento Roche 454 GS Junior de acuerdo con los protocolos del fabricante.

**Tabla 14: cebador de PCR de la primera ronda**

NOMBRE	SECUENCIA (SEQ ID NO)
3' mlgk externo	AAGAAGCACACGACTGAGGCAC (SEQ ID NO:129)

**Tabla 15: cebador de PCR de la segunda ronda**

NOMBRE	SECUENCIA (SEQ ID NO)
3' mlgk interno	GGAAGATGGATACAGTTGGTGC (SEQ ID NO:130)

Para el análisis bioinformático, las lecturas de secuencia 454 se separaron basándose en la total coincidencia del código de barras de la muestra y se acortaron con fines de calidad. Las secuencias se anotaron basándose en el alineamiento de las secuencias de Ig reordenadas con respecto a la base de datos de los segmentos V y J de la línea germinal humana utilizando la instalación de igblast (NCBI, v2.2.25+). Una secuencia se marcó como ambigua y se eliminó de los análisis cuando se detectaron varios mejores resultados con idéntica puntuación. Se desarrolló un conjunto de programas en perl para analizar los resultados y los datos se guardaron en la base de datos mysql. La región CDR3 de la cadena ligera kappa se definió entre el motivo FGXG y el codón C conservado.

La **Fig. 38** representa alineaciones de la secuencia de aminoácidos codificada por la secuencia IGKV3-20 (**Fig. 38A**) o la secuencia IGKV1-39 (**Fig. 38B**) de línea germinal humana, con traducciones de aminoácidos de secuencias V $\kappa$  ejemplares obtenidas de anticuerpos reordenados productivamente generados en ratones que comprendían una DLC-5J modificada con histidina (que comprende un locus variable de cadena ligera que comprende los segmentos génicos V $\kappa$ 1-39 y V $\kappa$ 3-20, cada segmento con tres modificaciones de histidina como se ha descrito anteriormente). Las lecturas de secuencia mostraron que la mayor parte de las cadenas ligeras reordenadas productivamente conservaban al menos una histidina introducida en su CDR3 de la línea germinal. En algunos casos, en la mayor parte de todas las cadenas ligeras humanas reordenadas productivamente que comprendían la secuencia V $\kappa$ 3-20 que conservaba al menos un resto de histidina, las tres modificaciones de histidina introducidas en su CDR3 de la línea germinal se conservaban (véase la **Fig. 38A**). En algunos casos, en las cadenas ligeras humanas reordenadas productivamente que comprenden la secuencia V $\kappa$ 1-39 que conserva al menos un resto de histidina, aproximadamente el 50 % de las cadenas ligeras conservan las tres histidinas introducidas en su CDR3 de la línea germinal (véase la **Fig. 38B** alineamiento superior), mientras que aproximadamente el 50 % de las cadenas ligeras conservan dos de las tres histidinas introducidas en su CDR3 de la línea germinal (véase la **Fig. 38B**, alineamiento inferior). En algunos casos, las histidinas en la última posición de la secuencia del segmento V pueden perderse debido a una reordenación V-J.

**Ejemplo 5. Generación y análisis de ratones que comprenden una sola secuencia de nucleótidos de cadena pesada de inmunoglobulina humana reordenada y dos segmentos génicos kappa V**

Se generaron ratones que comprendían una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena pesada (MAID6031; "ratón UHC") como se ha descrito anteriormente. En resumen, en el ratón UHC, todos los segmentos génicos variables de cadena pesada funcional endógena se delecionaron y se remplazaron con una sola secuencia de ácido nucleico de región variable de cadena pesada reordenada que codifica hV $\mu$ 3-23/D/J $\mu$ 4, que está unida operativamente a una secuencia de ácido nucleico de región constante de cadena pesada endógena.

Se generaron ratones que comprendían locus de cadena ligera modificados por ingeniería genética que contenían dos segmentos génicos Vk humanos (por ejemplo, un segmento génico Vk1-39 humano y otro Vk3-20 humano) y uno cualquiera de un segmento Jk humano (Jk5; DLC-1J) o cinco segmentos génicos Jk humanos (hJk1-5; DLC-5J) como se ha descrito anteriormente. En resumen, un locus de cadena ligera modificado por ingeniería genética contenía dos segmentos génicos Vk humanos y cinco segmentos génicos Jk humanos (Jk1-5) en configuración no reordenada y estaban unidos operativamente a una secuencia de región constante κ de ratón endógena (MAID 1911 (DLC-5J); **Fig. 19E**). El otro locus de cadena ligera modificado por ingeniería genética contenía dos segmentos génicos Vk humanos y un segmento génico Jk humano (Jk1) en configuración no reordenada y estaban unidos operativamente a una secuencia de región constante κ de ratón endógena (MAID 1913 (DLC-1J); **Fig. 21D**). Para cada uno de los dos locus de cadena ligera adicionales modificados por ingeniería genética, los segmentos génicos humanos estaban flanqueados en 3' con secuencias señal de recombinación para permitir la reordenación *in vivo* de los segmentos génicos humanos en los linfocitos B.

Los ratones homocigotos UHC (MAID6031) descritos anteriormente se cruzaron con ratones homocigotos DLC-5J (MAID1911) para producir un ratón heterocigoto para el alelo UHC y el alelo DLC-5J. De manera similar, ratones homocigotos UHC (MAID6031) se cruzaron con ratones UHC DLC-1J (MAID1913) para generar un ratón heterocigoto para el alelo UHC y el alelo OLC-1J. A partir de estos cruces, se generaron ratones F1 heterocigotos que se cruzaron entre sí para obtener ratones homocigotos para cada alelo. La presencia de los alelos modificados genéticamente en los locus de cadena ligera y cadena pesada de inmunoglobulina se confirmó mediante exploración TAQMAN™ y cariotipado utilizando las sondas y cebadores específicos descritos anteriormente.

Los ratones heterocigotos para el alelo UHC y para el alelo DLC-5J se cruzaron entre sí para generar ratones homocigotos (MAID 1912HO 6032HO; "DLC x UHC") que expresaban cadenas "ligeras" de inmunoglobulina principalmente a partir del locus modificado genéticamente. Los ratones MAID 1912HO 6032HO (homocigotos DLC x UHC) comprenden una inserción de la cadena pesada universal descrita en el presente documento (por ejemplo, HV<sub>H</sub>3-23/hD/hJ<sub>H</sub>4) en el locus de cadena pesada de ratón en el que todos los genes VDJ de cadena pesada variable endógenos están delecionados y DLC-5J (*hVlc1-39 hVlc3-20 hJlc1-5*) el locus de cadena ligera kappa (κ) de ratón en los que todos los genes Vk y Jk de ratón se han delecionado.

Todos los ratones se instalaron en jaulas y se cruzaron en condiciones específicas sin gérmenes patógenos en Regeneron Pharmaceuticals. Tres ratones F5 VELOCIMMUNE® (MAID 12930 1640HO ("VI3"), véase la Patente de Estados Unidos n.º 8.502.018) (de 14 semanas de vida, macho; Fondo: 26,5 % C57/BL6, 22,75 % 129 y 50,75 % Balb/c) y tres ratones MAID 1912HO 6032HO F2 (**Fig. 39**; de 7-8 semanas de vida, hembra; fondo: 18,75 C57/BL6, 18,75 % 129 y 62,5 % Balb/c) se sacrificaron y se extrajeron los bazo y la médula ósea de los animales. La médula ósea se extrajo de los fémures lavando abundantemente con medio RPMI completo (medio RPMI complementado con suero bovino fetal, piruvato sódico, Hepes, 2-mercaptoetanol, aminoácidos no esenciales y gentamicina). Las preparaciones de eritrocitos de bazo y médula ósea se lisaron con tampón de lisis ACK y se lavaron con medio RPMI completo.

#### 40 Citometría de flujo

Para examinar la capacidad de los ratones homocigotos "DLC x UHC" (MAID 1912HO 6032HO) modificados genéticamente para producir anticuerpos procedentes de los alelos modificados genéticamente (por ejemplo, del alelo que contiene una sola copia de V<sub>H</sub>3-23/D/J<sub>H</sub>4 reordenada en el locus de cadena pesada y del alelo que contiene dos genes Jk humanos y cinco Vk humanos en el locus de cadena ligera), se realizó análisis de separación de células activadas por fluorescencia (FACS) como se indica en el Ejemplo 3.

En el compartimento esplénico, los ratones MAID 1912HO 6032HO demostraron números de linfocitos B CD19+ y números de linfocitos B maduros que eran sustancialmente iguales a los números observados en ratones VELOCIMMUNE® (VI3) (**Figs. 40A-40B**), que sirven como un control de los efectos específicos observados en ratones MAID 1912HO 6032HO con respecto a ratones con otras modificaciones genéticas en sus locus de inmunoglobulina; además, el sistema inmunitario humoral de los ratones VELOCIMMUNE® funciona de manera similar al de los ratones de tipo silvestre (*supra*). Los ratones MAID 1912HO 6032HO demostraron un aumento de 2 veces en el número de linfocitos B inmaduros en el bazo en comparación con ratones VI3 (**Figs. 40A-40B**). Los ratones MAID 1912HO 6032HO también eran sustancialmente similares a los ratones VI3 con respecto al uso de la cadena ligera kappa y gamma (**Figs. 41A-41B**). Los ratones MAID 1912HO 6032HO (DLC x UHC) también demostraron aumento de IgM en la superficie en linfocitos B esplénicos (es decir, más expresión en superficie de IgM por célula) en comparación con ratones VI3 (**Fig. 42**).

Además, los ratones MAID 1912HO 6032HO (DLC x UHC) demostraron desarrollo alterado de linfocitos B periféricos según progresaban los linfocitos B a través de los diversos estadios en el compartimento esplénico (por ejemplo, inmaduros, maduros, T1, T2 T3, precursores de zona marginal, zona marginal, folicular-I, folicular-II, etc.) que se producía de una manera diferente a la observada en ratones VI3 (**Fig. 43A**). Específicamente, los ratones MAID 1912HO 6032HO (DLC x UHC) demostraron más linfocitos B inmaduros, T1 y de zona marginal (ZM) en el compartimento esplénico en comparación con los ratones VI3. El número de linfocitos B foliculares-I y foliculares-II

en los ratones MAID 1912HO 6032HO (DLC x UHC) era sustancialmente igual al observado en ratones VI3 (**Fig. 43B**).

- 5 En el compartimento de la médula ósea, los ratones MAID 1912HO 6032HO (DLC x UHC) demostraron números similares de linfocitos B CD19+ en comparación con los controles de ratón VI3 (**Figs. 44A-44B**). Sin embargo, los ratones MAID 1912HO 6032HO (DLC x UHC) demostraron aproximadamente 25 veces menos linfocitos pro B en la médula ósea en comparación con los ratones VI3 (**Figs. 45A-45B**). Los ratones MAID 1912HO 6032HO (DLC x UHC) también demostraron aproximadamente 2 veces menos linfocitos B inmaduros y 2 veces menos linfocitos B maduros en la médula ósea en comparación con los ratones VI3 (**Figs. 46A-46B**). Asimismo, los ratones MAID 10 1912HO 6032HO (DLC x UHC) demostraron una preferencia (un aumento en 2 veces) para expresión de lambda en comparación con los ratones VI3 (**Fig. 47**).

### Estudios de inmunización

- 15 Cinco ratones TS (fondo 75 % C57BL6/25 % 129) y siete F2 MAID1912HO 6031 HET (homocigotos DLC x heterocigotos UHC) se inmunizaron en la almohadilla plantar con 0,025 ml de una mezcla que contenía 2,35 µg de un antígeno X, 10 µg de oligonucleótido CpG (ODN) 1826, InvivoGen, cat n.º tlr1-1826) y 25 µg de adyuvante de gel fosfato de aluminio (Brenntag cat n.º 7784-30-7). Los ratones recibieron un refuerzo seis veces con la misma dosificación. Los días 0, 15 y 23 posteriores a la primera inmunización, se extrajo sangre de ratones anestesiados 20 utilizando un sangrado retroorbital en tubos separadores con suero BD (BD, cat. n.º 365956), y el suero se recogió siguiendo las instrucciones del fabricante. Se realizó una segunda ronda de inmunización como se ha indicado anteriormente cinco semanas después de la primera ronda de inmunización.

- 25 Para medir los niveles de anticuerpos de IgG específicos de antígeno, se realizó un ELISA como en el Ejemplo 3. Como se muestra en la **Fig. 48**, los ratones modificados genéticamente, que son heterocigotos con respecto al alelo diana que contiene la secuencia de nucleótidos de V<sub>H</sub>3-23/D/J<sub>H</sub>4 reordenada y homocigotos con respecto al alelo diana que contiene DLC-5J, pudieron producir anticuerpos de IgG específicos de antígeno a niveles comparables a los producidos por los ratones de tipo silvestre tanto a los 23 días como a las 5 semanas después de la primera inmunización. Los ratones MAID1912HO 6031 HET (homocigotos DLC x heterocigotos UHC) también pudieron 30 producir anticuerpos de IgG específicos de antígeno a niveles comparables a los producidos por los ratones de tipo silvestre después de la segunda ronda de inmunización.

- Por tanto, estos ratones producen anticuerpos que comprenden una cadena ligera quimérica inversa (dominio variable de cadena ligera humana y C<sub>k</sub> de ratón) procedente de una reordenación de uno de los dos segmentos 35 génicos V<sub>L</sub> humanos (segmentos génicos V<sub>k</sub>1-39 o V<sub>k</sub>3-20) y segmentos J<sub>k</sub> humanos y una cadena pesada quimérica inversa (dominio variable de cadena pesada humana y C<sub>H</sub> de ratón) procedentes de un solo segmento génico variable de cadena pesada humana reordenada. Los anticuerpos quiméricos inversos (es decir anticuerpos que comprenden estas cadenas quiméricas inversas) se obtienen después de la inmunización con un antígeno de interés.

- 40 **Ejemplo 6. Generación y análisis de ratones que comprenden una sola secuencia de nucleótidos de cadena pesada de inmunoglobulina humana reordenada y dos segmentos génicos kappa V que contienen tres sustituciones de histidina**

- 45 De manera similar, ratones portadores de un locus de cadena ligera humana modificado genéticamente que comprenden una doble cadena ligera modificada con histidina (por ejemplo, ratones que comprenden dos segmentos génicos V<sub>L</sub> humanos con modificaciones de histidina descritos anteriormente en el presente documento) se cruzan con ratones que contienen un reemplazo del locus variable de cadena pesada de ratón endógena con locus de cadena pesada humana universal (locus que comprende un solo dominio variable de cadena pesada humana reordenada como se describe anteriormente en el presente documento). Por tanto, estos ratones producen anticuerpos que comprenden una cadena ligera quimérica inversa (dominio variable de cadena ligera humana y C<sub>k</sub> de ratón) procedente de una reordenación de uno de los dos segmentos génicos V<sub>L</sub> humanos modificados con histidina (segmentos génicos V<sub>k</sub>1-39 o V<sub>k</sub>3-20) y segmentos génicos J<sub>k</sub> humanos y una cadena pesada quimérica inversa (dominio variable de cadena pesada humana y C<sub>H</sub> de ratón) procedente de un solo dominio variable de 55 cadena pesada humana reordenada. Los anticuerpos quiméricos inversos se obtienen después de la inmunización con un antígeno de interés. Los anticuerpos humanos dependientes de pH generados en dichos ratones, se identifican utilizando métodos de exploración y de aislamiento de anticuerpos conocidos en la técnica o descritos anteriormente.

- 60 Las secuencias de nucleótidos de región variable de cadena pesada y ligera de los linfocitos B que expresan los anticuerpos se identifican, y las cadenas pesada y ligera completamente humanas se crean fusionando secuencias de nucleótidos de región variable de cadena pesada y ligera con secuencias de nucleótidos C<sub>L</sub> y C<sub>H</sub> humanas, respectivamente. Las cadenas ligeras de interés, por ejemplo, cadenas ligeras que se unen al antígeno de interés (por ejemplo, cadenas ligeras de anticuerpos que también demuestran propiedades antigénicas dependientes de pH 65 utilizando una variedad de ensayos conocidos en la técnica, por ejemplo, el ensayo BIACORE™) se coexpresan en un sistema de expresión adecuado con cadenas pesadas procedentes de otros anticuerpos, por ejemplo, cadenas

pesadas procedentes de anticuerpos que comprenden cadenas ligeras procedentes del mismo segmento génico  $V_L$  como el de la cadena ligera de interés (por ejemplo,  $V_{k1-39}$  o  $V_{k3-20}$ ) y la capacidad del anticuerpo reconstituido para conservar la unión a antígeno y las propiedades de unión a antígeno dependientes de pH, se somete a ensayo.

#### 5 Ejemplo 7. Construcción de ratones que comprenden un locus de cadena ligera de inmunoglobulina que contiene una secuencia VDJ de cadena pesada reordenada

Se generaron ratones que comprendían una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera kappa (MAID6079; "ratón UHC en kappa") mediante métodos similares a los descritos anteriormente para el direccionamiento del locus de cadena pesada. En resumen, en el ratón UHC en kappa, todos los segmentos génicos  $V_k$  y  $J_k$  variables kappa de cadena ligera funcional endógena se deleccionaron y se reemplazaron con una sola secuencia de ácido nucleico de región variable de cadena pesada reordenada que codificaba la secuencia de  $hV_{H3-23/D/J_{H4}}$ , que está unida operativamente a una secuencia de ácido nucleico de región constante de cadena ligera endógena. La construcción diana final para la creación de un locus genómico que contenía una secuencia de dominio variable de cadena pesada humana reordenada que contenía, de 5' a 3', (1) un brazo de homología 5' con aproximadamente 22500 pb de una secuencia genómica de ratón cadena arriba del locus de la cadena ligera de Ig endógena; (2) un sitio FRT 5'; (3) un casete de neomicina; (4) un sitio FRT 3', (5) un promotor  $hV_{H3-23}$  de 2239 pb (SEQ ID NO: 139); (6) una secuencia de nucleótidos de cadena pesada de inmunoglobulina humana reordenada ( $hV_{H3-23/D/J_{H4}}$ ; SEQ ID NO: 136); (7) un intrón  $hJ_{H4}$  (SEQ ID NO: 140); y (8) un brazo de homología 3' con aproximadamente 75000 pb de una secuencia genómica de ratón cadena abajo de los segmentos génicos  $J_L$  de ratón. Los ratones heterocigotos portadores de la modificación se cruzaron entre sí para generar homocigotos (MAID 6079HO) que eran capaces de crear cadenas "ligeras" de inmunoglobulina únicamente a partir del locus modificado genéticamente. Los ratones MAID 6079HO (homocigotos UHC en kappa) comprenden una inserción en la cadena pesada universal descrita en el presente documento (por ejemplo,  $hV_{H3-23/hD/hJ_{H4}}$ ) en el locus de cadena ligera kappa ( $\kappa$ ) de ratón en el que todos los genes  $V_k$  y  $J_k$  de ratón se han deleccionado.

Todos los ratones se instalaron en jaulas y se cruzaron en condiciones específicas sin gérmenes patógenos en Regeneron Pharmaceuticals. Cuatro ratones MAID 6079HO F1 (Fig. 49; 7-12,5 semanas de vida, machos y hembras) y cuatro ratones de control de la misma camada de tipo silvestre MAID 6079 F1 (7-12,5 semanas de vida, machos y hembras) se sacrificaron, y se extrajeron los bazo y la médula ósea de los animales. La médula ósea se extrajo de los fémures lavando abundantemente con medio RPMI completo (medio RPMI complementado con suero bovino fetal, piruvato sódico, Hepes, 2-mercaptoetanol, aminoácidos no esenciales y gentamicina). Las preparaciones de eritrocitos de bazo y médula ósea se lisaron con tampón de lisis ACK y se lavaron con medio RPMI completo.

#### 35 Citometría de flujo

Para examinar la capacidad de los ratones homocigotos "ratón UHC en kappa" (MAID 6079HO) modificados genéticamente descritos en el presente documento para producir anticuerpos procedentes del alelo modificado genéticamente (por ejemplo, del alelo que contiene una sola copia de la secuencia  $V_{H3-23/D/J_{H4}}$  reordenada en un locus de cadena ligera kappa), se realizó un análisis de separación de células activadas por fluorescencia (FACS) como se indica en el Ejemplo 3.

Los ratones MAID 6079HO demostraron números de pro y pre linfocitos B en el compartimento de la médula ósea que eran sustancialmente iguales a los observados en las crías de la camada de tipo silvestre (Figs. 50A-50B). En cambio, demostraron números más bajos de linfocitos B inmaduros y maduros en el compartimento de la médula ósea en comparación con las crías de la camada de tipo silvestre (Figs. 51A-51C). De hecho, los ratones tenían 2 veces menos linfocitos B inmaduros y casi 4 veces menos linfocitos B maduros. Los ratones MAID 6079HO utilizaron casi exclusivamente secuencias de cadena ligera lambda en los linfocitos B inmaduros y maduros en la médula ósea (Fig. 52).

En el compartimento esplénico, los ratones MAID 6079HO demostraron menos linfocitos B maduros en comparación con las crías de la camada de tipo silvestre (Figs. 53A-53B). Similar a lo observado en la médula ósea, los ratones MAID 6079HO utilizaron casi exclusivamente secuencias de cadena ligera lambda en el compartimento esplénico (Figs. 54A-54B). También demostraron menos linfocitos inmaduros, un aumento de linfocitos B en la zona marginal y una disminución de linfocitos B foliculares en comparación con las crías de la camada de tipo silvestre (Fig. 55).

#### 60 Ejemplo 8. Generación y análisis de ratones que comprenden un locus de cadena ligera de inmunoglobulina que contiene una secuencia VDJ de cadena pesada reordenada y un locus de cadena pesada de inmunoglobulina que contiene una secuencia de dominio variable de cadena ligera humana

Se generaron ratones homocigotos para una secuencia de ácido nucleico de región variable de cadena pesada reordenada en el locus de cadena ligera (MAID 6079HO, homocigotos "ratones UHC en kappa") como se ha descrito anteriormente. Estos ratones se cruzaron con ratones homocigotos (MAID 1994HO) para una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada (ratón con cadena kappa en cadena pesada ("KoH")). Los ratones KoH homocigotos MAID 1994 comprenden 40 genes  $V_k$  humanos y todos los

genes Jk humanos, con IGCR larga y ADAM6 de ratón, insertados en un locus de cadena constante de cadena pesada de Ig de ratón (es decir, un locus de cadena pesada de Ig de ratón deleciónado)). Los ratones KoH se han descrito anteriormente; véase, por ejemplo, la publicación previa a la concesión de patente de Estados Unidos 2012/0096572.

5 Todos los ratones se instalaron en jaulas y se cruzaron en condiciones específicas sin gérmenes patógenos en Regeneron Pharmaceuticals. Dos ratones VELOCIMMUNE® (MAID 1242HO 1640HO ("VI3"), véase la Patente de Estados Unidos n.º 8.502.018) (15 semanas de vida, hembras n=2; Fondo: 28 % C57/BL6, 13 % 129 y 59 % Balb/c), cuatro ratones MAID 1994HO 6079HO F2 (**Fig. 56**; 13-14 semanas de vida, machos n=2; Fondo: 25 % C57/BL6, 25 % 129, y 50 % Balb/c), y ratones de control de la crías de camada de tipo silvestre MAID 6079, se sacrificaron, y se extrajeron los bazo y la médula ósea de los animales. La médula ósea se extrajo de los fémures lavando abundantemente con medio RPMI completo (medio RPMI complementado con suero bovino fetal, piruvato sódico, Hepes, 2-mercaptoetanol, aminoácidos no esenciales y gentamicina). Las preparaciones de eritrocitos de bazo y médula ósea se lisaron con tampón de lisis ACK y se lavaron con medio RPMI completo.

### 15 Citometría de flujo

Para examinar la capacidad de los ratones homocigotos "KoH x UHC en kappa" (MAID 1994HO 6079HO) modificados genéticamente descritos en el presente documento para producir anticuerpos procedentes de los alelos modificados genéticamente (por ejemplo, del alelo que contiene una sola copia de la secuencia de V<sub>H</sub>3-23/D/J<sub>H</sub>4 reordenada y del alelo que contiene una secuencia de ácido nucleico de región variable de cadena ligera kappa en un locus de cadena pesada), se realizó análisis de separación de células activadas por fluorescencia (FACS) como se indica en el Ejemplo 3.

25 Los ratones MAID 1994HO 6079HO demostraron frecuencias más bajas de linfocitos pre B y CD19+ en el compartimento de la médula ósea en comparación con los ratones VI3 (**Fig. 57A**). Específicamente, los ratones MAID 1994HO 6079HO demostraron números aproximadamente 2 veces más bajos de linfocitos CD19+ y pre B en la médula ósea en comparación con los ratones VI3 (**Fig. 57B**). Adicionalmente, los ratones MAID 1994HO 6079HO demostraron aproximadamente 3 veces menos linfocitos B inmaduros en el compartimento de la médula ósea en comparación con los ratones VI3 (**Figs. 58A y 58B**). También se encontró que los linfocitos B de los ratones MAID 1994HO 6079HO carecían esencialmente de expresión de cadena ligera lambda en la médula ósea (**Fig. 59**).

Los ratones MAID1994HO 6079HO demostraron una frecuencia más baja de linfocitos B en el compartimento esplénico. Específicamente, los ratones MAID 1994HO 6079HO tenían menos números de linfocitos B esplénicos (aproximadamente 2 veces menos) y linfocitos B maduros (aproximadamente 3 veces menos) con respecto a los ratones VI3 (**Figs. 60A-60B**). De nuevo, estos demostraron carecer de expresión de cadena ligera lambda en comparación con los ratones VI3 (**Fig. 61**).

Teniendo en cuenta el desarrollo de linfocitos B periféricos en el bazo, el análisis FACS indicó que, en comparación con los ratones VI3, los ratones MAID1994HO 6079HO tenían una mayor frecuencia de linfocitos en fase T1 en el bazo (**Fig. 62**).

### LISTADO DE SECUENCIAS

45 <110> Regeneron Pharmaceuticals, Inc.

<120> ANIMALES NO HUMANOS CON SECUENCIAS DE CADENA PESADA DE INMUNOGLOBULINA MODIFICADAS

50 <130> 2010794-0456

<150> US 61/879.338

<151> 18-09-2013

55 <150> US 61/766.765

<151> 20-02-2013

<170> PatentIn versión 3.5

60 <210> 1

<211> 95

<212> PRT

<213> Artificial

65 <220>

<223> Sintética

ES 2 678 221 T3

<400> 1

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr  
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile  
35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly  
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr Pro  
85 90 95

5

<210> 2  
<211> 21  
<212> ADN  
<213> Artificial

10

<220>  
<223> Sintética

15

<220>  
<221> CDS  
<222> (1)..(21)

<400> 2

cag cag agc tac agc acc ccc  
Gln Gln Ser Tyr Ser Thr Pro  
1 5

20

21

<210> 3  
<211> 7  
<212> PRT  
<213> Artificial

25

<220>  
<223> Construcción sintética

30

<400> 3

Gln Gln Ser Tyr Ser Thr Pro  
1 5

35

<210>4  
<211>21  
<212> ADN  
<213> Artificial

40

<220>

<223> Sintética  
 <220>

5 <221> CDS  
 <222> (1)..(21)

<400>4

cac cat agc cac agc acc cac  
 His His Ser His Ser Thr His  
 1 5

10 <210> 5  
 <211>7  
 <212> PRT  
 <213> Artificial

15 <220>  
 <223> Construcción sintética

20 <400> 5

His His Ser His Ser Thr His  
 1 5

25 <210> 6  
 <211> 21  
 <212> ADN  
 <213> Artificial

30 <220>  
 <223> Sintética

<220>  
 <221> CDS  
 <222> (1)..(21)

35 <400> 6

cac cag agc tac agc acc ccc  
 His Gln Ser Tyr Ser Thr Pro  
 1 5

40 <210> 7  
 <211> 7  
 <212> PRT  
 <213> Artificial

45 <220>  
 <223> Construcción sintética

<400> 7

His Gln Ser Tyr Ser Thr Pro  
 1 5

50 <210> 8  
 <211> 21  
 <212> ADN  
 <213> Artificial

55 <220>  
 <223> Sintética

<220>



ES 2 678 221 T3

<400> 12

cag cag agc tac agc acc cac  
 Gln Gln Ser Tyr Ser Thr His  
 1 5

21

5 <210> 13  
 <211> 7  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Construcción sintética

<400> 13

Gln Gln Ser Tyr Ser Thr His  
 1 5

15

<210> 14  
 <211>21  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Sintética

25 <220>  
 <221> CDS  
 <222> (1)..(21)

<400> 14

30 cac cat agc tac agc acc ccc  
 His His Ser Tyr Ser Thr Pro  
 1 5

21

35 <210> 15  
 <211> 7  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Construcción sintética

40 <400> 15

His His Ser Tyr Ser Thr Pro  
 1 5

45 <210> 16  
 <211> 21  
 <212> ADN  
 <213> Artificial

50 <220>  
 <223> Sintética

<220>  
 <221> CDS  
 <222> (1)..(21)

55 <400> 16

cac cag agc cac agc acc ccc  
 His Gln Ser His Ser Thr Pro  
 1 5

5 <210> 17  
 <211> 7  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Construcción sintética  
 <400> 17

His Gln Ser His Ser Thr Pro  
 1 5

15 <210> 18  
 <211> 21  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Sintética

25 <220>  
 <221> CDS  
 <222> (1)..(21)  
 <400> 18

cac cag agc tac agc acc cac  
 His Gln Ser Tyr Ser Thr His  
 1 5

30 <210> 19  
 <211> 7  
 <212> PRT  
 <213> Artificial

35 <220>  
 <223> Construcción sintética  
 <400> 19

40 His Gln Ser Tyr Ser Thr His  
 1 5

45 <210> 20  
 <211> 21  
 <212> ADN  
 <213> Artificial

50 <220>  
 <223> Sintética

55 <220>  
 <221> CDS  
 <222> (1)..(21)  
 <400> 20

ES 2 678 221 T3

	cag cat agc cac agc acc ccc		21
	Gln His Ser His Ser Thr Pro		
	1	5	
5	<210> 21 <211> 7 <212> PRT <213> Artificial		
10	<220> <223> Construcción sintética <400> 21		
		Gln His Ser His Ser Thr Pro	
		1	5
15	<210> 22 <211> 21 <212> ADN <213> Artificial		
20	<220> <223> Sintética		
25	<220> <221> CDS <222> (1)..(21) <400> 22		
	cag cat agc tac agc acc cac		21
	Gln His Ser Tyr Ser Thr His		
	1	5	
30	<210> 23 <211> 7 <212> PRT <213> Artificial		
35	<220> <223> Construcción sintética		
40	<400> 23		
		Gln His Ser Tyr Ser Thr His	
		1	5
45	<210> 24 <211> 21 <212> ADN <213> Artificial		
50	<220> <223> Sintética		
55	<220> <221> CDS <222> (1)..(21) <400> 24		

cag cag agc cac agc acc cac  
 Gln Gln Ser His Ser Thr His  
 1 5

5 <210> 25  
 <211> 7  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Construcción sintética  
 <400> 25

Gln Gln Ser His Ser Thr His  
 1 5

15 <210> 26  
 <211> 21  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Sintética

25 <220>  
 <221> CDS  
 <222> (1)..(21)  
 <400> 26

cac cat agc cac agc acc ccc  
 His His Ser His Ser Thr Pro  
 1 5

30 <210> 27  
 <211> 7  
 <212> PRT  
 <213> Artificial

35 <220>  
 <223> Construcción sintética  
 <400> 27

40 His His Ser His Ser Thr Pro  
 1 5

45 <210> 28  
 <211> 21  
 <212> ADN  
 <213> Artificial

50 <220>  
 <223> Sintética

55 <220>  
 <221> CDS  
 <222> (1)..(21)  
 <400> 28

cac cat agc tac agc acc cac 21  
 His His Ser Tyr Ser Thr His  
 1 5

<210> 29  
 <211> 7  
 5 <212> PRT  
 <213> Artificial

<220>  
 <223> Construcción sintética  
 10 <400> 29

His His Ser Tyr Ser Thr His  
 1 5

15 <210> 30  
 <211> 21  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Sintética

<220>  
 <221> CDS  
 25 <222> (1)..(21)

<400> 30

cac cag agc cac agc acc cac 21  
 His Gln Ser His Ser Thr His  
 1 5

30 <210>31  
 <211> 7  
 <212> PRT  
 <213> Artificial

35 <220>  
 <223> Construcción sintética

<400> 31  
 40

His Gln Ser His Ser Thr His  
 1 5

<210> 32  
 <211> 21  
 45 <212> ADN  
 <213> Artificial

<220>  
 <223> Sintética  
 50

<220>  
 <221> CDS  
 <222> (1)..(21)

55 <400> 32

cag cat agc cac agc acc cac 21  
 Gln His Ser His Ser Thr His  
 1 5

ES 2 678 221 T3

<210> 33  
 <211> 7  
 <212> PRT  
 5 <213> Artificial  
  
 <220>  
 <223> Construcción sintética  
  
 10 <400> 33  
  

Gln His Ser His Ser Thr His  
1 5

  
 <210> 34  
 15 <211> 45  
 <212> ADN  
 <213> Artificial  
  
 <220>  
 20 <223> Sintética  
  
 <400> 34  
 cttactactg tcaacatagt cacagtacc atccgatcac cttcg 45  
  
 25 <210> 35  
 <211> 50  
 <212> ADN  
 <213> Artificial  
  
 30 <220>  
 <223> Sintética  
  
 <400> 35  
 caactacta ctgtcacat agtcacagta cccatccgat caccttcggc 50  
 35  
 <210> 36  
 <211> 45  
 <212> ADN  
 <213> Artificial  
 40  
 <220>  
 <223> Sintética  
  
 <400> 36  
 45 cgaaggtgat cggatgggta ctgtgactat gttgacagta gtaag 45  
  
 <210> 37  
 <211> 50  
 <212> ADN  
 50 <213> Artificial  
  
 <220>  
 <223> Sintética  
  
 <400> 37  
 55 gccgaaggtg atccgatggg tactgtgact atggtgacag tagtaagttg 50  
  
 <210> 38  
 <211> 23  
 <212> ADN  
 60 <213> Artificial  
  
 <220>

<223> Sintética  
 <400> 38  
 5 tgggcacaac agacaatcgg ctg 23  
 <210> 39  
 <211> 19  
 <212> ADN  
 <213> Artificial  
 10  
 <220>  
 <223> Sintética  
 <400> 39  
 15 ggtggagagg ctattcggc 19  
 <210> 40  
 <211> 17  
 <212> ADN  
 <213> Artificial  
 20  
 <220>  
 <223> Sintética  
 <400> 40  
 25 gaacacggcg gcatcag 17  
 <210>41  
 <211> 29  
 <212> ADN  
 <213> Artificial  
 30  
 <220>  
 <223> Sintética  
 <400> 41  
 35 ccattatgat gtcocatgcc tctctgtc 29  
 <210> 42  
 <211> 28  
 <212> ADN  
 <213> Artificial  
 40  
 <220>  
 <223> Sintética  
 <400> 42  
 45 aggtgagggt acagataagt gttatgag 28  
 <210> 43  
 <211> 27  
 <212> ADN  
 <213> Artificial  
 50  
 <220>  
 <223> Sintética  
 <400> 43  
 55 tgacaaatgc cctaattata gtgatca 27  
 <210> 44  
 <211> 26  
 <212> ADN  
 <213> Artificial  
 60

<220>  
 <223> Sintética

5 <400> 44  
 atcagcagaa accagggaaa gccct 26

<210> 45  
 <211> 21  
 <212> ADN  
 <213> Artificial

10

<220>  
 <223> Sintética

15 <400> 45  
 ggcaagtca gagcattagc a 21

<210> 46  
 <211> 21  
 <212> ADN  
 <213> Artificial

20

<220>  
 <223> Sintética

25 <400> 46  
 tgcaactgg atgcagcata g 21

<210> 47  
 <211> 19  
 <212> ADN  
 <213> Artificial

30

<220>  
 <223> Sintética

35 <400> 47  
 ggccacattc catgggttc 19

40 <210> 48  
 <211> 20  
 <212> ADN  
 <213> Artificial

45

<220>  
 <223> Sintética

50 <400> 48  
 gcaaacaaaa accactggcc 20

<210> 49  
 <211> 37  
 <212> ADN  
 <213> Artificial

55

<220>  
 <223> Sintética

60 <400> 49  
 ctgttcctct aaaactggac tccacagtaa atggaaa 37

<210> 50  
 <211> 22

<212> ADN  
 <213> Artificial  
  
 <220>  
 5 <223> Sintética  
  
 <400> 50  
 gggcactgga tacgatgat gg 22  
  
 10 <210> 51  
 <211> 19  
 <212> ADN  
 <213> Artificial  
  
 15 <220>  
 <223> Sintética  
  
 <400> 51  
 cacagcttgt gcagcctcc 19  
 20  
 <210> 52  
 <211> 37  
 <212> ADN  
 <213> Artificial  
 25  
 <220>  
 <223> Sintética  
  
 <400> 52  
 30 agaagaagcc tgtactacag catccgtttt acagtca 37  
  
 <210> 53  
 <211> 19  
 <212> ADN  
 35 <213> Artificial  
  
 <220>  
 <223> Sintética  
  
 <400> 53  
 40 accatagtca cagtaccca 19  
  
 <210> 54  
 <211> 23  
 <212> ADN  
 45 <213> Artificial  
  
 <220>  
 <223> Sintética  
 50  
 <400> 54  
 agcagtctgc aacctgaaga ttt 23  
  
 <210> 55  
 <211> 18  
 <212> ADN  
 55 <213> Artificial  
  
 <220>  
 60 <223> Sintética  
  
 <400> 55  
 cccttgccg aagtgat 18

# ES 2 678 221 T3

5  
<210> 56  
<211> 19  
<212> ADN  
<213> Artificial  
  
<220>  
<223> Sintética  
  
10 <400> 56  
atagtcacag tacccatcc 19  
  
15 <210> 57  
<211> 23  
<212> ADN  
<213> Artificial  
  
<220>  
<223> Sintética  
  
20 <400> 57  
agtctgcaac ctgaagattt tgc 23  
  
25 <210> 58  
<211> 18  
<212> ADN  
<213> Artificial  
  
<220>  
<223> Sintética  
  
30 <400> 58  
cccttgccg aaggtgat 18  
  
35 <210> 59  
<211> 108  
<212> PRT  
<213> Artificial  
  
40 <220>  
<223> Sintética  
  
<400> 59

ES 2 678 221 T3

Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly  
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser  
 20 25 30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu  
 35 40 45

Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Ile Pro Asp Arg Phe Ser  
 50 55 60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu  
 65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser Pro  
 85 90 95

Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys  
 100 105

5 <210> 60  
 <211> 54  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Sintética

<400> 60  
 gatttgcag tgtattactg tcagcagat gtagctcac ctggacgtt cggc 54

15 <210> 61  
 <211> 54  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Sintética

<400> 61  
 gatttgcag tgtattactg tcaccat ggtcactcac ctggacgtt cggc 54

25 <210> 62  
 <211> 54  
 <212> ADN  
 <213> Artificial

30 <220>  
 <223> Sintética

<400> 62  
 gccgaacgtc caagtgagt gaccatggg atgacagtaa tacactgcaa aatc 54

35 <210> 63  
 <211> 47  
 <212> ADN

<213> Artificial  
 <220>  
 <223> Sintética  
 5  
 <400> 63  
 gcagtgatt actgtcatca ctatggcac tcacctgga cgttcgg 47  
 <210> 64  
 <211> 47  
 <212> ADN  
 <213> Artificial  
 10  
 <220>  
 <223> Sintética  
 15  
 <400> 64  
 ccgaacgtcc aagtgagtg accatagtga tgacagtaat aactgc 47  
 <210> 65  
 <211> 24  
 <212> ADN  
 <213> Artificial  
 20  
 <220>  
 <223> Sintética  
 25  
 <400> 65  
 aaagagccac cctctctgc aggg 24  
 <210> 66  
 <211> 19  
 <212> ADN  
 <213> Artificial  
 30  
 <220>  
 <223> Sintética  
 35  
 <400> 66  
 tccaggcacc ctgtcttg 19  
 <210> 67  
 <211> 26  
 <212> ADN  
 <213> Artificial  
 40  
 <220>  
 <223> Sintética  
 45  
 <400> 67  
 aagtagctgc tgctaact ctgact 26  
 <210> 68  
 <211> 15  
 <212> ADN  
 <213> Artificial  
 50  
 <220>  
 <223> Sintética  
 55  
 <400> 68  
 ctgtcatcac catgg 15  
 <210> 69

<211> 23  
 <212> ADN  
 <213> Artificial  
 5 <220>  
 <223> Sintética  
 <400> 69  
 gcagactgga gcctgaagat ttt 23  
 10 <210> 70  
 <211> 20  
 <212> ADN  
 <213> Artificial  
 15 <220>  
 <223> Sintética  
 <400> 70  
 20 ccgaacgtcc aaggtgagtg 20  
 <210>71  
 <211> 17  
 <212> ADN  
 25 <213> Artificial  
 <220>  
 <223> Sintética  
 30 <400> 71  
 tactgtcatc actatgg 17  
 <210> 72  
 <211> 22  
 <212> ADN  
 35 <213> Artificial  
 <220>  
 <223> Sintética  
 40 <400> 72  
 gcagactgga gcctgaagat tt 22  
 <210> 73  
 <211> 20  
 <212> ADN  
 45 <213> Artificial  
 <220>  
 <223> Sintética  
 50 <400> 73  
 ccgaacgtcc aaggtgagtg 20  
 <210> 74  
 <211>21  
 <212> ADN  
 <213> Artificial  
 60 <220>  
 <223> Sintética  
 <220>  
 <221> CDS  
 65 <222> (1)..(21)

ES 2 678 221 T3

<400> 74

cag cag tat ggt agc tca cct  
 Gln Gln Tyr Gly Ser Ser Pro  
 1 5

5

<210> 75  
 <211> 7  
 <212> PRT  
 <213> Artificial

10

<220>  
 <223> Construcción sintética

<400> 75

15

Gln Gln Tyr Gly Ser Ser Pro  
 1 5

<210> 76  
 <211> 21  
 <212> ADN  
 <213> Artificial

20

<220>  
 <223> Sintética

25

<220>  
 <221> CDS  
 <222> (1)..(21)

30

<400> 76

cat cac cat ggt cac tca cct  
 His His His Gly His Ser Pro

1 5

35

<210> 77  
 <211> 7  
 <212> PRT  
 <213> Artificial

40

<220>  
 <223> Construcción sintética

<400> 77

His His His Gly His Ser Pro  
 1 5

45

<210> 78  
 <211> 21  
 <212> ADN  
 <213> Artificial

50

<220>  
 <223> Sintética

55

<220>  
 <221> CDS  
 <222> (1)..(21)

ES 2 678 221 T3

<400> 78

cat cac tat ggt cac tca cct  
His His Tyr Gly His Ser Pro  
1 5

21

5 <210> 79  
<211> 7  
<212> PRT  
<213> Artificial

10 <220>  
<223> Construcción sintética

<400> 79

15 His His Tyr Gly His Ser Pro  
1 5

<210> 80  
<211> 108  
<212> PRT  
<213> Artificial

<220>  
<223> Sintética

25 <400> 80

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
1 5 10 15  
Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr  
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile  
35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly  
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr Pro Pro  
85 90 95

Ile Thr Phe Gly Gln Gly Thr Arg Leu Glu Ile Lys  
100 105

30 <210> 81  
<211> 50  
<212> ADN  
<213> Artificial

35 <220>  
<223> Sintética

<400>81

ES 2 678 221 T3

caacttacta ctgtcaacag agttacagta ccctccgat caccttcggc 50

5 <210> 82  
 <211> 9652  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Sintética

<400> 82

gcaccacttc gtcgcagcgc aggctttggt ctcccttgtc cgtgcggcgc acgcccaccg	60
agtttacgca ccagcacacc gaggtctggt tggtagcga gttcctattc cgaagttcct	120
attctctaga aagtatagga acttctcgcg cgtctggcct ccgaggcctc cgcgccgggt	180
tttggcgcct cccgcgggcg cccccctcct cacggcgagc gctgccacgt cagacgaagg	240
gcgcagcgag cgtcctgatc ctccgccccg gacgctcagg acagcggccc gctgctcata	300
agactcggcc ttagaacccc agtatcagca gaaggacatt ttaggacggg acttgggtga	360
ctctagggca ctggttttct ttccagagag cggaacaggc gaggaaaagt agtcccttct	420
cggcgattct gcggagggat ctccgtgggg cggatgaacgc cgatgattat ataaggacgc	480
gccgggtgtg gcacagctag ttccgtcgca gccgggattt gggtcgcggg tcttgtttgt	540

ES 2 678 221 T3

ggatcgctgt gatcgctcact tggtagtag cgggctgctg ggctggccgg ggctttcgtg 600  
 gccgccgggc cgctcgggtg gacggaagcg tgtggagaga ccgccaaggg ctgtagtctg 660  
 ggtccgcgag caaggttgcc ctgaactggg ggttgggggg agcgcagcaa aatggcggct 720  
 gttcccagat cttgaatgga agacgcttgt gaggcgggct gtgaggctgt tgaaacaagg 780  
 tggggggcat ggtgggcggc aagaacccaa ggtcttgagg ccttcgctaa tgcgggaaag 840  
 ctcttattcg ggtgagatgg gctggggcac catctgggga ccctgacgtg aagtttgtca 900  
 ctgactggag aactcggttt gtcgtctgtt gcggggggcg cagttatggc ggtgccgttg 960  
 ggcaagtgcac ccgtacctt gggagcgcgc gccctcgtcg tgcgtgacg tcaccogttc 1020  
 tgttggctta taatgcaggg tggggccacc tgccggtagg tgtgcggtag gcttttctcc 1080  
 gtcgcaggac gcagggttcg ggcctagggt aggctctcct gaatcgacag gcgccggacc 1140  
 tctggtgagg ggagggataa gtgaggcgtc agtttctttg gtcggtttta tgtacctatc 1200  
 ttcttaagta gctgaagctc cggttttgaa ctatgcgctc ggggttggcg agtgtgtttt 1260  
 gtgaagtttt ttaggcacct tttgaaatgt aatcatttgg gtcaatatgt aattttcagt 1320  
 gttagactag taaattgtcc gctaaattct ggccgttttt ggcttttttg ttagacgtcg 1380  
 agctctagat tgggaaccog ggtctctcga attgttgaca attaatacgc ggcatagtat 1440  
 atcggcatag tataatacga caaggtgagg aactaaacca tccaccatga ttgaacaaga 1500  
 tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc 1560  
 acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc 1620  
 ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg acgaggcagc 1680  
 gcggctatcg tggctggcca cgacgggctt tccttgcgca gctgtgctcg acgttgtcac 1740  
 tgaagcggga aggactggc tgctattggg cgaagtgccg gggcaggatc tcctgtcatc 1800  
 tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc ggctgcatac 1860  
 gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg agcagacacg 1920  
 tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc atcaggggct 1980  
 cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt 2040  
 cgtgaccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg 2100  
 attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggctac 2160  
 ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttctcog tgctttacgg 2220  
 tatcgccgct cccgattcgc agcgcacgc cttctatcgc cttcttgacg agttcttctg 2280  
 aagatccgct gtaagtctgc agaaattgat gatctattaa acaataaaga tgtccactaa 2340  
 aatggaagtt tttcctgtca tactttgtta agaagggtga gaacagagta cctacatttt 2400

ES 2 678 221 T3

gaatggaagg attggagcta cgggggtggg ggtggggtgg gattagataa atgcctgctc 2460  
 tttactgaag gctctttact attgctttat gataatgttt catagttgga tatkataatt 2520  
 taaacaagca aaaccaaatt aagggccagc tcattcctcc cactcatgat ctatagatct 2580  
 atagatctct cgtgggatca ttgtttttct cttgattccc actttgtggg tctaagtact 2640  
 gtggtttcca aatgtgtcag tttcatagcc tgaagaacga gatcagcagc ctctgttcca 2700  
 catacacttc attctcagta ttgttttgcc aagttctaata tccatcagac ctcgacctgc 2760  
 agcccctaga gaagttccta ttccgaagtt cctattctct agaaagtata ggaacttcct 2820  
 agggtttcac cggtggcgcg ccgagagagg aaagagagca gcgataccga aatgtcctc 2880  
 agcgagaagc taccacagag gatgaatgga gatcaagccc acgtggaaac atgggaaaat 2940  
 gtctcagtat ttttccacct aagaagggag ggagatgggg tatgtataca cctccctgtc 3000  
 ctactgatt gagggctttc cgagaggatg ctcatccag gtgctgtgat aggccatgtg 3060  
 tacaggcagg gctgccttct ccagcttcag atagagcagt gaggaaaaga tatggccatg 3120  
 gggggtcagc agaagtacag caaagggaaa agggaaaggg tagcaagagt gacaactata 3180  
 ttcaccccc ccacacacac acacacacac acgaaattgt gtattgcaat ccagaactgc 3240  
 ttctctctga acctaaatct tagcaagcag tttaccagta actgcccttg aaattcaggc 3300  
 ccctggaaag gagcaggggg ttgtgtacag gctataccac agcagtctgc ccacccttag 3360  
 tgatgcatga gtaatgctcc ctggactccc caggttctag tcttctcatg tcgatgtagt 3420  
 tgattccact tcccttgctg cacaaccagg ctgggatgcc tgggcagagg cagacatgtg 3480  
 aggtataggg gttcaaatct gtttccaagt tttatccagc ttcaaagcat ttctccgtgt 3540  
 acatgagcgg tggcttgaca ggagatggag actctctttc ctggatgtga ggcaaggagg 3600  
 caggcgtctg agtcaggatg atgtccctac tactgctaa agagaaaagt ggctttgatg 3660  
 gtgcagggca gggaaatgca ctgagtggtc gccaccctca cagaagagaa agtgttcact 3720  
 gacctggcct tccccaggg cctctccctc ccattgcttt ccagaaagcc atgatttttg 3780  
 agagccacac ctgaacactc acaaacatta tgggtgggaa agcagatcag agcattaggc 3840  
 aagttgcatt accttggcct tcttcttttg gagacaattg atgtgggggt ctagattgac 3900  
 ccagagtttc aagtttatcc tgattcaggc ttcaacagct ggaggaagaa acagagatgt 3960  
 tttttgaagt aacacagatct agcattacta atcaaccctt catactgatg acctatggga 4020  
 aataataccc aagggcagaa aatgggcag aataagggga gccccaaacc aagacgaagc 4080  
 tgctgcccat tgagaccctg ggtattacag agacctatag ctctggataa tggaagatct 4140  
 atgagtggca caggcgtga ggaatcacag catcattatc gtgcatctgc agggaattgc 4200  
 ttgtaaatat actggtaatt acaaatgttt aaggtcacta caaatacttt ggagtgtatt 4260  
 aaatatgctt ctgataaaga ctgtttttct cacatgaaac aatgggaacc atgtgacaat 4320

ES 2 678 221 T3

cacagaggtg ttgttactat agcaaaaggg attgttactc tccacatccc tttaagtaac 4380  
ttgaaggcct gatagaccca ccctctaaga cttcattaga cattccctac gaatggttat 4440  
actctcctgt atactcccaa tacaactcta aaatatatta ttccatatag tccttagggtt 4500  
tgtattaaag tttgactttt ttccttcaaa atatctcttg tcacaacagc ggctctagag 4560  
agaaatacat tccctccagg caaatctatg ctgcgctggt ctgacctggg accctgggga 4620  
cattgccct gtgctgagtt actaagatga gccagccctg cagctgtgct cagcctgccc 4680  
catgccctgc tgattgattt gcatgttcca gagcacagcc ccctgccctg aagacttttt 4740  
tatgggctgg tcgcaccctg tgcaggagtc agtctcagtc aggacacagc atggacatga 4800  
gggtccccgc tcagctcctg gggctcctgc tactctggct ccgaggtaag gatggagaac 4860  
actaggaatt tactcagcca gtgtgctcag tactgactgg aacttcaggg aagttctctg 4920  
ataacatgat taatagtaag aatatttggt tttatgtttc caatctcagg tgccagatgt 4980  
gacatccaga tgaccagtc tccatcctcc ctgtctgcat ctgtaggaga cagagtcacc 5040  
atcacttgcc gggcaagtca gagcattagc agctatttaa attggtatca gcagaaacca 5100  
gggaaagccc ctaagctcct gatctatgct gcatccagtt tgcaaagtgg ggtcccatca 5160  
aggttcagtg gcagtgatc tgggacagat ttcactctca ccatcagcag tctgcaacct 5220  
gaagattttg caacttacta ctgtcaacag agttacagta cccctccac agtgttacia 5280  
gtcataacat aaacctccaa ggaagcagat gtgtgaggac gagccacccc agatgctcct 5340  
cctggtgcct ccatctgctg agagcatttc tcaaactcag tcaggttttg aaagtcattg 5400  
ggagactttt gtagagggga ccagggaggc tcctctgaac tctaagcctc ttttgccccg 5460  
gatccgatca ggaaaatagc aaccgagtga ctggactatt aaagcatccc ctctgtggag 5520  
aaaaggacta tggcaacagc ttgccaccta taagagacia atatgtcact cagctggatg 5580  
ctgggtctac catgaccacc cttgagatga acttatgcc a tgtttttgga tgcccaactga 5640  
gactttgctt tgaccaagga aagttcttta ctgccc aaat gacgtcaatg gacacactct 5700  
catgggacat gatggatttt ccatacacct tgttatccaa aggccaatgg atctatttaa 5760  
tgctagaaca gctgactcac acggcaactg aagagtgtac atcagggcaa cctgctagta 5820  
gagtgttacc ccatctgact acagcaattg gacattaaac actgcaactgc agagcaaggg 5880  
aaagacggca cggaggtgca ggttgaaaaa cactgagttg agtgaggaag aagttgggca 5940  
aggccagttt cctgatagag ctgccactgt gaaatcccaa actcagtgtt tccaatcatt 6000  
tttttccttt tatagtgatg ttcttggggg tgttgcagtt taggccacca tgataaccca 6060  
gacaggcctc ctaactctaa ttggatggaa tgattcccct gggtccttc tatgggttcc 6120  
acaggatcaa gaaatatcag ggtgttaatt tccttctgct ggtaggaggg tcatctgatt 6180

ES 2 678 221 T3

cctccattag ggtggatgac gttatcagac atgtcaatth tctacagtac ttgacctgcc 6240  
 ttctagaact ggacaaccga ggggtgaaagg actggatcaa gcaacaatgg caatgggtgc 6300  
 ccaagaggt agtagtctca gaacagggag agacagactg agtccctcca ctaactcagc 6360  
 ccaacacctg tggatgagta ggggatgcct gagatcctgg ggtggatggg aggtggggca 6420  
 ctgatctgtc aatctgcttt ttcttcaagg atcaggcagc agagaccag aagcttcatg 6480  
 tctttgtaag gctcttccaa gccaacacag ataatttgac aaaacactgt atctgcatcc 6540  
 cagacctcac cctgtcacag actgaagtgg ttgtttcatc ctactaaggg taaactatac 6600  
 cagctataca gaataaaaag actggatagt ttagaggatc acccaagaaa tagttctttg 6660  
 cctgcatgga caaaaccatc ttctgtcttt agggaaatgg tatcactacc ctgaggattt 6720  
 ggagcccag tctcacttat ctgtgcagtt gtgaaagtcc tcacaccac agtgctgagg 6780  
 ttaattgaat gttactcttt taatttctgc aaagaatgag acagcttctg gacctcagg 6840  
 aaagatcact aacaagtaaa tacaagtata tccggaagat aaagttgtaa tagactcttc 6900  
 ctttcaacct gatccatcat gcatttaggg agctgactgg gcacaagttg gagcagaaaag 6960  
 agaaaaatga aaccacagcc ttctatthtg tttctaacag acttgtacca aacattctgt 7020  
 ggctcaatct aggtgatggg gagacaagag gacacagggg ttaaattctg tggccgcagg 7080  
 ggagaagttc taccctcaga ctgagccaac ggctthttct ggctgatca cctgggcatg 7140  
 ggctgctgag agcagaaaagg ggaggcagat tgtctctgca gctgcaagcc cagcaccgc 7200  
 cccagctgct ttgcatgtcc ctcccagccg ccctgcagtc cagagccat atcaatgcct 7260  
 gggtcagagc tctggagaag agctgctcag ttaggacca gagggaaacca tggaaacccc 7320  
 agcgcagctt ctcttctcc tgctactctg gctcccaggt gaggggaaca tgggatggtt 7380  
 ttgcatgtca gtgaaaacc tctcaagtcc tgttacctgg caactctgct cagtcaatac 7440  
 aataattaa gctcaatata aagcaataat tctggctctt ctgggaagac aatgggtttg 7500  
 atttagatta catgggtgac ttttctgtht tatttccaat ctgagatacc accggagaaa 7560  
 ttgtgttgac gcagtctcca ggcaccctgt ctttgtctcc aggggaaaga gccaccctct 7620  
 cctgcagggc cagtcaagat gttagcagca gctacttagc ctggtaccag cagaaacctg 7680  
 gccaggctcc caggctctc atctatggg catccagcag ggccactggc atcccagaca 7740  
 ggttcagtgg cagtgggtct gggacagact tcaactctcac catcagcaga ctggagcctg 7800  
 aagatthtg agtgtattac tgtcagcagt atggtagctc acctcccaca gtgattcagc 7860  
 ttgaaacaaa aacctctgca agaccttcat tgtttactag attataccag ctgcttctt 7920  
 tacagatagc tgctgcaatg acaactcaat ttagcatct cttctctgct tgggcatttt 7980  
 ggggatctta aaaaagtaat cccttgatat atthttgact ctgattcctg cattthtct 8040  
 cagaccaaga tggacagcca ggtttaagca cagthtcaca gtaatggcca ctggatcaga 8100

ES 2 678 221 T3

ttacatcag tggatgtcag taaaggtccc aaccagagcc ataaggcaac aacaatagca 8160  
 acaataatc aaaattggaa aagaagaatt aaagctgtca taattcactg atgaaggatt 8220  
 gtgtgcagat aaaattcaaa tttgtctaca gagaaactac taaaattgac atgagaaata 8280  
 gaaaatcatt agattcaaga tcaatattatt aattcataga ttcaaaaatc aatttcattt 8340  
 ctgcataata aaaaatgcta aaaattaaca ttataaaaca aacacaccat ttacaaaaac 8400  
 atcaaagtat caattattta aaaaaaatag actaaatata ctgacgtctc cagaatatta 8460  
 ttttgaaaaa taaaagaaaa cctaagtaaa tagaaattca gttcaaagac tgaatgtctc 8520  
 agtactataa aaatgtcaat tcttcaaaga ttaaaatatt gattacatat aagaaaaatc 8580  
 aaaatcctaa agtatactcc aatttaaatt aagaagctaa tctaaatatt atatgggaat 8640  
 gtcaaggatg tagaatagcc acagtgaacc tgaagaaaca ccaaatgag aacttcagat 8700  
 gcctgaatac ctggaatata gtgtgggtgg cagtatgggtg atggtgagat cagaagtta 8760  
 aaaatttgca aacgtgctta tttttgaaa taatcactac gcagatgtag ccaaaccctc 8820  
 ttcaactgtg gcgtacgctc aggtcaattc caaagagtac cagattcttt caaaaagtca 8880  
 gatgagtaag ggatagaaaa ttagttcatc ttaaggaaca gccaagcgct agccagttaa 8940  
 gtgaggcatc tcaattgcaa gattttctct gcatcgggtca ggtagtgat attaacagcg 9000  
 aaaagagatt tttgttaagg ggaaagtaat taagttaaca ctgtggatca cttcggcca 9060  
 agggacacga ctggagatta aacgtaagta atttttcact attgtcttct gaaatttggg 9120  
 tctgatggcc agtattgact tttagaggct taaataggag tttggtaaag attggtaaat 9180  
 gagggcattt aagatttgcc atgggttgca aaagttaaac tcagcttcaa aatggattt 9240  
 ggagaaaaaa agattaaatt gctctaaact gaatgacaca aagtaaaaaa aaaaagtgta 9300  
 actaaaaagg aacccttgta tttctaagga gcaaaaagtaa atttattttt gttcactctt 9360  
 gccaaatatt gtattgggtg ttgctgatta tgcatgatac agaaaagtgg aaaaatacat 9420  
 tttttagtct ttctcccttt tgtttgataa attattttgt cagacaaca taaaatcaa 9480  
 tagcacgccc taagatctag atgcatgctc gagtgccatt tcattacctc tttctccgca 9540  
 cccgacatag ataagcttat cgataccgct gacctcgagg gggggcccgg taccctgtgg 9600  
 catacagtgt cagattttct gtttatcaag ctagtgagat taggggcaaa aa 9652

<210> 83  
 <211> 10867  
 <212> ADN  
 <213> Artificial

5

<220>  
 <223> Sintética

10

<400> 83

ES 2 678 221 T3

gcaccacttc gtcgcagcgc aggctttggt ctcccttgtc cgtgcgggcg acgccaccg 60  
 agtttacgca ccagcacacc gaggtctggt tggtagcga gttcctatc cgaagttcct 120  
 attctctaga aagtatagga acttctcgcg cgtctggcct ccgaggcctc cgcgccgggt 180  
 tttggcgcct cccgcggggg cccccctcct cacggcgagc gctgccacgt cagacgaagg 240  
 gcgcagcgag cgtcctgatc cttccgcccg gacgctcagg acagcggccc gctgctcata 300  
 agactcggcc ttagaacccc agtatcagca gaaggacatt ttaggacggg acttgggtga 360  
 ctctagggca ctggttttct ttccagagag cggaacaggc gaggaaaagt agtcccttct 420  
 cggcgattct gcggagggat ctccgtgggg cggtagaacgc cgatgattat ataaggacgc 480  
 gccgggtgtg gcacagctag ttccgctcga gccgggattt gggtagcggg tcttgtttgt 540  
 ggatcgcgtg gatcgtcaact tggtagtag cgggctgctg ggctggccgg ggctttcgtg 600  
 gccgccgggc cgctcgggtg gacggaagcg tgtggagaga ccgccaaggg ctgtagtctg 660  
 ggtccgcgag caaggttgcc ctgaactggg ggttgggggg agcgcagcaa aatggcggct 720  
 gttcccagat cttgaatgga agacgcttgt gaggcgggct gtgaggcgt tgaacaagg 780  
 tggggggcat ggtgggcggc aagaacccaa ggtcttgagg ccttcgctaa tgcgggaaag 840  
 ctcttattcg ggtgagatgg gctggggcac catctgggga ccctgacgtg aagtttgtca 900  
 ctgactggag aactcggttt gtcgtctgtt gcggggcg cagttatggc ggtgccgttg 960  
 ggcagtgcac ccgtaccttt gggagcgcgc gccctcgtcg tgtcgtgacg tcaccgcttc 1020  
 tgttggtta taatgcaggg tggggccacc tgccgtagg tgtgcggtag gcttttctcc 1080  
 gtcgcaggac gcagggttcg ggcctagggt aggctctcct gaatcgacag gcgccggacc 1140  
 tctggtgagg ggagggataa gtgaggcgtc agtttctttg gtcggtttta tgtacctatc 1200  
 ttcttaagta gctgaagctc cggttttgaa ctatgcgtc ggggttggcg agtgtgtttt 1260  
 gtgaagtttt ttaggcacct tttgaaatgt aatcatttgg gtcaatatgt aattttcagt 1320  
 gttagactag taaattgtcc gctaaattct ggccgttttt ggcttttttg ttagacgtcg 1380  
 agctctagat tgggaacccg ggtctctcga attgttgaca attaatcctc ggcatagtat 1440  
 atcggcatag tataatacga caaggtgagg aactaaacca tccaccatga ttgaacaaga 1500  
 tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct atgactgggc 1560  
 acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc 1620  
 ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg acgaggcagc 1680  
 gcggctatcg tggctggcca cgacgggctg tccttgcgca gctgtgctcg acgttgtcac 1740  
 tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc tcctgtcatc 1800  
 tcacctgtct cctgccgaga aagtatccat catggctgat gcaatgcggc ggctgcatac 1860  
 gcttgatccg gctacctgcc cattcgacca ccaagcгаа catcgcatcg agcagacagc 1920

ES 2 678 221 T3

tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc atcaggggct 1980  
 cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt 2040  
 cgtgaccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg 2100  
 attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggctac 2160  
 ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg 2220  
 tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg agttcttctg 2280  
 aagatccgct gtaagtctgc agaaattgat gatctattaa acaataaaga tgtccactaa 2340  
 aatggaagtt tttcctgtca tactttgtta agaagggtga gaacagagta cctacatttt 2400  
 gaatggaagg attggagcta cgggggtggg ggtgggggtgg gattagataa atgcctgctc 2460  
 tttactgaag gctctttact attgctttat gataatgttt catagttgga tatcataatt 2520  
 taacaagca aaaccaaatt aagggccagc tcattcctcc cactcatgat ctatagatct 2580  
 atagatctct cgtgggatca ttgtttttct ctgattccc actttgtggt tctaagtact 2640  
 gtggtttcca aatgtgtcag tttcatagcc tgaagaacga gatcagcagc ctctgttcca 2700  
 catacacttc attctcagta ttgttttgcc aagttctaata tccatcagac ctcgacctgc 2760  
 agcccctaga gaagttccta ttccgaagtt cctattctct agaaagtata ggaacttcct 2820  
 agggtttcac cgggtggcgc cagagagagg aaagagagca gcgataccga aatgtcctc 2880  
 agcgagaagc taccacagag gatgaatgga gatcaagccc acgtggaaac atgggaaaat 2940  
 gtctcagtat tttccacct aagaaggag ggagatgggg tatgtataca cctccctgtc 3000  
 ctactgatt gagggctttc cgagaggatg ctattccag gtgctgtgat aggccatgtg 3060  
 tacaggcagg gctgccttct ccagcttcag atagagcagt gaggaaaaga tatggccatg 3120  
 gggggtcagc agaagtacag caaagggaaa agggaaagg tagcaagagt gacaactata 3180  
 ttcaccccc ccacacacac acacacacac acgaaattgt gtattgcaat ccagaactgc 3240  
 ttctctctga acctaaatct tagcaagcag tttaccagta actgcccttg aaattcaggc 3300  
 ccctggaag gagcaggggg ttgtgtacag gctataccac agcagtctgc ccacccttag 3360  
 tgatgcatga gtaatgctcc ctggactccc caggttctag tcttctcatg tcgatgtagt 3420  
 tgattccact tcccttgctg cacaaccagg ctgggatgcc tgggcagagg cagacatgtg 3480  
 aggtataggg gttcaaatct gtttccaagt tttatccagc ttcaaagcat ttctccgtgt 3540  
 acatgagcgg tggcttgaca ggagatggag actctctttc ctggatgtga ggcaaggagg 3600  
 caggcgtctg agtcaggatg atgtccctac tctactgctaa agagaaaagt ggctttgatg 3660  
 gtgcagggca gggaaatgca ctgagtggc gccaccctca cagaagagaa agtgttcact 3720  
 gacctggcct ttcccaggg cctctccctc ccattgcttt ccagaaagcc atgatttttg 3780

ES 2 678 221 T3

agagccacac	ctgaacactc	acaaacatta	tggtgggaaa	agcagatcag	agcattaggc	3840
aagttgcatt	accttggcct	tcttcctttg	gagacaattg	atgtgggggtt	ctagattgac	3900
ccagagtttc	aagtttatcc	tgattcaggc	ttcaacagct	ggaggaagaa	acagagatgt	3960
tttttgaagt	aaacagatct	agcattacta	atcaaccctt	catactgatg	acctatggga	4020
aataataccc	aagggcagaa	aaatgggcag	aataagggga	gccccaaacc	aagacgaagc	4080
tgctgcccat	tgagaccctg	ggtattacag	agacctatag	ctctggataa	tggaagatct	4140
atgagtggca	caggcgctga	ggaatcacag	catcattatc	gtgcatctgc	aggggaattgc	4200
ttgtaaatat	actggttaatt	acaaatgttt	aaggtcacta	caaatacttt	ggagtgtatt	4260
aaatatgctt	ctgataaaga	ctgtttttct	cacatgaaac	aatgggaacc	atgtgacaat	4320
cacagaggtg	ttgttactat	agcaaaaggg	attgttactc	tccacatccc	tttaagtaac	4380
ttgaaggcct	gatagaccca	ccctctaaga	cttcattaga	cattccctac	gaatggttat	4440
actctcctgt	atactcccaa	tacaactcta	aaatatatta	ttccatatag	tccttaggtt	4500
tgtattaaag	tttgactttt	ttccttcaaa	atatctcttg	tcacaacagc	ggctctagag	4560
agaaatacat	tccctccagg	caaatctatg	ctgcgctggt	ctgacctggg	accctgggga	4620
cattgcccct	gtgctgagtt	actaagatga	gccagccctg	cagctgtgct	cagcctgcc	4680
catgccctgc	tgattgattt	gcatgttcca	gagcacagcc	ccctgcctg	aagacttttt	4740
tatgggctg	tcgcaccctg	tgaggagtc	agtctcagtc	aggacacagc	atggacatga	4800
gggtccccgc	tcagctcctg	gggctcctgc	tactctggct	ccgaggtaag	gatggagaac	4860
actaggaatt	tactcagcca	gtgtgctcag	tactgactgg	aacttcaggg	aagttctctg	4920
ataacatgat	taatagtaag	aatatgtgtt	tttatgtttc	caatctcagg	tgccagatgt	4980
gacatccaga	tgaccagtc	tccatcctcc	ctgtctgcat	ctgtaggaga	cagagtcacc	5040
atcacttgcc	gggcaagtca	gagcattagc	agctatttaa	attggtatca	gcagaaacca	5100
gggaaagccc	ctaagctcct	gatctatgct	gcatccagtt	tgcaaagtgg	ggtcccacatca	5160
aggttcagtg	gcagtggatc	tgggacagat	ttcactctca	ccatcagcag	tctgcaacct	5220
gaagattttg	caacttacta	ctgtcaacag	agttacagta	cccctccac	agtgttacia	5280
gtcataacat	aaacctccaa	ggaagcagat	gtgtgaggac	gagccacccc	agatgctcct	5340
cctggtgcct	ccatctgctg	agagcatttc	tcaaactcag	tcaggttttg	aaagtcattg	5400
ggagactttt	gtagagggga	ccagggaggc	tcctctgaac	tctaagcctc	ttttgccccg	5460
gatccgatca	ggaaaatagc	aaccgagtga	ctggactatt	aaagcatccc	ctctgtggag	5520
aaaaggacta	tggaacagc	ttgccaccta	taagagacia	atatgtcact	cagctggatg	5580
ctgggtctac	catgaccacc	cttgagatga	acttatgcca	tgtttttggga	tgcccactga	5640
gactttgctt	tgaccaagga	aagttcttta	ctgcccacaa	gacgtcaatg	gacacactct	5700

ES 2 678 221 T3

catgggacat gatggatfff ccatacacct tgttatccaa aggccaatgg atctatttaa 5760  
 tgctagaaca gctgactcac acggcaactg aagagtgtac atcagggcaa cctgctagta 5820  
 gagtggtagc ccatctgact acagcaattg gacattaaac actgcactgc agagcaaggg 5880  
 aaagacggca cggaggtgca ggttgaaaaa cactgagttg agtgaggaag aagttgggca 5940  
 aggccagttt cctgatagag ctgccactgt gaaatcccaa actcagtgtt tccaatcatt 6000  
 tttttccttt tatagtgatg ttcttggggg tgttgtagtt taggccacca tgataacca 6060  
 gacaggcctc ctaactctaa ttggatggaa tgattcccct ggggtgccttc tatgggttcc 6120  
 acaggatcaa gaaatatcag ggtgttaatt tccttctgct ggtaggaggg tcatctgatt 6180  
 cctccattag ggtggatgac gttatcagac atgtcaattt tctacagtac ttgacctgcc 6240  
 ttctagaact ggacaaccga gggtgaaagg actggatcaa gcaacaatgg caatgggtgc 6300  
 ccaaagaggt agtagtctca gaacaggag agacagactg agtccctcca ctaactcagc 6360  
 ccaacacctg tggatgagta ggggatgcct gagatcctgg ggtggatggg aggtggggca 6420  
 ctgatctgtc aatctgcttt ttcttcaagg atcaggcagc agagaccag aagcttcatg 6480  
 tctttgtaag gctcttcaa gccaacacag ataatttgac aaaacactgt atctgcatcc 6540  
 cagacctcac cctgtcacag actgaagtgg ttgtttcatc ctactaaggg taaactatac 6600  
 cagctataca gaataaaaag actggatagt ttagaggatc acccaagaaa tagttctttg 6660  
 cctgcatgga caaaaccatc ttctgtcttt agggaaatgg taccactacc ctgaggattt 6720  
 ggagcccagg tctcacttat ctgtgcagtt gtgaaagtcc tcacaccac agtgctgagg 6780  
 ttaattgaat gttactcttt taatttctgc aaagaatgag acagcttctg gaccctcagg 6840  
 aaagatcact aacaagtaaa tacaagtata tccggaagat aaagttgtaa tagactcttc 6900  
 ctttcaacct gatccatcat gcatttaggg agctgactgg gcacaagttg gagcagaaag 6960  
 agaaaaatga aaccacagcc ttctatfff tttctaacag acttgtacca aacattctgt 7020  
 ggctcaatct aggtgatggg gagacaagag gacacagggg ttaaattctg tggccgcagg 7080  
 ggagaagttc taccctcaga ctgagccaac ggccttttct ggctgatca cctgggcatg 7140  
 ggctgctgag agcagaaagg ggaggcagat tgtctctgca gctgcaagcc cagcaccgc 7200  
 cccagctgct ttgcatgtcc ctcccagccg cctgtagtc cagagccat atcaatgcct 7260  
 gggctcagagc tctggagaag agctgctcag ttaggacca gagggacca tggaaacccc 7320  
 agcgcagctt ctcttctcc tgctactctg gctcccaggt gaggggaaca tgggatggtt 7380  
 ttgcatgtca gtgaaaaccc tctcaagtcc tgttacctgg caactctgct cagtcaatac 7440  
 aataattaa gctcaatata aagcaataat tctggctctt ctgggaagac aatgggtttg 7500  
 attagatta catgggtgac ttttctgfff tatttccaat ctcagatacc accggagaaa 7560

ES 2 678 221 T3

ttgtggtgac	gcagtctcca	ggcaccctgt	ctttgtctcc	aggggaaaga	gccaccctct	7620
cctgcagggc	cagtcagagt	gttagcagca	gctacttagc	ctggtaccag	cagaaacctg	7680
gccaggctcc	caggctcctc	atctatggtg	catccagcag	ggccactggc	atcccagaca	7740
ggttcagtg	cagtgggtct	gggacagact	tcactctcac	catcagcaga	ctggagcctg	7800
aagatthttgc	agtgtattac	tgtcagcagt	atggtagctc	acctcccaca	gtgattcagc	7860
ttgaaacaaa	aacctctgca	agaccttcat	tgtttactag	attataccag	ctgcttcctt	7920
tacagatagc	tgctgcaatg	acaactcaat	tttagcatct	cttctctgct	tgggcatttt	7980
ggggatctta	aaaaagtaat	cccttgatat	atthtttgact	ctgattcctg	cattthttcct	8040
cagaccaaga	tggacagcca	ggthtaagca	cagthtcaca	gtaatggcca	ctggatcaga	8100
thttacatcag	tggatgtcag	taaaggthccc	aaccagagcc	ataaggcaac	aacaatagca	8160
acaaataatc	aaaattggaa	aagaagaatt	aaagctgtca	taattcactg	atgaaggatt	8220
gtgtgcagat	aaaattcaaa	thttgtctaca	gagaaactac	taaaattgac	atgagaaata	8280
gaaaatcatt	agattcaaga	tcaattthatt	aattcataga	thcaaaaatc	aattthcattt	8340
ctgcataata	aaaaatgcta	aaaattaaca	thataaaaaca	aacacaccat	thacaaaaac	8400
atcaaagtat	caattatthta	aaaaaaatag	actaaataca	ctgacgtctc	cagaatatta	8460
thtttgaaaa	taaaagaaaa	cctaagthaaa	tagaaattca	gthcaaaagac	tgaatgtctc	8520
agtactataa	aaatgtcaat	thttthcaaaga	thaaaatatt	gattacatat	aagaaaaatc	8580
aaaatcctaa	agtatactcc	aattthaaatt	aagaagctaa	thctaaatatt	atathgggaat	8640
gtcaaggatg	tagaatagcc	acagtgaacc	tgaagaaaca	ccaaaatgag	aactthcagth	8700
gcctgaatac	ctggaatata	gtgtgggtgg	cagthatggtg	atggtgagat	cagaagthta	8760
aaaatttgca	aacgtgctta	thttthggaaa	thaatcactac	gcagatgtag	ccaaaccctc	8820
thtcaactgtg	gcgtacgcac	ctgctcgaaa	agggagthtga	gctthcagcag	ctgaccctagg	8880
actctgtthcc	cctthtggtga	gaagggtthtt	tgtthcagcaa	gacaatggag	agctctcact	8940
gtggtggacg	thtcggccaag	ggaccaaggt	ggaaatcaaa	cgtgagthaga	atthtaaactth	9000
tgctthcctca	gttgtctgtg	thttctgtthc	cctgtgtctca	tgaagthgathc	tataagctga	9060
ctctgcaatc	agcctctgat	atcctthcagg	gaaaagataa	agataagthct	gthagthcaaac	9120
thcgagaattg	atthgcacatt	thctththgaag	agcaagcaag	atthcagthcat	tgggtgagaa	9180
thaaactthgtct	aagthaatagc	thtcagaaatg	thcctggggaa	cataacatgt	thctggacaga	9240
gcctthggtca	atthgtcagaa	agggagthttt	tgtataggag	ggaagthtaag	aggaaccatt	9300
gtgtgtacac	thttthggccag	gggaccaagc	tggagatcaa	acgthaaagthac	thttththccac	9360
tgattctthca	ctgtthgctaa	thagthththact	thgtgtthcct	thgtgtggat	ththcattagth	9420
cggatgcccag	ggatthcaaca	aactthcattc	ccaggtthtagg	thacagaggag	gggaaattgt	9480

ES 2 678 221 T3

tccacaggaa gctagcttgt ggctaatttt taagatttct aatcaaaat aacttcattg 9540  
 ggggaaaagag gcttgctgag ctttcagggg ggtttttgta aagggaaaag ttaagacgaa 9600  
 tcaactgtgat tcaactttcgg ccctgggacc aaagtggata tcaaacgtaa gtacatctgt 9660  
 ctcaattatt cgtgagatth tagtgccatt gtatcatttg tgcaagtttt gtgatatttt 9720  
 ggttgaataa acctgggtgac ccagaagtaa atagcaggac accagaaaat gaacttaaaa 9780  
 agctgagcaa atagacgaat cattgggttt gagaggagaa taggattcat gggggaaatg 9840  
 ggggaagaaat agctagatth ttctctgaac aagcagccta tctcatatga ttggcttcaa 9900  
 gagaggthtt tgttgagggg aaagggtgag atccctcact gtggctcact ttcggcggag 9960  
 ggaccaaggt ggagatcaaa cgtaagtgca ctttcctaata gctttttctt ataaggthtt 10020  
 aaatttgag cgthttttgtg tttgagatat tagctcaggt caattccaaa gagtaccaga 10080  
 ttctttcaaa aagtcagatg agtaagggat agaaaattag ttcactctta ggaacagcca 10140  
 agcgctagcc agttaagtga ggcatctcaa ttgcaagatt ttctctgcat cggtcaggtt 10200  
 agtgatatta acagcgaaaa gagatthttg ttaaggggaa agtaattaag ttaactctgt 10260  
 ggatcacctt cggccaaggg acacgactgg agattaaacg taagtaattt ttcactattg 10320  
 tcttctgaaa tttgggtctg atggccagta ttgactthta gaggcttaa taggagthttg 10380  
 gtaaagattg gtaaagtgg gcatthtaaga tttgccatgg gttgcaaaag ttaactcag 10440  
 cttcaaaaat ggatttgag aaaaaaagat taaattgctc taaactgaat gacacaaagt 10500  
 aaaaaaaaa gtgtaactaa aaaggaacc ttgtatttct aaggagcaaa agtaaattta 10560  
 tttttgttca ctcttgccaa atattgtatt ggttggtgct gattatgcat gatacagaaa 10620  
 agtggaaaaa tacatthttt agtctthtct cctthttgtht gataaattat tttgtcagac 10680  
 aactataaaa atcaatagca cgcctaaga tctagatgca tgctcgagt ccatttcatt 10740  
 acctcttht cgcacccga catagataaa gcttatcgat accgtcgacc tcgagggggg 10800  
 gcccggtacc acgtggcata cagtgtcaga thttctgtht atcaagctag tgagattagg 10860  
 ggcaaaa 10867

5 <210> 84  
 <211> 23  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Sintética

<400> 84  
 tctttgccc cggatccgat cag 23

15 <210> 85  
 <211> 23  
 <212> ADN

<213> Artificial  
 <220>  
 <223> Sintética  
 5  
 <400> 85  
 gggaggctcc tctgaactct aag 23  
 <210> 86  
 <211> 21  
 <212> ADN  
 <213> Artificial  
 10  
 <220>  
 <223> Sintética  
 15  
 <400> 86  
 gtccagtcac toggttgcta t 21  
 20  
 <210> 87  
 <211> 22  
 <212> ADN  
 <213> Artificial  
 25  
 <220>  
 <223> Sintética  
 <400> 87  
 ctcaactgt ggcgtacgca cc 22  
 30  
 <210> 88  
 <211>21  
 <212> ADN  
 <213> Artificial  
 35  
 <220>  
 <223> Sintética  
 <400> 88  
 acgcagatgt agccaaacc t 21  
 40  
 <210> 89  
 <211> 20  
 <212> ADN  
 <213> Artificial  
 45  
 <220>  
 <223> Sintética  
 <400> 89  
 cagctgctga agctcaactc 20  
 50  
 <210> 90  
 <211> 22  
 <212> ADN  
 <213> Artificial  
 55  
 <220>  
 <223> Sintética  
 60  
 <400> 90  
 aagaagcaca cgactgaggc ac 22  
 <210>91

<211> 24  
 <212> ADN  
 <213> Artificial

5

<220>  
 <223> Sintética

<400> 91  
 ctactggat ggtggaaga tgga 24

10

<210> 92  
 <211> 16  
 <212> ADN  
 <213> Artificial

15

<220>  
 <223> Sintética

<400> 92  
 gtaaacgac ggccag 16

20

<210> 93  
 <211> 17  
 <212> ADN  
 <213> Artificial

25

<220>  
 <223> Sintética

<400> 93  
 caggaaacag ctatgac 17

30

<210> 94  
 <211> 39  
 <212> PRT  
 <213> Artificial

35

<220>  
 <223> Sintética

40

<400> 94

Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr  
 1 5 10 15

Pro Tyr Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Ala Asp  
 20 25 30

Ala Ala Pro Thr Val Ser Ile  
 35

45

<210> 95  
 <211> 39  
 <212> PRT  
 <213> Artificial

50

<220>  
 <223> Sintética

<400> 95

55

ES 2 678 221 T3

Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr  
 1 5 10 15

Pro Phe Thr Phe Gly Pro Gly Thr Lys Val Asp Ile Lys Arg Ala Asp  
 20 25 30

Ala Ala Pro Thr Val Ser Ile  
 35

5 <210> 96  
 <211> 39  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Sintética

<400> 96

Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
 1 5 10 15

Pro Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Asp  
 20 25 30

Ala Ala Pro Thr Val Ser Ile  
 35

15 <210> 97  
 <211> 38  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> Sintética

<400> 97

Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
 1 5 10 15

Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Asp Ala  
 20 25 30

Ala Pro Thr Val Ser Ile  
 35

25 <210> 98  
 <211> 39  
 <212> PRT  
 <213> Artificial

30 <220>  
 <223> Sintética

<400> 98

35 <400> 98

ES 2 678 221 T3

Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
1 5 10 15

Pro Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Asp  
20 25 30

Ala Ala Pro Thr Val Ser Ile  
35

5 <210> 99  
<211> 39  
<212> PRT  
<213> Artificial

10 <220>  
<223> Sintética  
<400> 99

Asp Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
1 5 10 15

Pro Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Asp  
20 25 30

Ala Ala Pro Thr Val Ser Ile  
35

15 <210> 100  
<211> 39  
<212> PRT  
<213> Artificial

20 <220>  
<223> Sintética  
<400> 100

Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
1 5 10 15

Pro Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Asp  
20 25 30

Ala Ala Pro Thr Val Ser Ile  
35

25 <210> 101  
<211> 39  
<212> PRT  
<213> Artificial

30 <220>  
<223> Sintética  
<400> 101

ES 2 678 221 T3

Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
 1 5 10 15

Pro Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Asp  
 20 25 30

Ala Ala Pro Thr Val Ser Ile  
 35

5 <210> 102  
 <211> 39  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Sintética  
 <400> 102

Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
 1 5 10 15

Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Ala Asp  
 20 25 30

Ala Ala Pro Thr Val Ser Ile  
 35

15 <210> 103  
 <211> 39  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> Sintética  
 <400> 103

Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser  
 1 5 10 15

Pro Tyr Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Ala Asp  
 20 25 30

Ala Ala Pro Thr Val Ser Ile  
 35

30 <210> 104  
 <211> 39  
 <212> PRT  
 <213> Artificial

35 <220>  
 <223> Sintética  
 <400> 104



ES 2 678 221 T3

agattttgca ggtattact gtcacatca tggctactca cctccacag tgattcagct 60

5 <210> 110  
<211> 60  
<212> ADN  
<213> Artificial

10 <220>  
<223> Sintética

<400> 110  
agattttgca ggtattact gtcacatca tggctactca cctccacag tgattcagct 60

15 <210> 111  
<211> 18  
<212> ADN  
<213> Artificial

20 <220>  
<223> Sintética

<400> 111  
aacttactac tgtcacca 18

25 <210> 112  
<211> 19  
<212> ADN  
<213> Artificial

30 <220>  
<223> Sintética

<400> 112  
cagcagtctg caacctgaa 19

35 <210> 113  
<211> 18  
<212> ADN  
<213> Artificial

40 <220>  
<223> Sintética

<400> 113  
ggctgtcct cacacatc 18

45 <210> 114  
<211> 17  
<212> ADN  
<213> Artificial

50 <220>

<223> Sintética

55 <400> 114  
ttactgtcac catcatg 17

60 <210> 115  
<211> 19  
<212> ADN  
<213> Artificial

65 <220>  
<223> Sintética

<400> 115  
 gcagactgga gctgaaga 19

5 <210> 116  
 <211> 23  
 <212> ADN  
 <213> Artificial

10 <220>  
 <223> Sintética

<400> 116  
 aagctgaatc actgtgggag gtg 23

15 <210> 117  
 <211> 51  
 <212> ADN  
 <213> Artificial

20 <220>  
 <223> Sintética

<400> 117  
 caactfacta ctgtcaacag agttacagta cccctccac agtgttacaa g 51

25 <210> 118  
 <211> 54  
 <212> ADN  
 <213> Artificial

30 <220>  
 <223> Sintética

35 <400> 118  
 atttgcaac ttactactgt caacatagtc acagtacca tcccacagtg ttac 54

40 <210> 119  
 <211> 54  
 <212> ADN  
 <213> Artificial

45 <220>  
 <223> Sintética

<400> 119  
 gtaacactgt gggatgggta ctgtgactat gttgacagta gtaagttgca aaat 54

50 <210> 120  
 <211> 41  
 <212> ADN  
 <213> Artificial

55 <220>  
 <223> Sintética

<400> 120  
 gcagtgatt actgcagca gtatgtagc tcacctcca c 41

60 <210> 121  
 <211> 56  
 <212> ADN  
 <213> Artificial

65 <220>

ES 2 678 221 T3

<223> Sintética

<400> 121  
 attttgcagt gtattactgt caccattatg gtcactcacc tcccacagtg attcag 56

5

<210> 122  
 <211> 56  
 <212> ADN  
 <213> Artificial

10

<220>  
 <223> Sintética

<400> 122  
 ctgaatcact gtgggaggtg agtgaccata atggtgacag taatacactg caaaat 56

15

<210> 123  
 <211> 19  
 <212> ADN  
 <213> Artificial

20

<220>  
 <223> Sintética

25

<400> 123  
 cttactactg tcaacatag 19

30

<210> 124  
 <211> 19  
 <212> ADN  
 <213> Artificial

35

<220>  
 <223> Sintética

<400> 124  
 cagcagctg caacctgaa 19

40

<210> 125  
 <211> 18  
 <212> ADN  
 <213> Artificial

45

<220>  
 <223> Sintética

<400> 125  
 ggctcgtcct cacacatc 18

50

<210> 126  
 <211> 17  
 <212> ADN  
 <213> Artificial

55

<220>  
 <223> Sintética

<400> 126  
 tactgtcacc attatgg 17

60

<210> 127  
 <211> 19  
 <212> ADN  
 <213> Artificial

<220>  
 <223> Sintética

5 <400> 127  
 gcagactgga gcctgaaga 19

<210> 128  
 <211> 23  
 <212> ADN  
 <213> Artificial

10

<220>  
 <223> Sintética

15 <400> 128  
 aagctgaatc actgtgggag gtg 23

<210> 129  
 <211> 22  
 <212> ADN  
 <213> Artificial

20

<220>  
 <223> Sintética

25 <400> 129  
 aagaagcaca cgactgaggc ac 22

<210> 130  
 <211> 22  
 <212> ADN  
 <213> Artificial

30

<220>  
 <223> Sintética

35 <400> 130  
 ggaagatgga tacagttggt gc 22

40 <210> 131  
 <211> 95  
 <212> PRT  
 <213> Artificial

45

<220>  
 <223> Sintética

50 <400> 131

ES 2 678 221 T3

Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly  
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser  
 20 25 30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu  
 35 40 45

Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Ile Pro Asp Arg Phe Ser  
 50 55 60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu  
 65 70 75 80

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

<210> 132  
 <211> 96  
 <212> PRT  
 <213> Artificial

5

<220>  
 <223> Sintética

10

<400> 132

Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly  
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser  
 20 25 30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu  
 35 40 45

Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Ile Pro Asp Arg Phe Ser  
 50 55 60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu  
 65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Ser Pro  
 85 90 95

15 <210> 133  
 <211> 95  
 <212> PRT  
 <213> Artificial

20 <220>

ES 2 678 221 T3

<223> Sintética

<400> 133

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr  
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile  
35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly  
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
65 70 75 80

5 Glu Asp Phe Ala Thr Tyr Tyr Cys Gln His Ser His Ser Thr His  
85 90 95

<210> 134

<211> 95

<212> PRT

10 <213> Artificial

<220>

<223> Sintética

15 <400> 134

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr  
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile  
35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly  
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr Pro  
85 90 95

<210> 135

ES 2 678 221 T3

<211> 95  
 <212> PRT  
 <213> Artificial

5 <220>  
 <223> Sintética

<400> 135

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr  
 20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile  
 35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly  
 50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
 65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln His Ser His Ser Thr Lys  
 85 90 95

10

<210> 136  
 <211> 506  
 <212> ADN  
 <213> Desconocido

15

<220>  
 <223> mamífero

20

<400> 136

atggagtttg ggctgagctg gctttttctt gtggctattt taaaaggtaa ttcattggaga 60  
 aatagaaaaa ttgagtgtga atggataaga gtgagagaaa cagtggatac gtgtggcagt 120  
 ttctgaccag ggtttctttt tgtttgcagg tgtccagtgt gaggtgcagc tgttggagtc 180  
 tgggggaggc ttggtacagc ctgggggggtc cctgagactc tcctgtgcag cctctggatt 240  
 caccttttagc aactctccaa tgagctgggt ccgccaggct ccaggggaagg ggctggagtg 300  
 ggtctcagct attagtggta gtggtagtaa cacattctac gcagactccg tgaagggccg 360  
 gttcaccatc tccagagaca attccaagaa cacgctgcat ctgcaagtga acagcctgag 420  
 agccgacgac acggccgtat attactgtgc gaaaggttac tactttgact actggggcca 480  
 gggaaacctg gtcaccgtct cctcag 506

25

<210> 137  
 <211> 134  
 <212> PRT

ES 2 678 221 T3

<213> Desconocido

<220>

<223> mamífero

5

<400> 137

Met Glu Phe Gly Leu Ser Trp Leu Phe Leu Val Ala Ile Leu Lys Gly  
1 5 10 15

Val Gln Cys Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln  
20 25 30

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe  
35 40 45

Ser Asn Ser Pro Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu  
50 55 60

Glu Trp Val Ser Ala Ile Ser Gly Ser Gly Ser Asn Thr Phe Tyr Ala  
65 70 75 80

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn  
85 90 95

Thr Leu His Leu Gln Val Asn Ser Leu Arg Ala Asp Asp Thr Ala Val  
100 105 110

Tyr Tyr Cys Ala Lys Gly Tyr Tyr Phe Asp Tyr Trp Gly Gln Gly Thr  
115 120 125

Leu Val Thr Val Ser Ser  
130

10

<210> 138

<211> 263

<212> ADN

<213> *Mus musculus*

15

<400> 138

catctacagc atctgtctag gggttacaaa aagtctatgg gatacaattc ctcagaaagg 60

aataggattt ggacctgagc atactgctgc ctaacacatg aaatggcagt tcttctccag 120

ctggactagg tccttaacta agaaatgcac tgctcatgaa tatgcaaatt acccaagtct 180

atggcagtaa atacagagat gtccacaccc tgaagacaac ctatggccaa tgcctctcc 240

acagtccctg aagacactga ttc 263

20

<210> 139

<211> 2239

<212> ADN

ES 2 678 221 T3

<213> *Homo sapiens*

<400> 139

gggtgtgtcct	tccaagtatc	cctgccccttc	cttcaggtgt	gatgctgatc	catgcatctt	60
cctcccattc	ctgggtagag	ggtctccttg	ttctttcccc	aatcttctcc	cagcagccac	120
tatgtcttcc	cactgatgtc	ttcagcttcc	tcttttctgc	ccttcccagc	agatgagtct	180
gttattcctg	aggggaagaga	gaggaggtgg	gagatgaggc	tttattcctg	aaagaatgga	240
cagagctttg	gagcttttct	ttcttctacc	ccagtttcta	cgagttcctc	cagtgcctgt	300
aacacagagc	tgtcagtggc	tttgcccctg	catgttaatt	ctttggtcat	gtagtgtgg	360
tgagggagct	gggtctggat	gcatttcagc	cacagctgct	gttccacttc	cccacacaga	420
accgtctcag	gaggaagagg	ttggagattt	ttctgatgtg	tcctcaatcc	tgggagaaaa	480
5 gttttcaata	gcaataggaa	tctctcagtt	gtatgtccgc	tgagaattta	aacaataact	540

ES 2 678 221 T3

tctttataat	actaaatfff	aaacaatcca	atgaatatat	ctaatttaat	tttttattag	600
cttatgtgac	atttgctggc	ctctccccc	gataaggtca	tattctcttc	ctgtttgttc	660
ctgcaagtca	ctgaatttca	caagatgtaa	ggtctcttga	ttatactata	acatcaataa	720
tctggtgtac	taaagaaaat	gtgctaattt	gcagcttact	aaaattagtt	gttttagtaa	780
aaataatata	tttgttatag	catgcctaca	tctccaggct	gagtagaagc	tttgtttgta	840
atactcagaa	actgcaaata	acacacatat	taaggagctt	tctaaataaa	tactatgggt	900
gaattcttta	actggaatac	tgttcctcat	tagaatcaac	acattcttta	tacacaacca	960
aattactgac	tcacaagtaa	ttatactgag	caagaagcca	cacaaatgaa	attgaaaata	1020
taagatttca	ctttaacaaa	atcttgaaaa	atagatctga	attaacaaag	atgagagggt	1080
gactcagtat	ggtgagaggg	aaataatggg	gaggcaggaa	tgagaggaac	acaggggaggc	1140
ttttaactgt	aatcatcgtt	atctcgattg	ttatfffgga	tacacaggtg	agcacacgtg	1200
aaattacgtt	ttttttcctt	cggagacatg	gttttgcttt	gtcgtccagg	ctggagctcc	1260
gtggtgagat	catagctctc	tgcagcctca	gaccctgggt	ctccagcaat	cctcctccct	1320
ctgccctatg	tagctgggac	tacaggtgtg	caccaccatg	ctcggcttca	gtgcgtatta	1380
aaccgcacac	ttcaattatg	cagtatttat	tatatagcaa	taatgcctca	ataaggggat	1440
tacaaataag	tgatagata	atftgttaaa	gattgatgga	aagatagata	ccaacatgag	1500
aaatgtatga	cactcaagaa	aataaaaactg	taggaaactt	gctfffcttt	atatttgfta	1560
ggtaatcacc	acagtgtgta	cacatcacac	catgftccca	ttacagagaa	aaggftctgc	1620
gaacctcacg	agctgtgacc	cctgtgtgct	gggcttggtt	cagggagaag	tcaggtccag	1680
tggtgagaag	cacaggccca	gatgccaggg	ctcactctga	ccaaaagtga	gcaactggga	1740
cattgtataaa	cccacctgtg	ctfffgtctga	taatfffcca	tctftaacat	ggaaataata	1800
ttgatactat	ataccatggt	ttctctgctg	atgtaaaaat	aaaagatgat	tggtgctaac	1860
tttaaaaata	tgcagfttat	gtagatctat	ggtacctcaa	taaaactgft	ttaaaataaa	1920
aattacaaaa	ttataagatt	tttaggtfff	aaggfttaag	tttatcacia	aacaaactga	1980
caataggaaa	gcacaatttc	ccaatgcttt	caatatcaca	gatctccccg	aggacattct	2040
gacatgctct	gagccccaact	atctccaaag	gcctctcacc	ccagagctta	ctatatagta	2100
ggagatatgc	aaatagagcc	ctccgtctgc	tgatgaaaac	cagcccagcc	ctgaccctgc	2160
agctctgaga	gaggagccca	gccctgggat	tttcaggtgt	tttcatttgg	tgatcaggac	2220
tgaacagaga	gaactcacc					2239

<210> 140  
 <211> 200  
 <212> ADN

ES 2 678 221 T3

<213> *Homo sapiens*

<400> 140

```

gtgagtcctc acaacctctc tcctgcttta actctgaagg gttttgctgc atttttgggg      60
ggaaataagg gtgctgggtc tcctgccaag agagccccgg agcagcctgg ggggctcagg      120
aggatgccct gaggcaacag cggccacaca gacgaggggc aagggtcca gatgctcctt      180
5  cctcctgagc ccagcagcac                                          200

```

<210> 141

<211> 255

<212> ADN

10 <213> *Mus musculus*

<400> 141

```

ctgagcattg cagactaatc ttggatattt gtccctgagg gagccggctg agagaagttg      60
ggaaataaac tgtctagggg tctcagagcc tttaggacag attatctcca catctttgaa      120
aaactaagaa tctgtgtgat ggtgttggtg gagtccctgg atgatgggat agggactttg      180
gaggctcatt tgaagaagat gctaaaacaa tcctatggct ggagggatag ttggggctgt      240
15 agttggagat tttca                                          255

```

<210> 142

<211> 4798

<212> ADN

20 <213> Desconocido

<220>

<223> mamífero

<400> 142

25

ES 2 678 221 T3

ctcgagcatc tacagcatct gtctaggggt tacaaaaagt ctatgggata caattcctca	60
gaaaggaata ggatttggac ctgagcatac tgctgcctaa cacatgaaat ggcagttctt	120
ctccagctgg actaggtcct taactaagaa atgcaactgct catgaatatg caaattaccc	180
aagtctatgg cagtaaatac agagatgtcc acaccctgaa gacaacctat ggccaatgtc	240
ctctccacag tccctgaaga cactgattct aactataacg gtcctaaggt agcgatcgag	300
tgccatttca ttacctcttt ctccgcaccc gacatagatg catctgcaga attcgcctt	360
ggggtaccga aaccttgcgc tcgttcgcca gccaggacag aaatgcctcg acttcgctgc	420
tgcccaaggt tgccgggtga cgcacaccgt ggaaacggat gaaggcacga acccagtgga	480
cataagcctg ttcggttcgt aagctgtaat gcaagtagcg tatgcgctca cgcaactggt	540
ccagaacctt gaccgaacgc agcggtggtta acggcgcagt ggcggttttc atggcttgtt	600
atgactgttt ttttggggta cagtctatgc ctcgggcatc caagcagcaa gcgcgttacg	660
ccgtgggtcg atgtttgatg ttatggagca gcaacgatgt tacgcagcag ggcagtcgcc	720
ctaaaacaaa gttaaaccatc atgaggggaag cggatgatcgc cgaagtatcg actcaactat	780

ES 2 678 221 T3

cagaggtagt tggcgtcatc gagcgccatc tcgaaccgac gttgctggcc gtacatttgt 840  
 acggctccgc agtggatggc ggctgaagc cacacagtga tattgatttg ctggttacgg 900  
 tgaccgtaag gcttgatgaa acaacgcggc gagctttgat caacgacctt ttggaaactt 960  
 cggcttcccc tggagagagc gagattctcc gcgctgtaga agtcaccatt gttgtgcacg 1020  
 acgacatcat tccgtggcgt tatccagcta agcgcgaact gcaatttga gaatggcagc 1080  
 gcaatgacat tcttgcaggt atcttcgagc cagccacgat cgacattgat ctggctatct 1140  
 tgctgacaaa agcaagagaa catagcgttg ccttggtagg tccagcggcg gaggaactct 1200  
 ttgatccggt tcctgaacag gatctatttg aggcgctaaa tgaaacctta acgctatgga 1260  
 actcgccgcc cgactgggct ggcgatgagc gaaatgtagt gcttacgttg tcccgcattt 1320  
 ggtacagcgc agtaaccggc aaaatcgcg cgaaggatgt cgctgccgac tgggcaatgg 1380  
 agcgcctgcc ggcccagtat cagcccgtca tacttgaagc tagacaggct tatcttggac 1440  
 aagaagaaga tcgcttggcc tcgcgcgcag atcagttgga agaatttgtc cactacgtga 1500  
 aaggcgagat caccaaggta gtcggcaaat aatgtctcct aggcacgaag ggcgaattcc 1560  
 agcacactgg cggccgttac tagtggcgcg ccggtgtgtc cttccaagta tccctgccct 1620  
 tccttcaggt gtgatgctga tccatgcac ttcctcccat tccctggtag agggctctct 1680  
 tgttctttcc ccaatcttct cccagcagcc actatgtctt cccactgatg tcttcagctt 1740  
 cctcttttct gcccttccca gcagatgagt ctgttattcc tgagggaaga gagaggaggt 1800  
 gggagatgag gctttattcc tgaaagaatg gacagagctt tggagctttt ctttcttcta 1860  
 ccccagttc tacgagttcc tccagtgcct gtaacacaga gctgtcagtg gctttgcccc 1920  
 tgcatgttaa ttctttggtc atgtagtggg ggtgaggag ctgggtctgg atgcatttca 1980  
 gccacagctg ctgttccact tccccacaca gaaccgtctc aggaggaaga ggttggagat 2040  
 tttctgatg tgcctcaat cctgggagaa aagttttcaa tagcaatagg aatctctcag 2100  
 ttgtatgtcc gctgagaatt taaacaataa cttctttata atactaaatt ttaaacaatc 2160  
 caatgaatat atctaattta attttttatt agcttatgtg acatttgctg gcctctcccc 2220  
 cagataaggt catattctct tcctgtttgt tcctgcaagt cactgaattt cacaagatgt 2280  
 aaggctctct gattatacta taacatcaat aatctgggtg actaaagaaa atgtgctaat 2340  
 ttgcagctta ctaaaattag ttgttttagt aaaaataata tatttggtat agcatgccta 2400  
 catctccagg ctgagtagaa gctttgtttg taatactcag aaactgcaaa taacacacat 2460  
 attaaggagc tttctaaata aatactatgg gtgaattctt taactggaat actgttctct 2520  
 attagaatca acacattctt tatacacaac caaattactg actcacaagt aattatactg 2580  
 agcaagaagc cacacaaatg aaattgaaaa tataagattt cactttaaca aaatcttgaa 2640

ES 2 678 221 T3

aaatagatct gaattaacaa agatgagagg ttgactcagt atggtgagag ggaataatg 2700  
 gggaggcagg aatgagagga acacagggag gcttttaact gtaatcatcg ttatctcgat 2760  
 tgttatthttg gatacacagg tgagcacacg tgaattacg tttttttcc ttcggagaca 2820  
 tggthttgct ttgtcgtcca ggctggagct ccgtggtgag atcatagctc tctgcagcct 2880  
 cagaccctgg gtctccagca atcctcctcc ctctgcccta tgtagctggg actacaggtg 2940  
 tgcaccacca tgctcggctt cagtgcgtat taaaccgcac acttcaatta tgcagtatth 3000  
 attatatagc aataatgcct caataagggt attacaaata agtggataga taatthgtta 3060  
 aagattgatg gaaagataga taccaacatg agaaatgtat gacactcaag aaaataaaac 3120  
 tgtaggaaac ttgctthttct ttatathttgt taggtaatca ccacagtgtg tacacatcac 3180  
 accatgthcc cattacagag aaaaggthct gcgaacctca cgagctgtga cccctgtgtg 3240  
 ctgggcttgg ttcagggaga agtcaggthcc agtggtgaga agcacaggcc cagatgcca 3300  
 ggctcactct gaccaaagt gagcactggg gacattgtaa aaccacactg tgctthttgct 3360  
 gataathttt catctthaac atggaaataa tattgatact atataccatg gthtctctgc 3420  
 gtatgtaaaa ataaaagatg attggtgcta actthaaaaa tatgcagtht atgtagatct 3480  
 atggtacctc aataaaactg ththaaaaata aaaattacaa aattataaga ththtaggtt 3540  
 thaaagthta agthtatcac aaaacaaact gacaatagga aagcacaatt thccaatgct 3600  
 thcaatatca cagatctccc cgaggacatt ctgacatgct ctgagcccca ctatctcaa 3660  
 aggcctctca cccagagct tactatatag taggagatat gcaaatagag cctccgtct 3720  
 gctgatgaaa accagcccag cctgaccct gcagctctga gagaggagcc cagccctggg 3780  
 atthtcaggt gththcattt ggtgatcagg actgaacaga gagaactcac catggagtht 3840  
 gggctgagct ggctthttct tgtggctatt thaaaaggta atthcatggag aatagaaaa 3900  
 atthgagtgt aatggataag agtgagagaa acagthgata cgtgtggcag thtctgacca 3960  
 gggthtctth ttgtthgcag gtgtccagthg tgagthgcag ctgtthgagth ctgggggag 4020  
 ctthggtacag cctgggggggt cctgagact ctctgtgca gcctctggat thacctthtag 4080  
 caactctcca atgagctggg thccgccaggc thcagggaaag gggctggagth gggthctcagc 4140  
 thattagthgt agthgtagta acacattcta cgcagactcc gthgaagggcc gththcaccat 4200  
 ctccagagac aatthcaaga acacgctgca thctgcaagthg aacagcctga gagccgacga 4260  
 cacggccgta thactctgtg cgaaaggthta ctactthgac thactggggcc agggaaacct 4320  
 ggtcacctgc thctcaggtg agthctcaca acctctctcc tgctthtaact ctgaagggth 4380  
 thgtctcatt ththggggga aataaggthg ctgggtctcc thccaagaga gccccggagc 4440  
 agcctggggg gctcaggagthg atgcccctgag gcaacagcgg ccacacagac gaggggcaag 4500  
 ggctccagat gctcctctct cctgagccca gcagcacctg agcattgcag actaatctthg 4560

ES 2 678 221 T3

gatatttgtc cctgagggag ccggctgaga gaagttggga aataaactgt ctagggatct 4620  
 cagagccttt aggacagatt atctccacat ctttgaaaaa ctaagaatct gtgtgatggt 4680  
 gttggtggag tccctggatg atgggatagg gactttggag gctcatttga agaagatgct 4740  
 aaaacaatcc tatggctgga gggatagttg gggctgtagt tggagatfff cactcgag 4798

5 <210> 143  
 <211> 8  
 <212> PRT  
 <213> Desconocido

10 <220>  
 <223> mamífero  
 <400> 143

Ala Lys Gly Tyr Tyr Phe Asp Tyr  
 1 5

15 <210> 144  
 <211> 5  
 <212> PRT  
 <213> Desconocido

20 <220>  
 <223> mamífero  
 <400> 144

25 Asp Tyr Ser Asn Tyr  
 1 5

30 <210> 145  
 <211> 445  
 <212> ADN  
 <213> Desconocido

35 <220>  
 <223> mamífero  
 <400> 145

ctcgagccac catggagttc gggctgtcct ggctttttct tgtggccatt ctgaaggggtg 60  
 tgcaatgtga agtgcagttg ctggagtccg gggggggcct tgtgcagcct ggaggatcac 120  
 tccggctgag ctgctcagcc agtggtttca cgttcagtag ttatgctatg tcttgggtgc 180  
 gccaggcccc tggtaagggg ctggaatggg tgtcagctat ttccggcagc ggcggatcta 240  
 cttattacgc tgatagcgtg aagggacgct tcacaatctc tcgggacaac tccaaaaaca 300  
 ccctctatct tcagatgaat agcctccgcg ctgaggacac cgctgtttat tactgctcca 360  
 aaaccacagt gacatactac tactattatg gcatggacgt ctggggtcag ggaacaaccg 420  
 tcaccgtgtc cagcgctga agagc 445

40 <210> 146  
 <211> 448  
 <212> ADN

ES 2 678 221 T3

	<213> Desconocido	
	<220>	
	<223> mamífero	
5	<400> 146	
	ctcgagccac catggagttc gggctgtcct ggctttttct tgtggccatt ctgaaggggtg	60
	tgcaatgtga agtgcagttg ctggagtccg gggggggcct tgtgcagcct ggaggatcac	120
	tccggctgag ctgvcgagcc agtggtttca cgttcagtag ttatgctatg tcttgggtgc	180
	gccaggcccc tggttaagggg ctggaatggg tgtcagctat ttccggcagc ggcggatcta	240
	cttattacgc tgatagcgtg aagggacgct tcacaatctc tcgggacaac tccaaaaaca	300
	ccctctatct tcagatgaat agcctccgcg ctgaggacac cgctgtttat tactgcgcca	360
	aagattattc aaattactac tactactatt atggcatgga cgtctgggggt caggaacaa	420
	ccgtcaccgt gtccagcgcc tgaagagc	448
10	<210> 147	
	<211> 430	
	<212> ADN	
	<213> Desconocido	
15	<220>	
	<223> mamífero	
	<400> 147	
	ctcgagccac catggagttc gggctgtcct ggctttttct tgtggccatt ctgaaggggtg	60
	tgcaatgtga agtgcagttg ctggagtccg gggggggcct tgtgcagcct ggaggatcac	120
	tccggctgag ctgvcgagcc agtggtttca cgttcagtag ttatgctatg tcttgggtgc	180
	gccaggcccc tggttaagggg ctggaatggg tgtcagctat ttccggcagc ggcggatcta	240
	cttattacgc tgatagcgtg aagggacgct tcacaatctc tcgggacaac tccaaaaaca	300
	ccctctatct tcagatgaat agcctccgcg ctgaggacac cgctgtttat tactgcgcca	360
	aaaccacagt gacatacttt gattactggg gacaggggac cctcgtcacc gtctcttctg	420
20	cctgaagagc	430
	<210> 148	
	<211> 433	
	<212> ADN	
25	<213> Desconocido	
	<220>	
	<223> mamífero	
30	<400> 148	

ES 2 678 221 T3

```

ctcgagccac catggagttc gggctgtcct ggctttttct tgtggccatt ctgaagggtg      60

tgcaatgtga agtgcagttg ctggagtccg gggggggcct tgtgcagcct ggaggatcac      120
tccggctgag ctgcgcagcc agtggtttca cgttcagtag ttatgctatg tcttgggtgc      180
gccaggcccc tggttaagggg ctggaatggg tgtcagctat ttccggcagc ggcggatcta      240
cttattacgc tgatagcgtg aagggacgct tcacaatctc tcgggacaac tccaaaaaca      300
ccctctatct tcagatgaat agcctccgcg ctgaggacac cgctgtttat tactgcgcca      360
aagattattc aaattactac tttgattact ggggacaggg gaccctcgtc accgtctctt      420
ctgcctgaag agc                                                                433

```

<210> 149  
 <211> 141  
 <212> PRT  
 <213> Desconocido

<220>  
 <223> mamífero

<400> 149

```

Met Glu Phe Gly Leu Ser Trp Leu Phe Leu Val Ala Ile Leu Lys Gly
 1                               5                               10                               15

Val Gln Cys Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
                20                               25                               30

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
                35                               40                               45

Ser Ser Tyr Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
 50                               55                               60

Glu Trp Val Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala
65                               70                               75                               80

Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn
                85                               90                               95

Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
                100                               105                               110

Tyr Tyr Cys Ala Lys Thr Thr Val Thr Tyr Tyr Tyr Tyr Tyr Gly Met
                115                               120                               125

Asp Val Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser
130                               135                               140

```

ES 2 678 221 T3

<210> 150  
 <211> 140  
 <212> PRT  
 <213> Desconocido  
 5  
 <220>  
 <223> mamífero  
 <400> 150  
 10

Met	Glu	Phe	Gly	Leu	Ser	Trp	Leu	Phe	Leu	Val	Ala	Ile	Leu	Lys	Gly
1				5					10					15	
Val	Gln	Cys	Glu	Val	Gln	Leu	Leu	Glu	Ser	Gly	Gly	Gly	Leu	Val	Gln
			20					25					30		
Pro	Gly	Gly	Ser	Leu	Arg	Leu	Ser	Cys	Ala	Ala	Ser	Gly	Phe	Thr	Phe
		35					40					45			
Ser	Ser	Tyr	Ala	Met	Ser	Trp	Val	Arg	Gln	Ala	Pro	Gly	Lys	Gly	Leu
	50					55					60				
Glu	Trp	Val	Ser	Ala	Ile	Ser	Gly	Ser	Gly	Gly	Ser	Thr	Tyr	Tyr	Ala
65					70					75					80
Asp	Ser	Val	Lys	Gly	Arg	Phe	Thr	Ile	Ser	Arg	Asp	Asn	Ser	Lys	Asn
				85					90					95	
Thr	Leu	Tyr	Leu	Gln	Met	Asn	Ser	Leu	Arg	Ala	Glu	Asp	Thr	Ala	Val
			100					105					110		
Tyr	Tyr	Cys	Ala	Lys	Asp	Tyr	Ser	Asn	Tyr	Tyr	Tyr	Tyr	Tyr	Tyr	Gly
		115					120					125			
Met	Asp	Val	Trp	Gly	Gln	Gly	Thr	Thr	Val	Thr	Val				
	130					135					140				

<210> 151  
 <211> 17  
 <212> ADN  
 <213> Desconocido  
 15

<220>  
 <223> mamífero  
 20

<400> 151  
 tgcggccgat ctagcc 17

<210> 152  
 <211> 21  
 <212> ADN  
 <213> Desconocido  
 25

<220>  
 <223> mamífero  
 30

<400> 152  
 acgagcgggt tcggcccatt c 21

5 <210> 153  
 <211> 18  
 <212> ADN  
 <213> Desconocido

10 <220>  
 <223> mamífero

<400> 153  
 ttgaccgatt ccttgcg 18

15 <210> 154  
 <211> 20  
 <212> ADN  
 <213> Desconocido

20 <220>  
 <223> mamífero

<400> 154  
 tcaggtgtga tgctgatcca 20

25 <210> 155  
 <211> 25  
 <212> ADN  
 <213> Desconocido

30 <220>  
 <223> mamífero

<400> 155  
 tgcatctcc tccattcct ggga 25

35 <210> 156  
 <211> 20  
 <212> ADN  
 <213> Desconocido

40 <220>  
 <223> mamífero

<400> 156  
 tggctgctgg gagaagattg 20

45 <210> 157  
 <211> 18  
 <212> ADN  
 <213> Desconocido

50 <220>  
 <223> mamífero

<400> 157  
 cccagcagcc actatgtc 18

55 <210> 158  
 <211> 25  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> mamífero  
  
 <400> 158  
 5 tccactgat gtcttcagct tcctc 25  
  
 <210> 159  
 <211> 23  
 <212> ADN  
 10 <213> Desconocido  
  
 <220>  
 <223> mamífero  
  
 <400> 159  
 15 ccctcaggaa taacagactc atc 23  
  
 <210> 160  
 <211> 19  
 20 <212> ADN  
 <213> Desconocido  
  
 <220>  
 <223> mamífero  
  
 <400> 160  
 25 aagggccggt tcaccatct 19  
  
 <210> 161  
 <211> 26  
 30 <212> ADN  
 <213> Desconocido  
  
 <220>  
 <223> mamífero  
  
 <400> 161  
 35 ccagagacaa ttccaagaac acgctg 26  
  
 <210> 162  
 <211> 18  
 40 <212> ADN  
 <213> Desconocido  
  
 <220>  
 <223> mamífero  
  
 <400> 162  
 45 cggctctcag gctgttca 18  
  
 <210> 163  
 <211> 20  
 <212> ADN  
 <213> Desconocido  
  
 <220>  
 <223> mamífero  
  
 <400> 163  
 55 gcagcctctg gattcacctt 20  
  
 <210> 164  
 <211> 24  
 <212> ADN

<213> Desconocido  
 <220>  
 <223> mamífero  
 5  
 <400> 164  
 agcaactctc caatgagctg ggtc 24  
 <210> 165  
 <211> 21  
 <212> ADN  
 <213> Desconocido  
 10  
 <220>  
 <223> mamífero  
 15  
 <400> 165  
 tcacggagtc tgcgtagaat g 21  
 <210> 166  
 <211> 20  
 <212> ADN  
 <213> Desconocido  
 20  
 <220>  
 <223> mamífero  
 25  
 <400> 166  
 ctcaagtatt ctggccctgc 20  
 <210> 167  
 <211> 32  
 <212> ADN  
 <213> Desconocido  
 30  
 <220>  
 <223> mamífero  
 35  
 <400> 167  
 tgctccacag ctacaaacc cttcctataa tg 32  
 <210> 168  
 <211> 21  
 <212> ADN  
 <213> Desconocido  
 40  
 <220>  
 <223> mamífero  
 45  
 <400> 168  
 ggatgatggc tcagcacaga g 21  
 <210> 169  
 <211> 19  
 <212> ADN  
 <213> Desconocido  
 50  
 <220>  
 <223> mamífero  
 55  
 <400> 169  
 ggtggagagg ctattcggc 19  
 <210> 170

<211> 23  
 <212> ADN  
 <213> Desconocido

5

<220>  
 <223> mamífero

<400> 170  
 tgggcacaac agacaatcgg ctg 23

10

<210> 171  
 <211> 17  
 <212> ADN  
 <213> Desconocido

15

<220>  
 <223> mamífero

<400> 171  
 gaacacggcg gcatcag 17

20

<210> 172  
 <211> 23  
 <212> ADN  
 <213> Desconocido

25

<220>  
 <223> mamífero

<400> 172  
 ggtgtgcgat gtaccctctg aac 23

30

<210> 173  
 <211> 32  
 <212> ADN  
 <213> Desconocido

35

<220>

40

<223> mamífero

<400> 173  
 ctaaaaatgc tacacctggg gcaaaacacc tg 32

45

<210> 174  
 <211> 25  
 <212> ADN  
 <213> Desconocido

50

<220>  
 <223> mamífero

<400> 174  
 tgtggcagtt taatccagct ttatc 25

55

<210> 175  
 <211> 28  
 <212> ADN  
 <213> Desconocido

60

<220>  
 <223> mamífero

<400> 175

gatgggaaga gactggaac attgtac 28

5 <210> 176  
<211> 27  
<212> ADN  
<213> Desconocido

10 <220>  
<223> mamífero

<400> 176  
cctccactgt gtaatggct gccacaa 27

15 <210> 177  
<211> 26  
<212> ADN  
<213> Desconocido

20 <220>  
<223> mamífero

<400> 177  
ttcctctatt tcactcttg aggctc 26

25 <210> 178  
<211> 20  
<212> ADN  
<213> Desconocido

30 <220>  
<223> mamífero

<400> 178  
tggtcacctc caggagcctc 20

35 <210> 179  
<211> 30  
<212> ADN  
<213> Desconocido

40 <220>  
<223> mamífero

<400> 179  
agtctctgct tcccccttgt ggctatgagc 30

45 <210> 180  
<211> 21  
<212> ADN  
<213> Desconocido

50 <220>  
<223> mamífero

55 <400> 180  
gctgcagggt gtatcaggtg c 21

60 <210> 181  
<211> 17  
<212> ADN  
<213> Desconocido

<220>  
<223> mamífero

# ES 2 678 221 T3

<400> 181  
gccatgcaag gccaagc 17

5 <210> 182  
<211> 24  
<212> ADN  
<213> Desconocido

10 <220>  
<223> mamífero

<400> 182  
ccaggaaaat gctgccagag cctg 24

15 <210> 183  
<211> 24  
<212> ADN  
<213> Desconocido

20 <220>  
<223> mamífero

<400> 183  
agttcttgag ccttaggtg ctag 24

25

## REIVINDICACIONES

1. Una rata o ratón que en su genoma de línea germinal comprende, en un locus de cadena pesada de inmunoglobulina endógena, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada bajo el control de un promotor y unida operativamente a una secuencia génica de región constante de cadena pesada endógena, en la que la secuencia de nucleótidos de región variable de cadena pesada reordenada codifica la secuencia de  $V_{H3-23}/X_1X_2/J_H$ , en la que  $X_1$  y  $X_2$  es cualquier aminoácido, y en la que todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  funcionales endógenos están delecionados del locus de cadena pesada de inmunoglobulina endógena o se han hecho no funcionales.
2. La rata o el ratón de la reivindicación 1, en la que:
- (a)  $X_1$  es Gly y  $X_2$  es Tyr;
  - (b) el segmento génico  $J_H$  humano se selecciona del grupo que consiste en  $J_{H1}$ ,  $J_{H2}$ ,  $J_{H3}$ ,  $J_{H4}$ ,  $J_{H5}$ ,  $J_{H6}$  y una variante polimórfica de los mismos; o
  - (c) la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada codifica la secuencia de  $V_{H3-23}/GY/J_{H4-4}$  humana (SEQ ID NO: 137).
3. El ratón según la reivindicación 1.
4. La rata o el ratón de la reivindicación 1, en el que la secuencia génica de la región constante de cadena pesada se selecciona de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$  y una combinación de los mismos.
5. La rata o el ratón de la reivindicación 1, en la que:
- (a) todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  funcionales endógenos se reemplazan por la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada;
  - (b) un dominio variable de cadena pesada de inmunoglobulina codificado por la secuencia de nucleótidos de región variable de cadena pesada reordenada, no es inmunógeno para la rata o el ratón; o
  - (c) la rata o el ratón comprende un gen Adam6a, un gen Adam6b, o ambos.
6. La rata o el ratón de la reivindicación 1, en la que la rata o el ratón comprende adicionalmente una secuencia de nucleótidos que codifica un segmento génico V de cadena ligera ( $V_L$ ) de inmunoglobulina humana no reordenada y un segmento génico J de cadena ligera ( $J_L$ ) de inmunoglobulina humana no reordenada.
7. La rata o el ratón de la reivindicación 6, en la que:
- (a) la secuencia de nucleótidos que codifica el segmento génico V de cadena ligera ( $V_L$ ) no reordenada y el segmento génico de cadena ligera ( $J_L$ ) no reordenada, está unida operativamente a una secuencia génica de región constante de cadena ligera de inmunoglobulina; o
  - (b) el segmento génico de cadena ligera ( $V_L$ ) de inmunoglobulina humana no reordenada y el segmento génico ( $J_L$ ) de inmunoglobulina humana no reordenada están unidos operativamente, en un locus endógeno de rata o de ratón, a una secuencia génica de región constante de inmunoglobulina de rata o de ratón.
8. La rata o el ratón de la reivindicación 7, en la que la secuencia génica de región constante de cadena ligera se selecciona de una secuencia génica de región constante de rata, una secuencia génica de región constante de ratón o una secuencia génica de región constante humana.
9. La rata o el ratón de la reivindicación 8, en la que la secuencia génica de región constante de rata o de ratón se selecciona de una secuencia génica de región constante de ratón y una secuencia génica de región constante de rata.
10. Un método de creación de una rata o un ratón, comprendiendo el método:
- (a) modificar un genoma de una rata o de un ratón para delecionar o hacer no funcionales los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  de cadena pesada de inmunoglobulina funcional endógena; y
  - (b) colocar en el genoma, en el locus de cadena pesada de inmunoglobulina endógena, una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina reordenada, en el que la secuencia de nucleótidos de la región variable de cadena pesada reordenada, codifica la secuencia de  $V_{H3-23}/X_1X_2/J_H$  unida operativamente a una secuencia de región constante de cadena pesada endógena, en la que  $X_1$  y  $X_2$  es cualquier aminoácido.
11. La rata o el ratón de la reivindicación 1, que adicionalmente comprende en su genoma:

un locus de cadena ligera de inmunoglobulina que comprende uno o más segmentos génicos  $V_L$  y  $J_L$  de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera.

5 12. La rata o el ratón de la reivindicación 11, en la que:

(a)  $X_1$  es Gly y  $X_2$  es Tyr;

(b) el segmento génico  $J_H$  humano se selecciona del grupo que consiste en  $J_{H1}$ ,  $J_{H2}$ ,  $J_{H3}$ ,  $J_{H4}$ ,  $J_{H5}$ ,  $J_{H6}$  y una variante polimórfica de los mismos;

10 (c) la secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada codifica la secuencia de  $V_{H3-23}/GY/J_{H4-4}$  humana (SEQ ID NO: 137);

(d) el uno o más segmentos génicos  $V_L$  y  $J_L$  de cadena ligera de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, están unidos operativamente a la secuencia de ácido nucleico de región constante de cadena ligera de un locus endógeno;

15 (e) el locus de cadena ligera de inmunoglobulina comprende dos segmentos génicos  $V_L$  humanos,  $V_{K1-39}$  y  $V_{K3-20}$ ;

(f) al menos uno de los segmentos génicos  $V_L$  o  $J_L$  de cadena ligera humana, codifica uno o más codones de histidina que no están codificados por un segmento génico de región variable de cadena ligera de la línea germinal humana correspondiente; o

20 (g) al menos uno de los segmentos génicos  $V_L$  comprende una adición o sustitución de al menos un codón no de histidina codificado por la secuencia del segmento  $V_L$  de la línea germinal humana correspondiente, con un codón de histidina.

25 13. La rata o el ratón modificado de la reivindicación 12, en la que el codón de histidina añadido o sustituido está presente en la CDR3.

14. Un método para obtener una secuencia de ácido nucleico que codifica un dominio variable de cadena ligera ( $V_L$ ) de inmunoglobulina capaz de unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

30 (a) inmunizar a una rata o a un ratón según la reivindicación 11, 12 o 13, con un antígeno de interés o con un inmunógeno del mismo;

(b) permitir que la rata o el ratón genere una respuesta inmunitaria;

35 (c) aislar, de la rata o del ratón inmunizado, una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que pueda unirse al antígeno; y

(d) obtener de la célula, una secuencia de ácido nucleico que codifique el dominio variable de cadena ligera (dominio  $V_L$ ), que pueda unirse al antígeno.

15. El método de la reivindicación 14, en el que:

40 (i) la etapa de aislamiento (c) se realiza por separación de células activadas por fluorescencia (FACS) o citometría de flujo;

45 (ii) la célula que comprende la secuencia de ácido nucleico que codifica el dominio variable de cadena ligera que se une al antígeno, es un linfocito, en el que opcionalmente el linfocito comprende linfocitos citotóxicos naturales, linfocitos T o linfocitos B;

(iii) el método comprende adicionalmente: (c)' la fusión del linfocito con una célula cancerosa, siendo opcionalmente la célula cancerosa una célula de mieloma;

(iv) la secuencia de ácido nucleico de (d) está fusionada con una secuencia de ácido nucleico que codifica una secuencia de ácido nucleico de región constante de inmunoglobulina;

50 (v) la secuencia de ácido nucleico de región constante de cadena ligera es una secuencia kappa humana o una secuencia lambda humana;

(vi) la secuencia de ácido nucleico de (d) comprende una o más sustituciones o inserciones de codones de histidina que proceden del segmento génico  $V_L$  no reordenado en el genoma del roedor.

55 16. Un método para crear una proteína de unión a antígeno, que comprende un dominio variable de cadena ligera de inmunoglobulina que puede unirse a un antígeno independientemente de un dominio variable de cadena pesada, que comprende:

60 (a) inmunizar a una rata o a un ratón modificado genéticamente según la reivindicación 11, 12 o 13, con un primer antígeno que comprende un primer epítipo o una parte inmunógena del mismo, en el que el roedor comprende adicionalmente en su genoma:

65 (i) dos o más segmentos génicos de región variable de cadena ligera ( $V_L$  y  $J_L$ ) de inmunoglobulina humana, pero menos que el número de segmentos de tipo silvestre, unidos operativamente a una secuencia de ácido nucleico de región constante de cadena ligera de inmunoglobulina;

(b) permitir que la rata o el ratón genere una respuesta inmunitaria contra el primer epítipo o parte inmunógena del mismo;

(c) aislar, de la rata o del ratón, una célula que comprenda una secuencia de ácido nucleico que codifique un dominio variable de cadena ligera que se una específicamente al primer epítipo o a la parte inmunógena del mismo;

(d) obtener, de la célula de (c), la secuencia de ácido nucleico que codifique el dominio variable de cadena ligera que se una específicamente al primer epítipo o a la parte inmunógena del mismo;

(e) emplear la secuencia de ácido nucleico de (d), en una construcción de expresión, fusionada a una secuencia de ácido nucleico de región constante de inmunoglobulina humana; y

(f) expresar la secuencia de ácido nucleico de (d), en una línea celular de producción que exprese una cadena pesada de inmunoglobulina humana que se una específicamente a un segundo antígeno o epítipo para formar una proteína de unión a un antígeno, cuya cadena ligera esté codificada por el ácido nucleico de (d), y que se una al primer epítipo o a la parte inmunógena del mismo independientemente de la cadena pesada, y cuya cadena pesada se una específicamente al segundo antígeno o epítipo.

17. El método de la reivindicación 16, en el que:

(a) al menos uno de los segmentos génicos  $V_L$  o  $J_L$  de cadena ligera humana, codifica uno o más codones de histidina que no están codificados por un segmento génico variable de cadena ligera de la línea germinal humana correspondiente;

(b) el primer epítipo procede de un receptor de superficie celular, opcionalmente en el que el receptor de superficie celular es un receptor Fc, en el que preferentemente el receptor Fc es FcRn;

(c) el segundo antígeno o epítipo procede de un antígeno soluble;

(d) el segundo antígeno o epítipo procede de un receptor de superficie celular; o

(e) el primer antígeno es un receptor Fc, el segundo antígeno es una proteína soluble y la proteína de unión a antígeno comprende una o más sustituciones e inserciones de histidina procedentes del segmento génico  $V_L$  en el genoma de la rata o del ratón.

18. Una célula procedente de la rata o del ratón de una cualquiera de las reivindicaciones 1-13, en la que la célula comprende en su genoma, un locus de cadena pesada de inmunoglobulina que comprende una secuencia de nucleótidos de región variable de cadena pesada de inmunoglobulina humana reordenada, en la que la secuencia de nucleótidos de región variable de cadena pesada reordenada codifica la secuencia de  $V_H3-23/X_1X_2/J_H$ , en la que  $X_1$  y  $X_2$  es cualquier aminoácido y en la que todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  funcionales endógenos están deletados del locus de cadena pesada de inmunoglobulina endógena o se hacen no funcionales.