

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 612 459**

51 Int. Cl.:

**A01K 67/027** (2006.01)

**C12N 15/85** (2006.01)

**C07K 16/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.08.2011 PCT/US2011/046196**

87 Fecha y número de publicación internacional: **09.02.2012 WO2012018764**

96 Fecha de presentación y número de la solicitud europea: **02.08.2011 E 11741730 (3)**

97 Fecha y número de publicación de la concesión europea: **30.11.2016 EP 2601298**

54 Título: **Ratones que producen proteínas de unión que comprenden dominios VL**

30 Prioridad:

**02.08.2010 US 369909 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**17.05.2017**

73 Titular/es:

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

72 Inventor/es:

**MACDONALD, LYNN;  
STEVENS, SEAN;  
GURER, CAGAN;  
HOSIAWA, KAROLINA, A. y  
MURPHY, ANDREW, J.**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

ES 2 612 459 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

Ratones que producen proteínas de unión que comprenden dominios VL.

### 5 CAMPO DE LA INVENCIÓN

En el presente documento se describen proteínas de unión de tipo inmunoglobulina que comprenden una región constante de cadena pesada de inmunoglobulina fusionada con un dominio variable de cadena ligera de inmunoglobulina, así como proteínas de unión que tienen un dominio variable de cadena ligera de inmunoglobulina fusionado con un dominio constante de cadena ligera y un dominio variable de cadena ligera de inmunoglobulina fusionado con un dominio constante de cadena pesada. También se describen células que expresan tales proteínas de unión, ratones que las producen, y métodos y composiciones relacionados.

### ANTECEDENTES

Los anticuerpos típicamente comprenden una estructura tetramérica que tiene dos cadenas pesadas idénticas, cada una de ellas comprendiendo una región constante de cadena pesada ( $C_H$ ) fusionada con un dominio variable de cadena pesada ( $V_H$ ) asociado con una región constante de cadena ligera ( $C_L$ ) fusionada con un dominio variable de cadena ligera ( $V_L$ ). Una IgG humana típica, que es una molécula de anticuerpo, tiene aproximadamente de 150 kDa a aproximadamente 170 kDa de tamaño (por ejemplo, para IgG3, que comprende una región de bisagra más grande), dependiendo de la subclase de IgG (por ejemplo IgG1, IgG3, IgG4) y de la longitud (variable) de la región variable.

En un anticuerpo típico, los dominios  $V_H$  y  $V_L$  se asocian para formar un sitio de unión que se une a un antígeno diana. Las características del anticuerpo con respecto a la afinidad y especificidad, por lo tanto, pueden depender en gran medida de las características de los dominios  $V_H$  y  $V_L$ . En los anticuerpos típicos en seres humanos y en ratones, los dominios  $V_H$  se acoplan con dominios  $V_L$   $\lambda$  o  $\kappa$ . Sin embargo, también se sabe que se pueden producir dominios  $V_L$  que se unan específicamente a un antígeno diana en ausencia de un dominio  $V_H$  equivalente (por ejemplo, un dominio  $V_H$  que se exprese naturalmente en el contexto de un anticuerpo y esté asociado al dominio  $V_L$  particular), y que se pueden aislar dominios  $V_H$  que se unan específicamente a un antígeno diana en ausencia de un dominio  $V_L$  equivalente. Por lo tanto, la útil diversidad en las proteínas de unión a base de inmunoglobulina, generalmente es conferida por una recombinación que conduce a una  $V_H$  o  $V_L$  particular (e hipermutación somática, en la medida en que ocurra), así como por la combinación de un par  $V_H/V_L$  equivalente. Será útil desarrollar composiciones y métodos para explotar otras fuentes de diversidad.

Existe una necesidad en la técnica de proteínas de unión basadas en estructuras de inmunoglobulina, incluyendo regiones variables de inmunoglobulina tales como las regiones variables de cadena ligera, e incluyendo proteínas de unión que exhiban una mayor diversidad sobre los anticuerpos tradicionales. También existe una necesidad de otros métodos y animales para preparar proteínas de unión útiles, incluyendo proteínas de unión que comprendan las diversas secuencias de región variable de inmunoglobulina de cadena ligera. También existe una necesidad de formatos útiles de proteínas de unión basadas en inmunoglobulina, que proporcionen una mayor diversidad de proteínas de unión a partir de las cuales elegir, y una mayor diversidad de dominios variables de inmunoglobulina, incluyendo composiciones y métodos para generar dominios variables de inmunoglobulina somáticamente mutados y seleccionados por clonación, para su uso, por ejemplo, en la preparación de agentes terapéuticos para seres humanos.

La solicitud WO2009/143472 describe una estrategia con el fin de generar un ratón con inserción génica para generar anticuerpos quiméricos de cadena sencilla que tienen un dominio  $V_L$  somáticamente hipermutado humano, y una región CH de ratón. Tal ratón no se ha obtenido realmente.

La solicitud WO2011/158009 describe la inserción de la región VDJ de IgH humana, y/o las regiones VJ de cadena ligera, en unión operativa en el genoma con una región constante  $\mu$  de una especie no humana. Esta solicitud no describe la generación real de ningún animal no humano que tiene un genoma que codifica una cadena de anticuerpo que tiene una región variable de cadena ligera humana en 5' de una región constante de cadena pesada no humana y *a fortiori*, la expresión real de una proteína de unión a antígeno que comprende un polipéptido que comprende una secuencia de región variable de cadena ligera humana fusionada con una región constante de cadena pesada de inmunoglobulina.

**Resumen**

- La invención proporciona un ratón que comprende en su línea germinal un primer segmento génico ( $V_K$ ) variable de cadena ligera kappa humano no reordenado y un segmento génico ( $J_K$ ) de unión de cadena ligera kappa humano no reordenado unidos operativamente a la región constante ( $C_H$ ) de cadena pesada de ratón endógena en el locus de cadena pesada de ratón endógeno, donde el primer segmento génico  $V_K$  humano no reordenado y el segmento génico  $J_K$  humano no reordenado reemplazan todos los segmentos génicos ( $V_H$ ) variables de cadena pesada de ratón endógenos funcionales, todos los segmentos génicos de diversidad ( $D_H$ ) de ratón endógenos funcionales y todos los segmentos génicos ( $J_H$ ) de unión de cadena pesada de ratón endógenos funcionales, donde el primer segmento génico  $V_K$  humano no reordenado y el segmento génico  $J_K$  humano no reordenado participan en el reordenamiento para formar una secuencia  $V_K/J_K$  reordenada unida operativamente a la región constante de cadena pesada de ratón endógena en el ratón, y donde el ratón comprende adicionalmente en su línea germinal un segundo segmento génico ( $V_L$ ) variable de cadena ligera humano y un segmento génico ( $J_L$ ) J de cadena ligera humano unidos operativamente a un gen ( $C_L$ ) constante de cadena ligera de ratón.
- 15 La invención proporciona adicionalmente el uso de un ratón de la invención para producir una proteína de unión a antígeno que comprende un primer polipéptido que comprende un primer dominio variable de cadena ligera kappa humano fusionado con una región constante de cadena pesada de inmunoglobulina de ratón, y un segundo polipéptido que comprende un segundo dominio variable de cadena ligera humano fusionado con una región constante de cadena ligera de inmunoglobulina de ratón. La proteína de unión a antígeno puede ser un anticuerpo.
- 20 La invención proporciona adicionalmente el uso de un ratón de la invención para producir un hibridoma o un cuadroma para producir un anticuerpo como se ha definido anteriormente.
- La invención también proporciona una célula de ratón que comprende en su genoma un primer segmento génico  $V_K$  humano no reordenado y un segmento génico  $J_K$  humano no reordenado unidos operativamente a la región constante de cadena pesada endógena en el locus de cadena pesada de ratón endógeno, donde el primer segmento génico  $V_K$  humano no reordenado y el segmento génico  $J_K$  humano no reordenado reemplazan todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  endógenos funcionales, y donde la célula de ratón comprende adicionalmente un segundo segmento  $V_L$  humano y un segmento génico  $J_L$  humano unidos operativamente a un gen constante de cadena ligera de ratón.
- 30 La invención proporciona adicionalmente un tejido obtenido a partir de un ratón de la invención, que comprende células como se ha definido anteriormente.
- 35 La invención proporciona adicionalmente un embrión de ratón que comprende una célula que comprende en su genoma un primer segmento génico  $V_K$  humano no reordenado y un segmento génico  $J_K$  humano no reordenado unidos operativamente a la región constante de cadena pesada endógena en el locus de cadena pesada de ratón endógeno, donde el primer segmento génico  $V_K$  humano no reordenado y el segmento génico  $J_K$  humano no reordenado reemplazan todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  endógenos funcionales, y donde la célula comprende adicionalmente un segundo segmento génico  $V_L$  humano y un segmento génico  $J_L$  humano unidos operativamente a un gen constante de cadena ligera de ratón.
- 40 La invención proporciona un método para preparar un ratón modificado genéticamente de la invención, que comprende reemplazar en un locus de cadena pesada endógeno del ratón todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  funcionales del ratón con un primer segmento génico  $V_K$  humano no reordenado y un segmento génico  $J_K$  humano no reordenado, para unir operativamente de esta manera el primer segmento génico  $V_K$  humano no reordenado y el segmento génico  $J_K$  humano no reordenado a la región constante de cadena pesada endógena, y también insertando en la línea germinal del ratón un segundo segmento génico  $V_L$  humano y un segmento génico  $J_L$  humano unidos operativamente a un gen de región constante de cadena ligera de ratón.
- 45 La invención también proporciona una proteína de unión a antígeno obtenible a partir de un ratón de la invención, donde la proteína de unión a antígeno comprende un dominio variable de cadena ligera de inmunoglobulina humana fusionado a un dominio constante de cadena ligera de ratón y un dominio variable de cadena ligera inmunoglobulina kappa humana fusionado a un dominio constante de cadena pesada de ratón.
- 50 La invención proporciona adicionalmente un método para hacer una proteína de unión a antígeno, donde dicho método comprende obtener una secuencia nucleotídica que codifica un dominio  $V_K$  de un gen que codifica un dominio  $V_K$  fusionado a una región  $C_H$  de una célula de un ratón de la invención, clonar la secuencia nucleotídica que codifica el dominio  $V_K$  en marco con un gen que codifica una región  $C_H$  humana para formar una secuencia de

proteína de unión a antígeno humana, y expresar la secuencia de proteína de unión a antígeno humana en una célula adecuada.

5 La invención también proporciona un método para hacer una proteína de unión a antígeno que comprende un dominio  $V_K$  humano, donde dicho método comprende exponer un ratón de la invención a un antígeno de interés, permitir que el ratón desarrolle una respuesta inmune al antígeno de interés, y aislar dicha proteína de unión a antígeno, o aislar el dominio  $V_K$  humano de dicha proteína de unión a antígeno.

10 La invención proporciona adicionalmente un método para obtener una secuencia génica  $V_K$  humana, donde dicho método comprende exponer un ratón de la invención a un antígeno de interés, y aislar de dicho ratón una secuencia génica  $V_K$  humana reordenada, donde la secuencia génica  $V_K$  humana reordenada se fusiona con una secuencia nucleotídica que codifica una región  $C_H$  en dicho ratón.

15 En un aspecto, se describen proteínas de unión que comprenden dominios variables de inmunoglobulina que se derivan de la cadena ligera (es decir, dominios variables de inmunoglobulina kappa ( $\kappa$ ) y/o lambda ( $\lambda$ ), pero no de dominios variables de inmunoglobulina de cadena pesada de longitud completa. También se describen métodos y composiciones para preparar proteínas de unión, incluyendo ratones genéticamente modificados.

20 En un aspecto, se describen construcciones de ácidos nucleicos, células, embriones, ratones y métodos, para preparar proteínas que comprenden uno o más dominios variables de cadena ligera de inmunoglobulina  $\kappa$  y/o  $\lambda$  y una región constante de cadena pesada de inmunoglobulina, incluyendo proteínas que comprendan un dominio variable de cadena ligera  $\lambda$  o  $\kappa$  y una secuencia de región constante de cadena pesada humana o murina.

25 En un aspecto, se describe un ratón que comprende un locus de cadena pesada de inmunoglobulina que comprende un reemplazo de uno o más segmentos génicos variables de cadena pesada ( $V_H$ ) de inmunoglobulina, segmentos génicos de diversidad de cadena pesada ( $D_H$ ) y segmentos génicos de unión de cadena pesada ( $J_H$ ), en un locus de cadena pesada de inmunoglobulina murina endógeno, con uno o más segmentos génicos variables de cadena ligera ( $V_L$ ) y uno o más segmentos génicos de región de unión de cadena ligera ( $J_L$ ).

30 En un aspecto, se describe un ratón que comprende un locus de cadena pesada de inmunoglobulina que comprende un reemplazo de la totalidad o sustancialmente la totalidad de segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  con uno o más segmentos génicos  $V_L$  y uno o más segmentos génicos  $J_L$ , para formar una secuencia de segmento génico  $V_L$  en un locus de cadena pesada endógena (locus  $V_{LH}$ ), donde el locus  $V_{LH}$  es capaz de recombinarse con un gen  $C_H$  murino endógeno, para formar un gen reordenado que se obtiene de un segmento génico  $V_L$ , un segmento génico  $J_L$ , y un gen  $C_H$  de ratón endógeno.

40 En un aspecto, los segmentos génicos  $V_L$  son segmentos génicos  $V_L$  humanos. En un aspecto, los segmentos  $J_L$  son segmentos génicos  $J_L$  humanos. En un aspecto específico, los segmentos génicos  $V_L$  y  $J_L$  son segmentos génicos  $V_L$  humanos y segmentos  $J_L$  humanos.

45 En un aspecto, la totalidad o sustancialmente la totalidad de segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  se reemplazan con al menos seis segmentos génicos  $V_K$  humanos y al menos un segmento génico  $J_K$ . En un aspecto, la totalidad o sustancialmente la totalidad de segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  se reemplazan con al menos 16 segmentos génicos  $V_K$  humanos ( $V_K$  humanos) y al menos un segmento génico  $J_K$ . En un aspecto, la totalidad o sustancialmente la totalidad de segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  se reemplazan con al menos 30 segmentos génicos  $V_K$  humanos y al menos un segmento génico  $J_K$ . En un aspecto, la totalidad o sustancialmente la totalidad de segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  se reemplazan con al menos 40 segmentos génicos  $V_K$  humanos y al menos un segmento génico  $J_K$ . En un aspecto, el al menos un segmento génico  $J_K$  comprende dos, tres, cuatro o cinco segmentos génicos  $J_K$  humanos.

50 En un aspecto, los segmentos génicos  $V_K$  humanos comprenden 4-1, 5-2, 7-3, 2-4, 1-5 y 1-6. En una realización, el segmento  $\kappa V_L$  comprende 3-7, 1-8, 1-9, 2-10, 3-11, 1-12, 1-13, 2-14, 3-15, 1-16. En un aspecto, los segmentos génicos  $V_K$  humanos comprenden 1-17, 2-18, 2-19, 3-20, 6-21, 1-22, 1-23, 2-24, 3-25, 2-26, 1-27, 2-28, 2-29 y 2-30. En un aspecto, los segmentos génicos  $V_K$  humanos comprenden 3-31, 1-32, 1-33, 3-34, 1-35, 2-36, 1-37, 2-38, 1-39 y 2-40.

55 En un aspecto, los segmentos génicos  $V_K$  humanos comprenden 4-1, 5-2, 7-3, 2-4, 1-5, 1-6, 3-7, 1-8, 1-9, 2-10, 3-11, 1-12, 1-13, 2-14, 3-15 y 1-16. En un aspecto, los segmentos génicos  $V_K$  además comprenden 1-17, 2-18, 2-19, 3-20, 6-21, 1-22, 1-23, 2-24, 3-25, 2-26, 1-27, 2-28, 2-29 y 2-30. En una realización, los segmentos génicos  $V_K$

comprenden además 3-31, 1-32, 1-33, 3-34, 1-35, 2-36, 1-37, 2-38, 1-39 y 2-40.

En un aspecto, los segmentos génicos  $V_L$  son segmentos génicos  $V\lambda$  humanos y comprenden un fragmento de un grupo A del locus de cadena ligera  $\lambda$  humana. En un aspecto específico, el fragmento del grupo A del locus de cadena ligera  $\lambda$  humana se extiende desde el  $hV\lambda$ 3-27 hasta el  $hV\lambda$ 3-1.

En un aspecto, los segmentos génicos  $V_L$  comprenden un fragmento del grupo B del locus de cadena ligera  $\lambda$  humana. En un aspecto específico, el fragmento del grupo B del locus de cadena ligera  $\lambda$  humana se extiende desde el  $hV\lambda$ 5-52 hasta  $hV\lambda$ 1-40.

10

En un aspecto, los segmentos génicos  $V_L$  comprenden una secuencia de región variable de cadena ligera  $\lambda$  humana que comprende un fragmento genómico del grupo A y un fragmento genómico del grupo B. En un aspecto, la secuencia de región variable de cadena ligera  $\lambda$  humana comprende al menos un segmento génico del grupo A y al menos un segmento génico del grupo B.

15

En un aspecto, los segmentos génicos  $V_L$  comprenden al menos un segmento génico del grupo B y al menos un segmento génico del grupo C.

En un aspecto, los segmentos génicos  $V_L$  comprenden  $hV\lambda$ , 3-1, 4-3, 2-8, 3-9, 3-10, 2-11 y 3-12. En un aspecto específico, los segmentos génicos  $V_L$  comprenden una secuencia contigua del locus de cadena ligera  $\lambda$  humana que se extiende desde  $V\lambda$ 3-12 hasta  $V\lambda$ 3-1. En un aspecto, la secuencia contigua comprende al menos los segmentos génicos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 o 12  $hV\lambda$ . En un aspecto específico, las  $hV\lambda$ s incluyen 3-1, 4-3, 2-8, 3-9, 3-10, 2-11 y 3-12. En un aspecto específico, las  $hV\lambda$  comprenden una secuencia contigua del locus  $\lambda$  humano que se extiende desde  $V\lambda$ 3-12 hasta  $V\lambda$ 3-1.

25

En un aspecto, las  $hV\lambda$  comprenden 13 a 28 o más  $hV\lambda$ . En un aspecto específico, las  $hV\lambda$  incluyen 2-14, 3-16, 2-18, 3-19, 3-21, 3-22, 2-23, 3-25 y 3-27. En un aspecto específico, las  $hV\lambda$  comprenden una secuencia contigua del locus  $\lambda$  humano, que se extiende desde  $V\lambda$ 3-27 hasta  $V\lambda$ 3-1.

En un aspecto, los segmentos génicos  $V_L$  comprenden 29 a 40  $hV\lambda$ . En un aspecto específico, los segmentos génicos  $V_L$  comprenden una secuencia contigua del locus  $\lambda$  humano que se extiende desde  $V\lambda$ 3-29 hasta  $V\lambda$ 3-1, y una secuencia contigua del locus  $\lambda$  humano que se extiende desde  $V\lambda$ 5-52 hasta  $V\lambda$ 1-40. En un aspecto específico, la totalidad o sustancialmente la totalidad de las secuencias entre  $hV\lambda$ 1-40 y  $hV\lambda$ 3-29 en el ratón genéticamente modificado, consiste básicamente en una secuencia  $\lambda$  humana de aproximadamente 959 pb encontrada en la naturaleza (por ejemplo, en la población humana), en 3' del segmento génico  $hV\lambda$ 1-40 (en 3' de la porción 3' no traducida), que es un sitio de enzima de restricción (por ejemplo PI-SceI), seguido de una secuencia  $\lambda$  humana de aproximadamente 3.431 pb en 5' del segmento génico  $hV\lambda$ 3-29 encontrado en la naturaleza.

35

En un aspecto, el segmento génico  $J_k$  humano se selecciona del grupo que consiste de  $J_k$ 1,  $J_k$ 2,  $J_k$ 3,  $J_k$ 4,  $J_k$ 5, y una combinación de los mismos. En una realización específica, los segmentos génicos  $J_k$  comprenden  $J_k$ 1 a  $J_k$ 5.

40

En un aspecto, los segmentos génicos  $V_L$  son segmentos génicos  $V\lambda$  humanos, y el segmento génico  $J_k$  comprende una RSS que tiene un separador 12-mer, donde la RSS está yuxtapuesta en el extremo 5' del segmento génico  $J_k$ . En un aspecto, los segmentos génicos  $V_L$  son segmentos génicos  $V\lambda$  humanos y el locus  $V_{LH}$  comprende dos o más segmentos génicos  $J_k$ , comprendiendo cada uno una RSS que tiene un separador 12-mer, donde la RSS está yuxtapuesta en el extremo 5' de cada segmento génico  $J_k$ .

45

En una realización específica, los segmentos génicos  $V_k$  comprenden segmentos génicos  $\kappa$  humanos contiguos que abarcan el locus  $\kappa$  humano desde  $V_k$ 4-1 hasta  $V_k$ 2-40, y los segmentos génicos  $J_k$  comprenden segmentos génicos contiguos que abarcan el locus  $\kappa$  humano desde los segmentos génicos  $J_k$ 1 hasta  $J_k$ 5.

50

En un aspecto, cuando los segmentos génicos  $V_L$  son segmentos génicos  $V\lambda$ , y no hay presente ningún segmento génico  $D_H$  entre los segmentos génicos  $V_L$  y segmentos génicos  $J$ , los segmentos génicos  $V_L$  están flanqueados en 3' (es decir, yuxtapuestos en el lado 3') con RSS 23-mer, y los segmentos génicos  $J_k$ , si los hay, o los segmentos génicos  $J\lambda$ , si los hay, están flanqueados en 5' (es decir, yuxtapuestos en el lado 5') con RSS 12-mer.

55

En un aspecto, cuando hay segmentos génicos  $V_{\kappa}$  humanos y no hay presente ningún segmento génico  $D_H$  entre los segmentos génicos  $V_H$  y los segmentos génicos  $J_H$ , cada segmento génico  $V_{\kappa}$  está yuxtapuesto en el lado 3' con una RSS 12-mer, y los segmentos génicos  $J_{\kappa}$ , si los hay, o los segmentos génicos  $J_{\lambda}$ , si los hay, cada uno está yuxtapuesto en el lado 5' con una RSS 23-mer.

5

En un aspecto, el ratón comprende un gen reordenado que se obtiene de un segmento génico  $V_L$ , un segmento génico  $J_L$  y un gen  $C_H$  murino endógeno. En un aspecto, el gen reordenado está somáticamente mutado. En un aspecto, el gen reordenado comprende 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más N adiciones. En un aspecto, las N adiciones y/o las mutaciones somáticas observadas en el gen reordenado derivado del segmento génico  $V_L$  y el segmento génico  $J_L$  superan en un factor de 1,5, 2, 2,5, 3, 3,5, 4, 4,5, o al menos 5, el número de N adiciones y/o mutaciones somáticas que se observan en un dominio variable de cadena ligera reordenado (derivado del mismo segmento génico  $V_L$  y el mismo segmento génico  $J_L$ ), que se redisponen en un locus de cadena ligera endógeno. En un aspecto, el gen reordenado está en una célula B que se une específicamente a un antígeno de interés, donde la célula B se une al antígeno de interés con una  $K_D$  en el rango nanomolar bajo o menor (es decir, una  $K_D$  de 10 nanomolar o menor). En un aspecto, el gen  $C_H$  de ratón se selecciona de IgM, IgD, IgG, IgA e IgE. En un aspecto específico, la IgG de ratón se selecciona de IgG1, IgG2A, IgG2B, IgG2C e IgG3. En otro aspecto específico, la IgG de ratón es IgG1.

En un aspecto, el ratón comprende una célula B, donde la célula B produce a partir de un locus en un cromosoma de la célula B una proteína de unión que consiste esencialmente en cuatro cadenas polipeptídicas, donde las cuatro cadenas polipeptídicas consisten esencialmente en (a) dos polipéptidos idénticos que comprenden una región  $C_H$  de ratón endógena fusionada con un dominio  $V_L$ , y (b) dos polipéptidos idénticos que comprenden una región  $C_L$  de ratón endógena fusionada con un dominio  $V_L$  que es equivalente con respecto al dominio  $V_L$  que está fusionado con la región  $C_H$  de ratón, y, en un aspecto, es un dominio  $V_L$  humano (por ejemplo,  $\kappa$  humana). En un aspecto, el dominio  $V_L$  fusionado con la región  $C_H$  de ratón endógena, es un dominio  $V_L$  humano. En un aspecto específico, el dominio  $V_L$  humano fusionado con la región  $C_H$  de ratón es un dominio  $V_{\kappa}$ . En un aspecto específico, el dominio  $V_L$  humano fusionado con la región  $C_H$  de ratón es idéntico a un dominio V codificado por una secuencia nucleotídica de cadena ligera de línea germinal humana reordenada. En un aspecto específico, el dominio  $V_L$  humano fusionado con la región  $C_H$  de ratón comprende dos, tres, cuatro, cinco, seis o más hipermutaciones somáticas. En un aspecto, el dominio  $V_L$  humano fusionado con la región  $C_H$  de ratón está codificado por un gen reordenado que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9, ó 10 o más N adiciones.

En un aspecto, al menos el 50% de todas las moléculas de IgG producidas por el ratón comprenden un polipéptido que comprende una región  $C_H$  de isotipo IgG y un dominio  $V_L$ , donde la longitud de dicho polipéptido no es mayor de 535, 530, 525, 520 ó 515 aminoácidos. En un aspecto, al menos el 75% de todas las moléculas de IgG comprenden el polipéptido mencionado en este párrafo. En un aspecto, al menos el 80%, 85%, 90% ó 95% de todas las moléculas IgG comprenden el polipéptido mencionado en este párrafo. En un aspecto específico, todas las moléculas IgG producidas por el ratón comprenden un polipéptido que no es más largo que la longitud del polipéptido mencionado en este párrafo.

40

En un aspecto, el ratón produce una proteína de unión que comprende un primer polipéptido que comprende una región  $C_H$  de ratón endógena fusionada con un dominio variable codificado por un segmento génico V humano reordenado y un segmento génico J, pero no un segmento génico  $D_H$ , y un segundo polipéptido que comprende una región  $C_L$  de ratón endógena fusionada con un dominio V codificado por un segmento génico V humano reordenado y un segmento génico J, pero no un segmento génico  $D_H$ , y la proteína de unión se une específicamente a un antígeno con una afinidad en el rango micromolar, nanomolar o picomolar. En un aspecto, el segmento génico J es un segmento génico J humano (por ejemplo, un segmento génico  $\kappa$  humano). En un aspecto, el segmento génico V humano es un segmento génico  $V_{\kappa}$  humano. En un aspecto, el dominio variable que está fusionado con la región  $C_H$  de ratón endógena comprende un mayor número de hipermutaciones somáticas que el dominio variable que está fusionado con la región  $C_L$  de ratón endógena; en un aspecto específico, el dominio variable fusionado con la región  $C_H$  de ratón endógena comprende aproximadamente 1,5, 2, 3, 4 ó 5 veces o más hipermutaciones somáticas que el dominio V fusionado con la región  $C_L$  de ratón endógena; en un aspecto específico, el dominio V fusionado con la región  $C_H$  de ratón comprende al menos 6, 7, 8, 9, 10, 11, 12, 13, 14 ó 15, o más, hipermutaciones somáticas, que el dominio V fusionado con la región  $C_L$  de ratón. En un aspecto, el dominio V fusionado con la región  $C_H$  de ratón está codificado por un gen reordenado que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más N adiciones.

55

En un aspecto, el ratón expresa una proteína de unión que comprende un primer dominio variable de cadena ligera ( $V_L1$ ) fusionado con una secuencia de región constante de cadena pesada de inmunoglobulina, y un segundo dominio variable de cadena ligera ( $V_L2$ ) fusionado con una región constante de cadena ligera de inmunoglobulina,

donde  $V_L1$  comprende un número de hipermutaciones somáticas que es aproximadamente 1,5 a aproximadamente 5 mayor o más, que el número de hipermutaciones somáticas presentes en  $V_L2$ . En un aspecto, el número de hipermutaciones somáticas en  $V_L1$  es aproximadamente 2 a aproximadamente 4 veces mayor que en  $V_L2$ . En un aspecto, el número de hipermutaciones somáticas en el  $V_L1$  es aproximadamente 2 a aproximadamente 3 veces mayor que en  $V_L2$ . En un aspecto,  $V_L1$  está codificado por una secuencia que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más N adiciones.

En un aspecto, se describe un ratón genéticamente modificado que expresa una inmunoglobulina que consiste esencialmente en los siguientes polipéptidos: un primer par de polipéptidos idénticos, donde cada uno consiste esencialmente en una región  $C_H$  fusionada con un dominio variable que se obtiene de segmentos génicos que consisten esencialmente en un segmento génico  $V_L$  y un segmento génico  $J_L$ , y un segundo par de polipéptidos idénticos donde cada uno consiste esencialmente en una región  $C_L$  fusionada con un dominio variable que se obtiene de segmentos génicos que consisten esencialmente en un segmento génico  $V_L$  y un segmento génico  $J_L$ .

15 En un aspecto específico, los dos polipéptidos idénticos que tienen la región  $C_H$ , tienen una región  $C_H$  de ratón.

En un aspecto específico, los dos polipéptidos idénticos que tienen la región  $C_L$  tienen una región  $C_L$  de ratón.

En un aspecto, el dominio variable fusionado con la región  $C_L$  es un dominio variable que es equivalente al dominio variable fusionado con la región  $C_H$ .

En un aspecto, el dominio variable que está fusionado con la región  $C_H$  de ratón endógena comprende un mayor número de hipermutaciones somáticas que el dominio variable que está fusionado con la región  $C_L$  de ratón endógena; en un aspecto específico, el dominio variable fusionado con la región  $C_H$  de ratón endógena comprende aproximadamente 1,5, 2, 2,5, 3, 3,5, 4, 4,5 o 5 veces o más hipermutaciones somáticas, que el dominio variable fusionado con la región  $C_L$  de ratón endógena. En un aspecto, el dominio variable fusionado con la región  $C_L$  de ratón endógena está codificado por un gen que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más N adiciones.

En un aspecto, uno o más de los segmentos génicos V y los segmentos génicos J son segmentos génicos humanos. En un aspecto específico, tanto los segmentos génicos V como los segmentos génicos J son segmentos génicos  $\kappa$  humanos. En otro aspecto específico, tanto los segmentos génicos V como los segmentos génicos J son segmentos génicos  $\lambda$  humanos. En un aspecto, los segmentos génicos V y los segmentos génicos J se seleccionan independientemente de segmentos génicos  $\kappa$  humanos y  $\lambda$  humanos. En un aspecto específico, los segmentos génicos V son segmentos génicos  $V_\kappa$  y los segmentos génicos J son segmentos génicos  $J_\lambda$ . En otro aspecto específico, los segmentos génicos V son segmentos génicos  $V_\lambda$ , y los segmentos génicos J son segmentos génicos  $J_\kappa$ .

En un aspecto, uno o más de los dominios variables fusionados con la región  $C_L$  y los dominios variables fusionados con la región  $C_H$ , son dominios variables humanos. En un aspecto específico, los dominios variables humanos son dominios  $V_\kappa$  humanos. En otro aspecto específico, los dominios variables humanos son dominios  $V_\lambda$ . En un aspecto, los dominios humanos se seleccionan independientemente de dominios  $V_\kappa$  humanos y dominios  $V_\lambda$  humanos. En un aspecto específico, el dominio variable humano fusionado con la región  $C_L$  es un dominio  $V_\lambda$  humano, y el dominio variable humano fusionado con la región  $C_H$  es un dominio  $V_\kappa$  humano. En otro aspecto, el dominio variable humano fusionado con la región  $C_L$  es un dominio  $V_\kappa$  humano y el dominio variable humano fusionado con  $C_H$  es un dominio  $V_\lambda$  humano.

En un aspecto, el dominio  $V_L$  del primer par de polipéptidos idénticos se selecciona de un dominio  $V_\lambda$  humano y un dominio  $V_\kappa$  humano. En un aspecto, el dominio  $V_L$  del segundo par de polipéptidos idénticos se selecciona de un dominio  $V_\lambda$  humano y un dominio  $V_\kappa$  humano. En un aspecto específico, el dominio  $V_L$  del primer par de polipéptidos idénticos es un dominio  $V_\kappa$  humano y el dominio  $V_L$  del segundo par de polipéptidos idénticos se selecciona de un dominio  $V_\kappa$  humano y un dominio  $V_\lambda$  humano. En un aspecto específico, el dominio  $V_L$  del primer par de polipéptidos idénticos es un dominio  $V_\lambda$  humano y el dominio  $V_L$  del segundo par de polipéptidos idénticos se selecciona de un dominio  $V_\lambda$  humano y un dominio  $V_\kappa$  humano. En un aspecto específico, el dominio  $V_L$  humano del primer par de polipéptidos idénticos es un dominio  $V_\kappa$  humano, y el dominio  $V_L$  humano del segundo par de polipéptidos idénticos es un dominio  $V_\kappa$  humano.

En un aspecto, la IgG del ratón comprende una proteína de unión producida en respuesta a un antígeno, donde la proteína de unión comprende un polipéptido que consiste esencialmente en un dominio variable y una región  $C_H$ ,

donde el dominio variable está codificado por una secuencia nucleotídica que consiste esencialmente en un segmento génico  $V_L$  reordenado y un segmento génico J reordenado, y donde la proteína de unión se une específicamente a un epítipo del antígeno, con una  $K_D$  en el rango micromolar, nanomolar o picomolar.

5 En un aspecto, se describe un ratón, donde la totalidad o sustancialmente la totalidad de la IgG producida por el ratón en respuesta a un antígeno, comprende una cadena pesada que comprende un dominio variable, donde el dominio variable está codificado por un gen reordenado derivado de los segmentos génicos que consisten esencialmente en un segmento génico V y un segmento génico J. En un aspecto, el gen reordenado comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más N adiciones.

10

En un aspecto, el segmento génico V es un segmento génico V de una cadena ligera. En un aspecto, la cadena ligera se selecciona de una cadena ligera  $\kappa$  y una cadena ligera  $\lambda$ . En un aspecto específico, la cadena ligera es una cadena ligera  $\kappa$ . En un aspecto específico, el segmento génico V es un segmento génico V humano. En un aspecto específico, el segmento génico V es un segmento génico  $V_\kappa$  humano y el segmento génico J es un segmento génico

15  $J_\kappa$  humano.

En un aspecto, el segmento génico J es un segmento génico J de una cadena ligera. En un aspecto, la cadena ligera se selecciona de una cadena ligera  $\kappa$  y una cadena ligera  $\lambda$ . En un aspecto específico, la cadena ligera es una cadena ligera  $\kappa$ . En un aspecto específico, el segmento génico J es un segmento génico J humano. En otro aspecto, el segmento génico J es un segmento génico J de una cadena pesada (es decir,  $J_H$ ). En un aspecto específico, la

20

cadena pesada es de origen murino. En otro aspecto específico, la cadena pesada es de origen humano.

En un aspecto, el dominio variable de la cadena pesada se produce a partir de no más de un segmento génico V y un segmento génico J es un dominio variable somáticamente mutado.

25

En un aspecto, el dominio variable de la cadena pesada que se produce a partir de no más de un segmento génico V y un segmento génico J está fusionado a una región  $C_H$  de ratón.

En un aspecto específico, la totalidad o sustancialmente la totalidad de la IgG producida por el ratón en respuesta a un antígeno comprende un dominio variable que se obtiene de no más de un segmento génico V humano y no más de un segmento génico J humano, y el dominio variable está fusionado con una región constante de IgG murina, y la IgG además comprende una cadena ligera que comprende un dominio  $V_L$  humano fusionado con una región  $C_L$  de ratón. En un aspecto específico, el dominio  $V_L$  fusionado con la región  $C_L$  de ratón, se deriva de un segmento génico  $V_\kappa$  humano y un segmento génico  $J_\kappa$  humano. En un aspecto específico, el dominio  $V_L$  fusionado con la región  $C_L$  de ratón, se deriva de un segmento génico  $V_\lambda$  humano y un segmento génico  $J_\lambda$  humano.

30

35

En un aspecto, se describe un ratón que produce una IgG que comprende una primera CDR3 en un polipéptido que comprende una región  $C_H$ , y una segunda CDR3 en un polipéptido que comprende una región  $C_L$ , donde tanto la primera CDR3 como la segunda CDR3, se derivan cada una independientemente de no más de dos segmentos génicos, donde los dos segmentos génicos consisten esencialmente en un segmento génico  $V_L$  y un segmento génico  $J_L$ . En un aspecto, la CDR3 en el polipéptido que comprende la región  $C_H$  comprende una secuencia que se obtiene de una secuencia nucleotídica de CDR3 que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más N adiciones.

40

En un aspecto, el segmento génico  $V_L$  y el segmento génico  $J_L$  son segmentos génicos humanos. En un aspecto, el segmento génico  $V_L$  y el segmento génico  $J_L$  son segmentos génicos  $\kappa$ . En un aspecto, el segmento génico  $V_L$  y el segmento génico  $J_L$  son segmentos génicos  $\lambda$ .

45

En un aspecto, se describe un ratón que produce una IgG que comprende una primera CDR3 en un primer polipéptido que comprende una región  $C_H$ , y una segunda CDR3 en un segundo polipéptido que comprende una región  $C_L$ , donde tanto la primera CDR3 como la segunda CDR3 comprenden cada una una secuencia de aminoácidos donde más del 75% de los aminoácidos se derivan de un segmento génico V. En un aspecto, la CDR3 en el polipéptido que comprende la región  $C_H$  comprende una secuencia que se obtiene de una secuencia nucleotídica de CDR3 que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más N adiciones.

50

55 En un aspecto, más del 80%, más del 90%, o más del 95% de los aminoácidos de la primera CDR3; y más del 80%, más del 90%, o más del 95% de los aminoácidos de la segunda CDR3, se derivan de un segmento génico V de cadena ligera.



En un aspecto, no más de dos aminoácidos de la primera CDR3 se derivan de un segmento génico diferente del segmento génico V de cadena ligera. En un aspecto, no más de dos aminoácidos de la segunda CDR3 se derivan de un segmento génico diferente a un segmento génico V de cadena ligera. En un aspecto específico, no más de dos aminoácidos de la primera CDR3 y no más de dos aminoácidos de la segunda CDR3, se derivan de un segmento génico diferente a un segmento génico V de cadena ligera. En un aspecto, ninguna de las CDR3 de la IgG comprende una secuencia de aminoácidos derivada de un segmento génico D. En un aspecto, la CDR3 del primer polipéptido no comprende una secuencia derivada de un segmento génico D.

En un aspecto, el segmento génico V es un segmento génico V humano. En un aspecto específico, el segmento génico V es un segmento génico  $V_{\kappa}$  humano.

En un aspecto, la primera y/o la segunda CDR3 tiene al menos una, dos, tres, cuatro, cinco o seis hipermutaciones somáticas. En un aspecto, la primera CDR3 está codificada por una secuencia de ácido nucleico que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más N adiciones.

En un aspecto, la primera CDR3 consiste esencialmente en aminoácidos derivados de un segmento génico V de cadena ligera humana y un segmento génico J de cadena ligera humana, y la segunda CDR3 consiste esencialmente en aminoácidos derivados de un segmento génico V de cadena ligera humana y un segmento génico J de cadena ligera humana. En un aspecto, la primera CDR3 se deriva de una secuencia de ácido nucleico que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más N adiciones. En un aspecto, la primera CDR3 se deriva de no más de dos segmentos génicos, donde los no más de dos segmentos génicos son un segmento génico  $V_{\kappa}$  humano y un segmento génico  $J_{\kappa}$  humano; y la segunda CDR3 se deriva de no más de dos segmentos génicos, donde los no más de dos segmentos génicos son un segmento génico  $V_{\kappa}$  humano y un segmento génico J que se selecciona de un segmento génico  $J_{\kappa}$  humano, un segmento génico  $J_{\lambda}$  humano, y un segmento génico  $J_H$  humano. En un aspecto, la primera CDR3 se deriva de no más de dos segmentos génicos, donde los no más de dos segmentos génicos son un segmento génico  $V_{\lambda}$  humano y un segmento génico J que se selecciona de un segmento génico  $J_{\kappa}$  humano, un segmento génico  $J_{\lambda}$  humano, y un segmento génico  $J_H$  humano.

En un aspecto, se describe un ratón que produce una IgG que no contiene una secuencia de aminoácidos derivada de un segmento génico  $D_H$ , donde la IgG comprende un primer polipéptido que tiene un primer dominio  $V_L$  fusionado con una región  $C_L$  de ratón, y un segundo polipéptido que tiene un segundo dominio  $V_L$  fusionado con una región  $C_H$  de ratón, donde el primer dominio  $V_L$  y el segundo dominio  $V_L$  no son idénticos. En un aspecto, el primer y segundo dominios  $V_L$  se derivan de diferentes segmentos génicos V. En otro aspecto, el primer y segundo dominios  $V_L$  se derivan de diferentes segmentos génicos J. En un aspecto, el primer y segundo dominios  $V_L$  se derivan de segmentos génicos V y J idénticos, donde el segundo dominio  $V_L$  comprende un número más alto de hipermutaciones somáticas, en comparación con el primer dominio  $V_L$ .

En un aspecto, el primer y segundo dominios  $V_L$  se seleccionan independientemente de dominios  $V_L$  humanos y murinos. En un aspecto, el primer y segundo dominios  $V_L$  se seleccionan independientemente de dominios  $V_{\kappa}$  y  $V_{\lambda}$ . En un aspecto específico, el primer dominio  $V_L$  se selecciona de un dominio  $V_{\kappa}$  y un dominio  $V_{\lambda}$ , y el segundo dominio  $V_L$  es un dominio  $V_{\kappa}$ . En otro aspecto específico, el dominio  $V_{\kappa}$  es un dominio  $V_{\kappa}$  humano.

En un aspecto, se describe un ratón donde la totalidad o sustancialmente la totalidad de todas las IgG producidas por el ratón, consiste esencialmente en una cadena ligera que tiene un primer dominio  $V_L$  humano fusionado con un dominio  $C_L$  murino, y una cadena pesada que tiene un segundo dominio  $V_L$  humano fusionado con un dominio  $C_H$  murino.

En un aspecto, el dominio  $V_L$  humano fusionado con el dominio  $C_H$  murino es un dominio  $V_{\kappa}$  humano.

En un aspecto, el primero y el segundo dominios  $V_L$  humanos no son idénticos.

En un aspecto, se describe un ratón donde al menos el 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% o aproximadamente el 100 % de la inmunoglobulina G producida por el ratón, consiste esencialmente en un dímero de (a) un primer polipéptido que consiste esencialmente en un dominio  $V_L$  de inmunoglobulina y una región  $C_L$  de inmunoglobulina; y (b) un segundo polipéptido de no más de 535 aminoácidos de longitud, donde el segundo polipéptido consiste esencialmente en una región  $C_H$  y un dominio V que carece de una secuencia derivada de un segmento génico  $D_H$ .

En un aspecto, el segundo polipéptido tiene aproximadamente 435-535 aminoácidos de longitud. En un aspecto específico, el segundo polipéptido tiene aproximadamente 435-530 aminoácidos de longitud. En un aspecto específico, el segundo polipéptido tiene aproximadamente 435-525 aminoácidos de longitud. En un aspecto específico, el segundo polipéptido tiene aproximadamente 435-520 aminoácidos de longitud. En un aspecto  
5 específico, el segundo polipéptido tiene aproximadamente 435-515 aminoácidos de longitud.

En un aspecto, en aproximadamente el 90% o más de las IgG producidas por el ratón, el segundo polipéptido no tiene más de aproximadamente 535 aminoácidos de longitud.

10 En un aspecto, en aproximadamente el 50% o más de las IgG producidas por el ratón, el segundo polipéptido no tiene más de aproximadamente 535 aminoácidos de longitud. En un aspecto, en aproximadamente el 50% o más de la inmunoglobulina G producida por el ratón, el segundo polipéptido no tiene más de aproximadamente 530, 525,  
520, 515, 510, 505, 500, 495, 490, 485, 480, 475, 470, 465, 460, 455 ó 450 aminoácidos de longitud. En un aspecto,  
15 aproximadamente el 60%, 70%, 80%, 90% ó 95%, o más de la IgG producida por el ratón, tiene la longitud mencionada. En un aspecto específico, la totalidad o sustancialmente la totalidad de la IgG producida por el ratón, tiene la longitud mencionada.

En un aspecto, el dominio V del segundo polipéptido es un dominio  $V_L$ . En un aspecto específico, el dominio V del segundo polipéptido se selecciona de un dominio  $V_K$  y un dominio  $V_\lambda$ . En un aspecto específico, el dominio V del  
20 segundo polipéptido es un dominio  $V_K$  o  $V_\lambda$  humano.

En un aspecto, se describe un ratón que expresa, a partir de una secuencia nucleotídica en su línea germinal, un polipéptido que comprende una secuencia variable de cadena ligera (por ejemplo, una secuencia V y/o J), una secuencia  $D_H$ , y una región constante de cadena pesada.

25 En un aspecto, el ratón expresa el polipéptido a partir de un locus de cadena pesada de ratón endógeno, que comprende un reemplazo de la totalidad o sustancialmente la totalidad de segmentos génicos del locus variable de cadena pesada de ratón endógeno funcionales, por una pluralidad de segmentos génicos humanos en el locus de cadena pesada de ratón endógeno.

30 En un aspecto, el polipéptido comprende un dominio  $V_L$  derivado de un segmento génico  $V_\lambda$  o  $V_K$ , el polipéptido comprende una CDR3 derivada de un segmento génico  $D_H$ , y el polipéptido comprende una secuencia derivada de un segmento génico  $J_H$  o  $J_\lambda$  o  $J_K$ .

35 En un aspecto, el ratón comprende un locus de inmunoglobulina de cadena pesada murino endógeno que comprende un reemplazo de la totalidad de segmentos génicos  $V_H$  funcionales, por uno o más segmentos génicos  $V_\lambda$  de cadena ligera humana, donde el uno o más segmentos génicos  $V_\lambda$  humanos tiene cada uno yuxtapuesta en el lado 3' una secuencia de señal de recombinación (RSS) separada 23-mer, donde los segmentos génicos  $V_\lambda$  están unidos operativamente a un segmento génico  $D_H$  humano o murino que tiene yuxtapuesta en 5' y 3' una RSS  
40 separada 12-mer; el segmento génico  $D_H$  está unido operativamente con un segmento génico J yuxtapuesto en 5' con una RSS separada 23-mer que es adecuada para recombinarse con la RSS separada 12-mer yuxtapuesta con el segmento génico  $D_H$ , donde los segmentos génicos V,  $D_H$  y J están unidos operativamente a una secuencia de ácido nucleico que codifica una región constante de cadena pesada.

45 En un aspecto, el ratón comprende un locus de inmunoglobulina de cadena pesada de ratón endógeno que comprende un reemplazo de la totalidad de segmentos génicos  $V_H$  funcionales, por uno o más segmentos génicos  $V_K$  humanos, cada uno yuxtapuesto en el lado 3' con una secuencia de señal de recombinación (RSS) separada 12-mer, donde los segmentos génicos V están unidos operativamente a un segmento génico  $D_H$  humano o murino que está yuxtapuesto tanto en 5' como en 3' con una RSS separada 23-mer, el segmento génico  $D_H$  está unido  
50 operativamente con un segmento génico J yuxtapuesto en el lado 5' con una RSS separada 12-mer, que es adecuada para recombinarse con la RSS separada 23-mer yuxtapuesta con el segmento génico  $D_H$ , donde los segmentos génicos V,  $D_H$  y J están unidos operativamente a una secuencia de ácido nucleico que codifica una región constante de cadena pesada.

55 En un aspecto, la región constante de cadena pesada es una región constante de cadena pesada de ratón endógena. En un aspecto, la secuencia de ácido nucleico codifica un dominio que se selecciona de un  $C_{H1}$ , una bisagra, un  $C_{H2}$ , un  $C_{H3}$ , y una combinación de los mismos. En un aspecto, uno o más de los  $C_{H1}$ , bisagra,  $C_{H2}$  y  $C_{H3}$ , son humanos.

En un aspecto, el ratón comprende un locus de inmunoglobulina de cadena pesada de ratón endógeno que comprende un reemplazo de la totalidad de los segmentos génicos  $V_H$  funcionales, por una pluralidad de segmentos génicos  $V\lambda$  o  $V\kappa$  humanos, cada uno yuxtapuesto en 3' con una RSS separada 23-mer, una pluralidad de segmentos génicos  $D_H$  humanos yuxtapuestos tanto en 5' como en 3' con una RSS separada 12-mer, una pluralidad de segmentos génicos  $J$  humanos ( $J_H$  o  $J\lambda$  o  $J\kappa$ ) yuxtapuestos tanto en 5' como en 3' con una RSS separada 23-mer, donde el locus comprende una secuencia de región constante murina endógena que se selecciona de  $C_H1$ , bisagra,  $C_H2$ ,  $C_H3$ , y una combinación de las mismas. En un aspecto específico, el ratón comprende la totalidad o sustancialmente la totalidad de segmentos génicos  $V\lambda$  o  $V\kappa$  humanos funcionales, la totalidad o sustancialmente la totalidad de segmentos génicos  $D_H$  humanos, y la totalidad o sustancialmente la totalidad de segmentos génicos  $J_H$  o  $J\lambda$  o  $J\kappa$ .

En un aspecto, el ratón expresa una proteína de unión al antígeno que comprende: (a) un polipéptido que comprende un dominio de cadena ligera humana unido a una región constante de cadena pesada que comprende una secuencia murina; y (b) un polipéptido que comprende un dominio variable de cadena ligera humana unido a una región constante de cadena ligera humana o murina. En un aspecto específico, la secuencia de cadena ligera es una secuencia de cadena ligera humana, y al exponerse a una proteasa que sea capaz de escindir un anticuerpo en un Fc y un Fab, se forma un Fab completamente humano que comprende al menos cuatro CDR de cadena ligera, donde las al menos cuatro CDR de cadena ligera se seleccionan de secuencias  $\lambda$ , secuencias  $\kappa$ , y una combinación de las mismas. En un aspecto, el Fab comprende al menos cinco CDR de cadena ligera. En un aspecto, el Fab comprende seis CDR de cadena ligera. En un aspecto, al menos una CDR del Fab comprende una secuencia derivada de un segmento génico  $V\lambda$  o un segmento génico  $V\kappa$ , y la al menos una CDR además comprende una secuencia derivada de un segmento génico  $D$ . En un aspecto, la al menos una CDR es una CDR3 y la CDR se deriva de un segmento génico  $V\kappa$  humano, un segmento génico  $D$  humano, y un segmento génico  $J\kappa$  humano.

En un aspecto, el polipéptido comprende un dominio variable derivado de un segmento génico  $V\lambda$  o  $V\kappa$  humano, un segmento génico  $D_H$  humano, y un segmento génico  $J_H$  o  $J\lambda$  o  $J\kappa$  humano. En un aspecto específico, la región constante de cadena pesada se deriva de una secuencia  $C_H1$  humana y una secuencia  $C_H2$  murina y una secuencia  $C_H3$  murina.

En un aspecto, se describe un ratón que comprende en su línea germinal un segmento génico  $V\kappa$  o  $V\lambda$ , humano no reordenado, unido operativamente a un segmento génico  $J$  humano y una región constante de cadena pesada, donde el ratón expresa una proteína de unión  $V_L$  que comprende un dominio  $V\kappa$  humano fusionado con una región constante de cadena pesada, y donde el ratón exhibe una población de células B esplénicas que expresa proteínas de unión  $V_L$  en células B  $CD19^+$ , incluyendo células B de transición  $CD19^+IgM^{hi}IgD^{int}$ , y células B maduras ( $CD19^+IgM^{int}IgD^{hi}$ ).

En un aspecto, se describe un ratón que comprende en su línea germinal, un segmento génico  $V\kappa$  o  $V\lambda$ , humano no reordenado, unido operativamente a un segmento génico  $J$  humano y una región constante de cadena pesada, donde el ratón expresa en una célula B, una inmunoglobulina que comprende un dominio variable de cadena ligera fusionado con una región constante de cadena pesada, donde la población linfocítica de la médula ósea del ratón, exhibe una población de células pro/pre B que es aproximadamente la misma en número que una población de células pro/pre B de un ratón de tipo silvestre (linfocitos en la médula ósea).

En un aspecto, el ratón comprende al menos 6 segmentos génicos  $hV\kappa$  no reordenados y uno o más segmentos génicos  $hJ\kappa$  no reordenados, y los ratones comprenden una población de células esplénicas activadas por linfocitos e  $IgM^+$  que expresan una proteína de unión  $V_L$ , donde la población es al menos 75% tan grande como una población de células esplénicas activadas por linfocitos e  $IgM^+$  de un ratón de tipo silvestre.

En un aspecto, los ratones exhiben una población de esplenocitos activados por células B maduras ( $CD19^+$ ) de células  $IgD^+$  e  $IgM^+$ , que en total constituyen aproximadamente el 90%; en una realización, la población de esplenocitos activados por células B maduras ( $CD19^+$ ) de células  $IgD^+$  e  $IgM^+$  del ratón modificado, es aproximadamente igual (por ejemplo, dentro del 10% o dentro del 5%) que el total de células  $IgD^+$  y células  $IgM^+$  de un ratón de tipo silvestre que son esplenocitos activados por células B maduras ( $CD19^+$ ).

En un aspecto, se describe un ratón que expresa una proteína inmunoglobulina a partir de un locus de cadena pesada endógeno modificado en su línea germinal, donde el locus de cadena pesada endógeno modificado carece de un segmento génico  $V$  de cadena pesada de ratón funcional y el locus comprende segmentos génicos  $V$  de cadena ligera no reordenados y segmentos génicos  $J$  no reordenados, donde los segmentos génicos  $V$  de cadena

ligera no reordenados y los segmentos génicos J no reordenados, están unidos operativamente a una región constante de cadena pesada; donde la proteína inmunoglobulina consiste esencialmente en un primer polipéptido y un segundo polipéptido, donde el primer polipéptido comprende un dominio de cadena ligera de inmunoglobulina y una región constante de cadena pesada de inmunoglobulina, y el segundo polipéptido comprende un dominio variable de cadena ligera de inmunoglobulina y una región constante de cadena ligera.

En un aspecto, se describe un ratón que expresa una proteína inmunoglobulina, donde la proteína inmunoglobulina carece de un dominio variable de inmunoglobulina de cadena pesada, y la proteína inmunoglobulina comprende un primer dominio variable derivado de segmentos génicos de cadena ligera, y un segundo dominio variable derivado de segmentos génicos de cadena ligera, donde el primer dominio variable y el segundo dominio variable son equivalentes uno con respecto al otro, donde el primer y segundo dominios variables de cadena ligera no son idénticos, y donde el primer y segundo dominios variables de cadena ligera se asocian, y cuando están asociados, se unen específicamente al antígeno de interés.

En un aspecto, se describe un ratón que produce a partir de segmentos génicos no reordenados en su línea germinal, una proteína inmunoglobulina que comprende dominios variables que se derivan por completo de segmentos génicos que consisten esencialmente en segmentos génicos humanos no reordenados, donde la proteína inmunoglobulina comprende una región constante de cadena ligera de inmunoglobulina y una región constante de cadena pesada de inmunoglobulina que se selecciona del grupo que consiste en C<sub>H</sub>1, una bisagra, C<sub>H</sub>2, C<sub>H</sub>3, y una combinación de las mismas.

En un aspecto, se describe un ratón que produce, a partir de segmentos génicos no reordenados en su línea germinal, una proteína inmunoglobulina que comprende dominios variables, donde todas las CDR3 de todas las regiones variables son generadas por completo a partir de segmentos génicos V y J de cadena ligera, y opcionalmente una o más hipermutaciones somáticas, por ejemplo una o más N adiciones.

En un aspecto, se describe un ratón que produce una proteína inmunoglobulina somáticamente mutada derivada de segmentos génicos de región variable de cadena ligera de inmunoglobulina humana no reordenados, en la línea germinal del ratón, donde la proteína inmunoglobulina carece de una CDR que comprenda una secuencia derivada de un segmento génico D, donde la proteína inmunoglobulina comprende una primera CDR3 en un dominio variable de cadena ligera fusionado con una región constante de cadena ligera, comprende una segunda CDR3 en un dominio variable de cadena ligera fusionado con una región constante de cadena pesada, y donde la segunda CDR3 se deriva de la secuencia de región variable de cadena ligera reordenada que comprende 1, 2, 3, 4, 5, 6, 7, 9 ó 10, o más N adiciones.

En un aspecto, se proporciona un ratón como se describe en el presente documento, donde el ratón comprende un locus de cadena ligera funcionalmente silenciado que se selecciona de un locus  $\lambda$ , un locus  $\kappa$ , y una combinación de los mismos. En un aspecto, el ratón comprende una delección de un locus  $\lambda$  y/o  $\kappa$ , en su totalidad o en parte, de tal modo que el locus  $\lambda$  y/o  $\kappa$  no es funcional.

En un aspecto, se describe un embrión de ratón que comprende una célula que comprende un locus de inmunoglobulina modificado como se describe en el presente documento. En un aspecto, el ratón es una quimera y al menos el 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ó 95% de las células del embrión comprenden un locus de inmunoglobulina modificado como se describe en el presente documento. En un aspecto, al menos el 96%, 97%, 98%, 99% ó 99,8% de las células del embrión, comprenden un locus de inmunoglobulina modificado como se describe en el presente documento. En un aspecto, el embrión comprende una célula huésped y una célula derivada de una célula ME donante, donde la célula derivada de la célula ME donante comprende un locus de inmunoglobulina modificado como se describe en el presente documento. En un aspecto, el embrión es un embrión huésped en el estadio de 2, 4, 8, 16, 32 ó 64 células, o un blastocisto, y además comprende una célula ME donante que comprende un locus de inmunoglobulina modificado como se describe en el presente documento.

En un aspecto, se proporciona un ratón o una célula obtenida utilizando una construcción de ácido nucleico como se describe en el presente documento.

En un aspecto, se proporciona un ratón obtenido utilizando una célula como se describe en el presente documento. En un aspecto, la célula es una célula ME murina.

En un aspecto, uso de un ratón como se describe en el presente documento, para preparar una secuencia de ácido nucleico que codifica una primera secuencia variable de inmunoglobulina de cadena ligera humana (V<sub>L</sub>1), que es

equivalente a una segunda secuencia variable de inmunoglobulina de cadena ligera humana ( $V_L2$ ), donde la  $V_L1$  fusionada con una región constante de cadena ligera de inmunoglobulina humana (polipéptido 1) se expresa con  $V_L2$  fusionada con una región constante de cadena pesada de inmunoglobulina humana (polipéptido 2), en forma de un dímero del polipéptido 1/polipéptido 2, para formar un anticuerpo  $V_L1-V_L2$ .

5

En un aspecto, uso de un ratón como se describe en el presente documento para preparar una secuencia de ácido nucleico que codifica una secuencia variable de cadena ligera de inmunoglobulina humana que está fusionada con una secuencia de cadena pesada de inmunoglobulina humana, donde la secuencia de ácido nucleico codifica un polipéptido  $V_L-C_H$  humano, donde el polipéptido  $V_L-C_H$  humano se expresa en forma de un dímero, y donde el

10 dímero se expresa en ausencia de una cadena ligera de inmunoglobulina (por ejemplo, en ausencia de una cadena ligera  $\lambda$  humana o  $\kappa$  humana). En un aspecto, el dímero  $V_L-C_H$  se une específicamente a un antígeno de interés, en ausencia de una cadena ligera  $\lambda$  y en ausencia de una cadena ligera  $\kappa$ .

En un aspecto, uso de un ratón como se describe en el presente documento, para preparar una secuencia de ácido nucleico que codifica la totalidad o una porción de un dominio variable de inmunoglobulina. En un aspecto, el

15 dominio variable de inmunoglobulina es un dominio  $V_\lambda$ , humano o  $V_\kappa$  humano.

En un aspecto, se describe el uso de un ratón como se describe en el presente documento para preparar un Fab completamente humano (que comprende una primera  $V_L$  humana fusionada con una región constante de cadena pesada humana, y una segunda  $V_L$  humana fusionada con una región constante de cadena pesada humana), o un

20  $F(ab)_2$  completamente humano.

En un aspecto, se describe el uso de un ratón como se describe en el presente documento, para preparar una línea celular inmortalizada. En un aspecto, la línea celular inmortalizada comprende una secuencia de ácido nucleico que

25 codifica un dominio  $V_\lambda$  o  $V_\kappa$  humano, unido operativamente a una secuencia de ácido nucleico que comprende una secuencia de ácido nucleico de región constante murina.

En un aspecto, se describe el uso de un ratón como se describe en el presente documento, para preparar un

30 hibridoma o un cuadro.

30

En un aspecto, se describe una célula que comprende un locus de inmunoglobulina modificado como se describe en el presente documento. En un aspecto, la célula se selecciona de una célula totipotente, una célula pluripotente, una célula madre pluripotente inducida (CMPi), y una célula ME. En un aspecto específico, la célula es una célula de

35 ratón, por ejemplo una célula ME murina. En un aspecto, la célula es homocigota para el locus de inmunoglobulina modificado.

En un aspecto, se describe una célula que comprende una secuencia de ácido nucleico que codifica un primer polipéptido que comprende un primer dominio  $V_\kappa$  o  $V_\lambda$  humano somáticamente mutado, fusionado con un gen de

40 región constante de cadena pesada humana.

40

En un aspecto, la célula además comprende una segunda cadena polipeptídica que comprende un segundo dominio  $V_\kappa$  o  $V_\lambda$  somáticamente mutado, fusionado con una región constante de cadena ligera humana.

En un aspecto, el dominio  $V_\kappa$  o  $V_\lambda$  humano del primer polipéptido es equivalente al dominio  $V_\kappa$  o  $V_\lambda$  humano del

45 segundo polipéptido.

En un aspecto,  $V_\kappa$  o  $V_\lambda$  del primer polipéptido y  $V_\kappa$  o  $V_\lambda$  humano del segundo polipéptido, cuando se asocian, se unen específicamente a un antígeno de interés. En un aspecto específico, el primer polipéptido comprende un

50 dominio variable que consiste esencialmente en un dominio  $V_\kappa$  humano, y el segundo polipéptido comprende un dominio variable que consiste en un dominio  $V_\kappa$  humano que es equivalente al dominio  $V_\kappa$  humano del primer polipéptido, y la región constante humana, es una secuencia de IgG.

En un aspecto, la célula se selecciona de una célula CHO, una célula COS, una célula 293, una célula HeLa y una

55 célula retiniana humana que expresa una secuencia de ácido nucleico viral (por ejemplo, una célula PERC.6™).

55

En un aspecto, se describe una célula somática de ratón que comprende un cromosoma que comprende una modificación genética como se describe en el presente documento.

En un aspecto, se describe una célula germinal de ratón, que comprende una secuencia de ácido nucleico que comprende una modificación genética como se describe en el presente documento.

5 En un aspecto, se describe una célula pluripotente, pluripotente inducida o totipotente, derivada de un ratón como se describe en el presente documento. En un aspecto específico, la célula es una célula madre embrionaria (ME) murina.

10 En un aspecto, se describe el uso de una célula como se describe en el presente documento, para la fabricación de un ratón, una célula, o una proteína terapéutica (por ejemplo, un anticuerpo u otra proteína de unión al antígeno).

15 En un aspecto, se describe una construcción de ácido nucleico que comprende un segmento génico  $D_H$  humano yuxtapuesto en 5' y 3' con una RSS separada 23-mer. En un aspecto específico, la construcción de ácido nucleico comprende un brazo de homología que es homólogo a una secuencia genómica humana que comprende segmentos génicos  $V_K$  humanos. En un aspecto, la construcción de direccionamiento comprende la totalidad o sustancialmente la totalidad de los segmentos génicos  $D_H$  humanos, cada uno yuxtapuesto en 5' y 3' con una RSS separada 23-mer.

20 En un aspecto, se describe una construcción de ácido nucleico que comprende un segmento génico  $J_K$  humano yuxtapuesto en 5' con una RSS separada 12-mer. En un aspecto específico, la construcción de ácido nucleico comprende un primer brazo de homología que contiene homología con una secuencia génica  $D_H$  genómica humana que está yuxtapuesta en 5' y 3' con una RSS separada 23-mer. En un aspecto, la construcción de ácido nucleico comprende un segundo brazo de homología que contiene homología con una secuencia génica  $J$  genómica humana, o que contiene homología con una secuencia de región constante de cadena pesada de ratón, o que contiene homología con una secuencia intergénica J-C que se encuentra en 5' de una secuencia de cadena pesada de región constante murina.

25 En un aspecto, se describe una construcción de ácido nucleico que comprende un segmento génico  $V_L$  yuxtapuesto en 3' con una RSS separada 23-mer, un segmento génico  $D_H$  humano yuxtapuesto en 3' y 5' con una RSS separada 12-mer, y un segmento génico  $J$  humano que se selecciona de un segmento génico  $J_K$  yuxtapuesto en 5' con una RSS separada 23-mer, un segmento génico  $J_L$  humano yuxtapuesto en 5' con una RSS separada 23-mer, y un segmento génico  $J_H$  humano yuxtapuesto en 5' con una RSS separada 23-mer. En un aspecto, la construcción comprende un brazo de homología que contiene homología con una secuencia de región constante murina, una secuencia intergénica J-C murina, y/o una secuencia  $V_L$  humana.

30 En un aspecto, la construcción de ácido nucleico comprende una región variable de cadena ligera  $\lambda$  humana que comprende un fragmento del grupo A del locus de la cadena ligera  $\lambda$  humana. En un aspecto específico, el fragmento del grupo A del locus de cadena ligera  $\lambda$  humana, se extiende desde  $hV\lambda.3-27$  a  $hV\lambda.3-1$ .

35 En un aspecto, la construcción de ácido nucleico comprende una región variable de cadena ligera  $\lambda$  humana que comprende un fragmento del grupo B del locus de cadena ligera  $\lambda$  humana. En un aspecto específico, el fragmento del grupo B del locus de cadena ligera  $\lambda$  humana se extiende desde  $hV\lambda.5-52$  a  $hV\lambda.1-40$ .

40 En un aspecto, la construcción de ácido nucleico comprende una región variable de cadena ligera  $\lambda$  humana que comprende un fragmento genómico del grupo A y un fragmento genómico del grupo B. En un aspecto, la región variable de cadena ligera  $\lambda$  humana comprende al menos un segmento génico del grupo A y al menos un segmento génico del grupo B.

45 En un aspecto, la región variable de cadena ligera  $\lambda$  humana comprende al menos un segmento génico del grupo B y al menos un segmento génico del grupo C.

50 En un aspecto, se describe una construcción de ácido nucleico que comprende un segmento génico  $D_H$  humano yuxtapuesto en 5' y 3' con una RSS separada 23-mer, normalmente encontrada en la naturaleza flanqueando un segmento génico  $J_K$ ,  $J_H$ ,  $V_L$  o  $V_H$ . En un aspecto, la construcción de ácido nucleico comprende un primer brazo de homología que es homólogo a una región intergénica V-J humana, o es homólogo a una secuencia genómica humana que comprende un segmento génico  $V$  humano. En un aspecto, la construcción de ácido nucleico comprende un segundo brazo de homología que es homólogo a una secuencia de región constante de cadena pesada humana o murina. En un aspecto específico, la región constante de cadena pesada humana o murina se selecciona de  $C_H1$ , bisagra,  $C_H2$ ,  $C_H3$ , y una combinación de las mismas. En un aspecto, la construcción de ácido nucleico comprende un segmento génico  $J$  humano flanqueado en 5' por una RSS separada 12-mer. En un aspecto,

la construcción de ácido nucleico comprende un segundo brazo de homología que contiene homología con un segmento génico J flanqueado en 5' por una RSS separada 12-mer. En un aspecto, el segmento génico J se selecciona de un J $\kappa$  humano, J $\lambda$  humano, y J $\mu$  humano.

- 5 En un aspecto, se describe una construcción de ácido nucleico que comprende un segmento génico D $\mu$  humano yuxtapuesto en 5' y 3' con una RSS separada 23-mer, y una secuencia de reconocimiento de recombinasa específica de sitio, por ejemplo una secuencia reconocida por una recombinasa específica de sitio, tal como una proteína Cre, Flp o Dre.
- 10 En un aspecto, se describe una construcción de ácido nucleico que comprende un segmento génico V $\lambda$  humano o V $\kappa$  humano, un segmento génico D $\mu$  yuxtapuesto en 5' y 3' con una RSS separada 12-mer ó 23-mer, y un segmento génico J humano con una RSS separada 12-mer ó 23-mer, donde la RSS separada 12-mer ó 23-mer está situada inmediatamente en 5' respecto al segmento génico J humano (es decir, con respecto a la dirección de la transcripción). En un aspecto, la construcción comprende un segmento génico V $\lambda$  humano yuxtapuesto con una RSS separada 23-mer en 3', un segmento génico D $\mu$  humano yuxtapuesto en 5' y 3' con una RSS separada 12-mer, y un segmento génico J $\kappa$  humano yuxtapuesto en 5' con una RSS separada 23-mer. En un aspecto, la construcción comprende un segmento génico V $\kappa$  humano yuxtapuesto en 3' con una RSS separada 12-mer, un segmento génico D $\mu$  humano yuxtapuesto en 5' y 3' con una RSS separada 23-mer y un segmento génico J $\lambda$  humano yuxtapuesto con una RSS separada 12-mer en 5'.
- 15
- 20 En un aspecto, se describe un vector de direccionamiento, que comprende: (a) un primer brazo de direccionamiento y un segundo brazo de direccionamiento, donde el primer y segundo brazos de direccionamiento se seleccionan independientemente de brazos de direccionamiento humanos y murinos, donde los brazos de direccionamiento dirigen al vector hacia un locus génico de la región V de inmunoglobulina endógeno o modificado; y (b) una
- 25 secuencia contigua de segmentos génicos V $\mu$  humanos o una secuencia contigua de segmentos génicos V $\mu$  humanos y al menos un segmento génico J $\kappa$  humano, donde la secuencia contigua se selecciona del grupo que consiste en (i) hV $\kappa$ 4-1 a hV $\kappa$ 1-6 y J $\kappa$ 1, (ii) hV $\kappa$ 4-1 a hV $\kappa$ 1-6 y J $\kappa$ 1 a J $\kappa$ 2, (iii) hV $\kappa$ 4-1 a hV $\kappa$ 1-6 y J $\kappa$ 1 a J $\kappa$ 3, (iv) hV $\kappa$ 4-1 a hV $\kappa$ 1-6 y J $\kappa$ 1 a J $\kappa$ 4, (v) hV $\kappa$ 4-1 a hV $\kappa$ 1-6 y J $\kappa$ 1 a J $\kappa$ 5, (vi) hV $\kappa$ 3-7 a hV $\kappa$ 1-16, (vii) hV $\kappa$ 1-17 a hV $\kappa$ 2-30, (viii) hV $\kappa$ 3-31 a hV $\kappa$ 2-40, y (ix) una combinación de los mismos.
- 30
- En un aspecto, los brazos de direccionamiento que dirigen el vector hacia un locus de inmunoglobulina endógeno o modificado, son idénticos o sustancialmente idénticos a una secuencia en el locus de inmunoglobulina endógeno o modificado.
- 35 En un aspecto, se describe el uso de una construcción de ácido nucleico como se describe en el presente documento, para la fabricación de un ratón, una célula o una proteína terapéutica (por ejemplo, un anticuerpo u otra proteína de unión a un antígeno).
- En un aspecto, se describe el uso de una secuencia de ácido nucleico de un ratón, como se describe en el presente
- 40 documento, para preparar una línea celular para la fabricación de un agente terapéutico humano. En un aspecto, el agente terapéutico humano es una proteína de unión que comprende un dominio variable de cadena ligera humana (por ejemplo, derivado de un segmento génico V $\lambda$  humano o V $\kappa$  humano), fusionado con una región constante de cadena pesada humana. En un aspecto, el agente terapéutico humano comprende un primer polipéptido que es una cadena ligera de inmunoglobulina  $\lambda$  o  $\kappa$  humana, y un segundo polipéptido que comprende un dominio variable V $\lambda$
- 45 humano o V $\kappa$  humano fusionado con una región constante de cadena pesada humana.
- En un aspecto, se describe un sistema de expresión que comprende una célula de mamífero transfectada con una construcción de ADN que codifica un polipéptido que comprende un dominio V $\mu$  humano somáticamente mutado, fusionado con un dominio C $\mu$  humano.
- 50
- En un aspecto, el sistema de expresión comprende además una secuencia nucleotídica que codifica un dominio V $\mu$  de inmunoglobulina fusionado con un dominio C $\mu$  humano, donde el dominio V $\mu$  fusionado con el dominio C $\mu$  humano es una cadena ligera equivalente al dominio V $\mu$  fusionado con el dominio C $\mu$  humano.
- 55 En un aspecto, la célula de mamífero se selecciona de una célula CHO, una célula COS, una célula Vero, una célula 293, y una célula retiniana que expresa un gen viral (por ejemplo, una célula PER.C6<sup>TM</sup>).

En un aspecto, se describe un método para preparar una proteína de unión, que comprende obtener una secuencia

nucleotídica que codifica un dominio  $V_L$  a partir de un gen que codifica un dominio  $V_L$  fusionado con una región  $C_H$  de una célula de un ratón como se describe en el presente documento, y clonar la secuencia nucleotídica que codifica el dominio  $V_L$  en marco con un gen que codifica una región  $C_H$  humana para formar una secuencia de proteína de unión humana que expresa la proteína de unión humana en una célula adecuada.

5

En un aspecto, el ratón se ha inmunizado con un antígeno de interés, y el dominio  $V_L$  fusionado con la región  $C_H$  se une específicamente (por ejemplo, con una  $K_D$  en el rango micromolar, nanomolar o picomolar) a un epítipo del antígeno de interés. En un aspecto, la secuencia nucleotídica que codifica el dominio  $V_L$  fusionado con la región  $C_H$ , está somáticamente mutada en el ratón.

10

En un aspecto, la célula adecuada se selecciona de una célula B, un hibridoma, un cuadroma, una célula CHO, una célula COS, una célula 293, una célula HeLa y una célula retiniana humana que expresa una secuencia de ácido nucleico viral (por ejemplo, una célula PERC.6™).

15 En un aspecto, la región  $C_H$  comprende un isotipo IgG humano. En un aspecto específico, la IgG humana se selecciona de IgG1, IgG2 e IgG4. En otro aspecto específico, la IgG humana es la IgG 1. En otro aspecto específico, la IgG humana es la IgG4. En otro aspecto específico, la IgG4 humana es una IgG4 modificada. En un aspecto, la IgG4 modificada comprende una sustitución en la región de bisagra. En un aspecto específico, la IgG4 modificada comprende una sustitución en el residuo de aminoácido 228 en relación con una IgG4 humana de tipo silvestre,  
 20 numerada de acuerdo con el índice de numeración de la UE de Kabat. En un aspecto específico, la sustitución en el residuo de aminoácido 228 es una sustitución S228P, numerada de acuerdo con el índice de numeración de la UE de Kabat.

En un aspecto, la célula además comprende una secuencia nucleotídica que codifica un dominio  $V_L$  a partir de una  
 25 cadena ligera que es equivalente al dominio  $V_L$  fusionado con la región  $C_H$ , y el método además comprende expresar la secuencia nucleotídica que codifica para el dominio  $V_L$  equivalente fusionado con un dominio  $C_K$  o  $C_L$  humano.

En un aspecto, se describe un método para preparar un ratón genéticamente modificado, que comprende  
 30 reemplazar, en un locus de cadena pesada de ratón endógeno, uno o más segmentos génicos de cadena pesada de inmunoglobulina de un ratón, con uno o más segmentos génicos de cadena ligera de inmunoglobulina humanos. En un aspecto, el reemplazo es de la totalidad o sustancialmente la totalidad de segmentos génicos de cadena pesada de inmunoglobulina murina funcionales (es decir, los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$ ), por uno o más segmentos génicos de cadena ligera humana funcionales (es decir, segmentos  $V_L$  y  $J_L$ ). En un aspecto, el reemplazo es de la  
 35 totalidad o sustancialmente la totalidad de segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  de cadena pesada de ratón funcionales por la totalidad o sustancialmente la totalidad de segmentos génicos  $V_L$  o  $V_K$  humanos y al menos un segmento génico  $J_L$  o  $J_K$ . En un aspecto específico, el reemplazo incluye la totalidad o sustancialmente la totalidad de segmentos génicos  $J_L$  o  $J_K$  humanos funcionales.

40 En un aspecto, se describe un método para producir un ratón que expresa un polipéptido que comprende una secuencia derivada de un segmento génico  $V_L$  o  $V_K$  y/o  $J_L$  o  $J_K$  de inmunoglobulina humana fusionado con una región constante de cadena pesada de ratón, que comprende reemplazar los segmentos génicos variables de inmunoglobulina de cadena pesada de ratón endógenos ( $V_H$ ,  $D_H$  y  $J_H$ ) con al menos un segmento génico  $V_L$  o  $V_K$  humano, y al menos un segmento génico  $J_L$  o  $J_K$  humano, donde el reemplazo es en una célula de ratón  
 45 pluripotente, pluripotente inducida, o totipotente, para formar una célula progenitora murina genéticamente modificada; la célula progenitora murina genéticamente modificada se introduce en un ratón huésped; y el ratón huésped que comprende la célula progenitora genéticamente modificada, se somete a gestación para formar un ratón que comprende un genoma derivado de la célula progenitora de ratón genéticamente modificada. En un aspecto, el huésped es un embrión. En un aspecto específico, el huésped se selecciona de una premórula murina  
 50 (por ejemplo, estadio de 8 ó 4 células), un embrión tetraploide, un agregado de células embrionarias, o un blastocisto.

En un aspecto, se describe un método para preparar un ratón genéticamente modificado como se describe en el presente documento, que comprende introducir, por transferencia nuclear, un ácido nucleico que contiene una  
 55 modificación como se describe en el presente documento, en una célula, y mantener la célula bajo condiciones adecuadas (por ejemplo, incluyendo el cultivo de la célula y la gestación de un embrión que comprende la célula, en una madre sucedánea), para que se desarrolle y forme un ratón como se describe en el presente documento.

En un aspecto, se describe un método para producir un ratón modificado, que comprende modificar como se



describe en el presente documento una célula ME murina o una célula pluripotencial o totipotencial o pluripotencial inducida murina, para incluir uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina no reordenados, unidos operativamente a una secuencia constante de cadena pesada de inmunoglobulina, cultivar la célula ME, introducir la célula ME cultivada en un embrión huésped para formar un embrión quimérico, e introducir el embrión quimérico en un ratón huésped adecuado, para que se desarrolle y se forme un ratón modificado. En un aspecto, el uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina no reordenados son segmentos génicos  $\lambda$  humanos o  $\kappa$  humanos. En un aspecto, el uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina no reordenados comprenden segmentos génicos  $V\lambda$  humanos o  $V\kappa$  humanos, y uno o más segmentos génicos  $J\lambda$ ,  $J\kappa$  o  $J_H$ . En un aspecto, la secuencia génica de región constante de cadena pesada es una secuencia humana que se selecciona de  $C_H1$ , bisagra  $C_H2$ ,  $Ch3$ , y una combinación de las mismas. En un aspecto, el uno o más segmentos génicos de región variable de cadena ligera de inmunoglobulina no reordenados, reemplazan la totalidad o sustancialmente la totalidad de segmentos génicos de región variable de cadena pesada de ratón endógenos funcionales, en el locus de cadena pesada de ratón endógeno, y la secuencia constante de cadena pesada es una secuencia murina que comprende una  $C_H1$ , una bisagra, una  $C_H2$  y una  $C_H3$ .

En un aspecto, se describe una región variable de inmunoglobulina (VR) (por ejemplo, que comprende un segmento génico de  $V_L$  humana fusionado con una  $J_L$  o  $J_H$ , o  $D_H$  y  $J_H$ , o  $D_H$  y  $J_L$  humana) producida en un ratón como se describe en el presente documento. En un aspecto específico, la VR de inmunoglobulina se deriva de un segmento génico humano de línea germinal que se selecciona de un segmento génico  $V\kappa$  y un segmento génico  $V\lambda$ , donde la VR está codificada por una secuencia reordenada del ratón donde la secuencia reordenada está somáticamente hipermutada. En un aspecto, la secuencia reordenada comprende 1 a 5 hipermutaciones somáticas. En un aspecto, la secuencia reordenada comprende al menos 6, 7, 8, 9 ó 10 hipermutaciones somáticas. En un aspecto, la secuencia reordenada comprende más de 10 hipermutaciones somáticas. En un aspecto, la secuencia reordenada está fusionada con una o más secuencias de región constante de cadena pesada humana o murina (por ejemplo, que se seleccionan de  $C_H1$ , bisagra,  $C_H2$ ,  $C_H3$  humanas o murinas, y una combinación de las mismas).

En un aspecto, se describe una secuencia de aminoácidos de dominio variable de inmunoglobulina de una proteína de unión, producida en un ratón como se describe en el presente documento. En un aspecto, la VR está fusionada con una o más secuencias de región constante de cadena pesada humana o murina (por ejemplo, que se seleccionan de  $C_H1$ , bisagra,  $C_H2$ ,  $C_H3$ , humana o murina, y una combinación de las mismas).

En un aspecto, se describe un dominio variable de cadena ligera codificado por una secuencia de ácido nucleico derivada de un ratón como se describe en el presente documento.

En un aspecto, se describe un anticuerpo o un fragmento de unión a antígeno del mismo (por ejemplo Fab,  $F(ab)_2$ , scFv) producido en un ratón como se describe en el presente documento, o derivado de una secuencia producida en un ratón como se describe en el presente documento.

### Breve descripción de las figuras

La **fig. 1A** ilustra un esquema (no a escala) del locus de cadena pesada de ratón. El locus de cadena pesada de ratón tiene aproximadamente 3 Mb de longitud y contiene aproximadamente 200 segmentos génicos de región variable de cadena pesada ( $V_H$ ), 13 segmentos génicos de región de diversidad de cadena pesada ( $D_H$ ) y 4 segmentos génicos de región de unión de cadena pesada ( $J_H$ ), así como potenciadores (Pot) y regiones constantes de cadena pesada ( $C_H$ ).

La **fig. 1B** ilustra un esquema (no a escala) del locus de cadena ligera  $\kappa$  humana. El locus de cadena ligera  $\kappa$  humana está duplicado en fragmentos contiguos distal y proximal de polaridad opuesta que abarcan aproximadamente 440 kb y 600 kb, respectivamente. Entre los dos segmentos contiguos hay aproximadamente 800 kb de ADN que se cree que carece de segmentos génicos  $V\kappa$ . El locus de cadena ligera  $\kappa$  humana contiene aproximadamente 76 segmentos génicos  $V\kappa$ , 5 segmentos génicos  $J\kappa$ , un potenciador intrónico (Pot) y una única región constante ( $C\kappa$ ).

La **fig. 2** muestra una estrategia de direccionamiento para la inserción progresiva de 40 segmentos génicos  $V\kappa$  humanos y 5  $J\kappa$  humanos en el locus de cadena pesada de ratón. Se muestran los casetes de selección de higromicina (HYG) y neomicina (NEO) con los sitios de reconocimiento de recombinasa (R1, R2, etc.).

La **fig. 3** muestra un locus de cadena pesada de ratón modificado que comprende segmentos génicos  $V\kappa$  y  $J\kappa$

humanos unidos operativamente a las regiones  $C_H$  murinas.

La **fig. 4A** muestra un ejemplo de estrategia de direccionamiento para la inserción progresiva de un segmento génico  $V\lambda$  humano y un único segmento génico  $J\lambda$  humano en el locus de cadena pesada de ratón. Se muestran los casetes de selección de higromicina (HYG) y neomicina (NEO) con los sitios de reconocimiento de recombinasa (R1, R2, etc.).

La **fig. 4B** muestra un ejemplo de una estrategia de direccionamiento para la inserción progresiva de segmentos génicos  $V\lambda$  humanos y cuatro segmentos génicos  $J\lambda$  humanos en el locus de cadena pesada de ratón. Se muestran los casetes de selección de higromicina (HYG) y neomicina (NEO) con los sitios de reconocimiento de recombinasa (R1, R2, etc.).

La **fig. 5A** muestra un ejemplo de una estrategia de direccionamiento para la inserción progresiva de segmentos génicos  $V\lambda$ , humanos,  $D_H$  humanos y  $J_H$  humanos, en el locus de cadena pesada de ratón. Se muestran los casetes de selección de higromicina (HYG) y neomicina (NEO) con los sitios de reconocimiento de recombinasa (R1, R2, etc.).

La **fig. 5B** muestra un ejemplo de una estrategia de direccionamiento para la inserción progresiva de segmentos génicos  $V\lambda$ , humanos,  $D_H$  humanos y  $J_K$  humanos, en el locus de cadena pesada de ratón. Se muestran los casetes de selección de higromicina (HYG) y neomicina (NEO) con los sitios de reconocimiento de recombinasa (R1, R2, etc.).

La **fig. 6A** muestra gráficas de contorno de esplenocitos teñidos para la expresión de superficie de B220 e IgM de un ratón de tipo silvestre (TS) representativo y un ratón representativo homocigoto para seis segmentos génicos  $V_K$  humanos y cinco segmentos génicos  $J_K$  humanos situados en el locus de la cadena pesada endógena (6h $V_K$ -5h $J_K$  HO).

La **fig. 6B** muestra gráficas de contorno de esplenocitos activados en células B CD19<sup>+</sup> y teñidos para inmunoglobulina D (IgD) e inmunoglobulina M (IgM), procedentes de un ratón representativo de tipo silvestre (TS) y un ratón representativo homocigoto para seis segmentos génicos  $V_K$  humanos y cinco segmentos génicos  $J_K$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_K$ -5h $J_K$  HO).

La **fig. 6C** muestra el número total de células B CD19<sup>+</sup>, células B de transición (CD19<sup>+</sup>IgM<sup>hi</sup>IgD<sup>int</sup>) y células B maduras (CD19<sup>+</sup>IgM<sup>int</sup>IgD<sup>hi</sup>) en bazo extraídos de ratones de tipo silvestre (TS) y ratones homocigotos para seis segmentos génicos  $V_K$  humanos y cinco segmentos génicos  $J_K$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_K$ -5h $J_K$  HO).

La **fig. 7A** muestra gráficas de contorno de médula ósea activada en singletes teñidos por inmunoglobulina M (IgM) y B220, procedentes de un ratón de tipo silvestre (TS) y un ratón homocigoto para seis segmentos génicos  $V_K$  humanos y cinco segmentos génicos  $J_K$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_K$ -5h $J_K$  HO). Se observan células B inmaduras, maduras y pro/pre B en cada uno de los gráficos de puntos.

La **fig. 7B** muestra el número total de células B pre/Pro (B220<sup>+</sup>IgM<sup>-</sup>), inmaduras (B220<sup>int</sup>IgM<sup>+</sup>) y maduras (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 seis segmentos génicos  $V_K$  humanos y cinco segmentos génicos  $J_K$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_K$ -5h $J_K$  HO).

La **fig. 7C** muestra gráficas de contorno de médula ósea activada en CD19<sup>+</sup> y teñida para ckit<sup>+</sup> y CD43<sup>+</sup>, procedente de un ratón de tipo silvestre (TS) y un ratón homocigoto para seis segmentos génicos  $V_K$  humanos y cinco segmentos génicos  $J_K$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_K$ -5h $J_K$  HO). Se observan células pro y pre B en cada uno de los gráficos de puntos.

La **fig. 7D** muestra el número de células pro B (CD19<sup>+</sup>CD43<sup>+</sup>ckit<sup>+</sup>) y pre B (CD19<sup>+</sup>CD43<sup>-</sup>ckit<sup>-</sup>) en médula ósea extraída de los fémures de ratones de tipo silvestre (TS) y ratones homocigotos para seis segmentos génicos  $V_K$  humanos y cinco segmentos génicos  $J_K$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_K$ -5h $J_K$  HO).

La **fig. 7E** muestra gráficas de contorno de médula ósea activada en singletes teñida para CD19 y CD43 procedente

de un ratón de tipo silvestre (TS) y de un ratón homocigoto para seis segmentos génicos  $V_{\kappa}$  humanos y cinco segmentos  $J_{\kappa}$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_{\kappa}$ -5h $J_{\kappa}$  HO). Se observan células B inmaduras, pre y pro B en cada uno de los gráficos.

- 5 La **fig. 7F** muestra histogramas de médula ósea activada en células pre B (CD19<sup>+</sup>CD43<sup>int</sup>) y que expresa inmunoglobulina M (IgM) procedente de un ratón de tipo silvestre (TS) y un ratón homocigoto para seis segmentos génicos  $V_{\kappa}$  humanos y cinco segmentos génicos  $J_{\kappa}$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_{\kappa}$ -5h $J_{\kappa}$  HO).
- 10 La **fig. 7G** muestra el número de células pre B IgM<sup>+</sup> (CD19<sup>+</sup>IgM<sup>+</sup>CD43<sup>int</sup>) y células B inmaduras (CD19<sup>+</sup>IgM<sup>+</sup>CD43<sup>-</sup>), en médula ósea extraída de los fémures de ratones de tipo silvestre (TS) y ratones homocigotos para seis segmentos génicos  $V_{\kappa}$  humanos y cinco segmentos génicos  $J_{\kappa}$  humanos, posicionados en el locus de la cadena pesada endógena (6h $V_{\kappa}$ -5h $J_{\kappa}$  HO).
- 15 La **fig. 8A** muestra gráficas de contorno de esplenocitos activados en CD19<sup>+</sup> y teñidos para la expresión de Igλ<sup>+</sup> e Igκ<sup>+</sup>, procedentes de un ratón que contiene un locus de cadena pesada de tipo silvestre y un reemplazo de los segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  endógenos por segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  humanos (TS), y un ratón homocigoto para treinta segmentos génicos h $V_{\kappa}$  y cinco  $J_{\kappa}$ , en el locus de la cadena pesada endógena y un reemplazo de los segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  endógenos por segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  humanos (30h $V_{\kappa}$ -5h $J_{\kappa}$  HO).
- 20 La **fig. 8B** muestra gráficas de contorno de médula ósea activada en células B inmaduras (B220<sup>int</sup>IgM<sup>+</sup>) y maduras (B220<sup>hi</sup>IgM<sup>+</sup>), teñidas para la expresión de Igλ e Igκ, aislada de los fémures de un ratón que contenía un locus de cadena pesada de tipo silvestre y un reemplazo de los segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  endógenos por segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  humanos (TS) y un ratón homocigoto para treinta segmentos génicos h $V_{\kappa}$  y cinco  $J_{\kappa}$ , en el locus de
- 25 cadena pesada endógena y un reemplazo de los segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  endógenos por segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  humanos (30h $V_{\kappa}$ -5h $J_{\kappa}$  HO).

La **fig. 9** muestra una alineación de secuencia nucleotídica de la unión  $V_{\kappa}$ - $J_{\kappa}$ -mIgG de doce clones de RT-PCR independientes amplificados a partir de ARN de esplenocitos de ratones no tratados homocigotos para treinta

30 segmentos génicos h $V_{\kappa}$  y cinco segmentos génicos  $J_{\kappa}$  en el locus de cadena pesada de ratón y un reemplazo de los segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  endógenos por segmentos génicos  $V_{\kappa}$  y  $J_{\kappa}$  humanos. Las bases en letra minúscula indican bases que no son de línea germinal resultantes de la mutación y/o la N adición durante la recombinación. Se incluyen espacios artificiales (periodos) para alinear apropiadamente la región marco 4 y mostrar alineación de la secuencia nucleotídica IgG de cadena pesada de ratón para clones cebados con IgG1, IgG2a/c e IgG3.

### 35 Descripción detallada

La frase "proteína de unión biespecífica" incluye una proteína de unión capaz de unirse selectivamente a dos o más epítopos. Las proteínas de unión biespecíficas comprenden dos diferentes polipéptidos que comprenden un primer

40 dominio variable de cadena ligera ( $V_L1$ ), fusionado con una primera región  $C_H$  y un segundo dominio variable de cadena ligera ( $V_L2$ ) fusionado con una segunda región  $C_H$ . En general, la primera y segunda regiones  $C_H$  son idénticas, o difieren en una o más sustituciones de aminoácido (por ejemplo, como se describe en el presente documento). Las regiones  $V_L1$  y  $V_L2$  se unen específicamente a diferentes epítopos-ya sea en dos moléculas diferentes (por ejemplo, antígenos) o en la misma molécula (por ejemplo, en el mismo antígeno). Si una proteína de

45 unión biespecífica se une selectivamente a dos epítopos diferentes (un primer epítipo y un segundo epítipo), la afinidad de la  $V_L1$  por el primer epítipo generalmente será al menos de uno a dos o de tres a cuatro órdenes de magnitud más baja que la afinidad de  $V_L1$  por el segundo epítipo, y viceversa, con respecto a  $V_L2$ . Los epítopos reconocidos por la proteína de unión biespecífica pueden estar en la misma diana o en una diferente (por ejemplo, en el mismo antígeno o en uno diferente). Las proteínas de unión biespecíficas pueden prepararse, por ejemplo, al

50 combinar una  $V_L1$  y una  $V_L2$  que reconozcan diferentes epítopos del mismo antígeno. Por ejemplo, se pueden fusionar secuencias de ácido nucleico que codifican secuencias  $V_L$  que reconozcan diferentes epítopos del mismo antígeno, con secuencias de ácido nucleico que codifican diferentes regiones  $C_H$ , y tales secuencias pueden ser expresadas en una célula que exprese una cadena ligera de inmunoglobulina, o se pueden expresar en una célula que no exprese una cadena ligera de inmunoglobulina. Una proteína de unión biespecífica típica tiene dos cadenas

55 pesadas, teniendo cada una tres CDR de cadena ligera, seguidas de (del extremo N-terminal al extremo C-terminal) un dominio  $C_H1$ , una gbisagra, un dominio  $C_H2$ , y un dominio  $C_H3$ , y una cadena ligera de inmunoglobulina que no confiere especificidad de unión al antígeno, pero puede asociarse con cada una de las cadenas pesadas, o que puede asociarse con cada una de las cadenas pesadas y que puede unirse a uno o más de los epítopos unidos por

$V_L1$  y/o  $V_L2$ , o que puede asociarse con cada una de las cadenas pesadas y hacer posible la unión o ayudar a la unión de una o ambas de las cadenas pesadas a uno o ambos epítotos.

Por lo tanto, dos tipos generales de proteínas de unión biespecíficas son (1)  $V_L1-C_H$ (dímero), y (2)  $V_L1-C_H$ :cadena ligera +  $V_L2-C_H$ :cadena ligera, donde la cadena ligera es igual o diferente. En cualquiera de los casos, la  $C_H$  (es decir, la región constante de cadena pesada) puede ser diferencialmente modificada (por ejemplo, para unirse diferencialmente a la proteína A, para aumentar la semivida en suero, etc.) como se describe en el presente documento, o puede ser la misma.

10 El término "célula" cuando se utiliza con referencia a la expresión de una secuencia, incluye cualquier célula que sea adecuada para expresar una secuencia de ácido nucleico recombinante. Las células incluyen las de procariontes y eucariotes (de una única célula o de múltiples células), células bacterianas (por ejemplo, cepas de *E. coli*, *Bacillus* spp., *Streptomyces* spp., etc.), células de micobacterias, células de hongos, células de levaduras (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 por baculovirus, *Trichoplusia ni*, etc.), células de animales no humanos, células humanas, células B, o fusiones celulares, tales como por ejemplo, hibridomas o cuadromas. En algunas realizaciones, la célula es una célula humana, de mono, de simio, de hámster, de rata o de ratón. En algunas realizaciones, la célula es eucariótica y se selecciona de entre 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 (epidérmicas), 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 derivada de una célula anteriormente mencionada. En algunas realizaciones, la célula comprende uno o más genes víricos, por ejemplo una célula retiniana que expresa un gen vírico (por ejemplo, una célula PER.C6™).

25 El término "equivalente" cuando se utiliza en el sentido de "equivalente a", por ejemplo, un primer dominio  $V_L$  que es "equivalente a" un segundo dominio  $V_L$ , pretende incluir una referencia a la relación entre dos dominios  $V_L$  procedentes de una misma proteína de unión producida por un ratón de acuerdo con la invención. Por ejemplo, un ratón que se modifica genéticamente de acuerdo con un aspecto de la presente divulgación, por ejemplo un ratón que tiene un locus de cadena pesada en el que los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  se reemplazan por segmentos génicos  $V_L$  y  $J_L$ , produce proteínas de unión similares a anticuerpo que tienen dos cadenas polipeptídicas idénticas hechas de la misma región  $C_H$  de ratón (por ejemplo un isotipo IgG) fusionada con un primer dominio  $V_L$  humano, y dos cadenas polipeptídicas idénticas hechas de la misma región  $C_L$  murina fusionadas con un segundo dominio  $V_L$  humano. Durante la selección clonal en el ratón, el primer y segundo dominios  $V_L$  humanos se seleccionaron por el proceso de selección clonal para aparecer juntos en el contexto de una única proteína de unión similar a anticuerpo. Por lo tanto, el primer y segundo dominios  $V_L$  que aparecen juntos, como resultado del proceso de selección clonal, en una única molécula similar a anticuerpo, se refieren como "equivalentes". Por el contrario, un dominio  $V_L$  que aparece en una primera molécula similar a anticuerpo y un dominio  $V_L$  que aparece en una segunda molécula similar a anticuerpo, *no son* equivalentes, a *menos que* la primera y la segunda moléculas similares a anticuerpo tengan cadenas pesadas idénticas (es decir, a menos que el dominio  $V_L$  fusionado con la primera región de cadena pesada humana y el dominio  $V_L$  fusionado con la segunda región de cadena pesada humana, sean idénticos).

La frase "región determinante de complementariedad" o el término "CDR", incluye una secuencia de aminoácidos codificada por una secuencia de ácido nucleico de los genes de inmunoglobulina de un organismo, que normalmente aparece (es decir, en un animal de tipo silvestre) entre dos regiones marco en una región variable de una cadena ligera o pesada de una molécula de inmunoglobulina (por ejemplo, un anticuerpo o un receptor de células T). Una CDR puede codificarse, por ejemplo, por una secuencia de línea germinal, o por una secuencia reordenada o no reordenada, y, por ejemplo, por una célula B sin tratar o madura, o por una célula T. En algunas circunstancias (por ejemplo, para una CDR3), las CDR pueden codificarse por dos o más secuencias (por ejemplo, secuencias de 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 células B, por ejemplo, como resultado de corte y empalme o por la conexión de las secuencias (por ejemplo, una recombinación V-D-J para formar una CDR3 de cadena pesada).

La frase "segmento génico", o "segmento", incluye la referencia a un segmento génico de inmunoglobulina V (de cadena ligera o pesada) o D o J (de cadena ligera o pesada), que incluye secuencias no reordenadas en loci de inmunoglobulina (por ejemplo, en seres humanos y ratones) que pueden participar en un reordenamiento (mediado, por ejemplo, por recombinasas endógenas), para formar una secuencia V/J o V/D/J reordenada. A menos que se indique otra cosa, los segmentos génicos V, D y J comprenden las secuencias de señal de recombinación (RSS) que permiten la recombinación V/J o la recombinación V/D/J, de acuerdo con la regla 12/23. A menos que se indique otra

cosa, los segmentos génicos comprenden además secuencias con las que están asociadas en la naturaleza o equivalentes funcionales de las mismas (por ejemplo, para promotores y líderes de segmentos génicos V).

La frase "cadena pesada" o "cadena pesada de inmunoglobulina" incluye una región constante de cadena pesada de inmunoglobulina procedente de cualquier organismo, y a menos que se especifique de otro modo, incluye un dominio variable de cadena pesada ( $V_H$ ). Los dominios  $V_H$  incluyen tres CDR de cadena pesada y cuatro regiones marco (FR), a menos que se especifique otra cosa. Los fragmentos de cadenas pesadas incluyen CDR, CDR y FR, y combinaciones de los mismos. Una cadena pesada típica consiste esencialmente en, después del dominio variable (del extremo N-terminal al extremo C-terminal): un dominio  $C_H1$ , una región de bisagra, un dominio  $C_H2$ , un dominio  $C_H3$ , y opcionalmente un dominio  $C_H4$  (por ejemplo, en el caso de la IgM o IgE), y un dominio transmembrana (M) (por ejemplo, en el caso de inmunoglobulina unida a la membrana en los linfocitos). Una región constante de cadena pesada es una región de una cadena pesada que se extiende (desde el extremo N-terminal hacia el extremo C-terminal) desde fuera de FR4 hasta el extremo C-terminal de la cadena pesada. Las regiones constantes de cadena pesada con desviaciones menores, por ejemplo truncamientos de uno, dos, tres o varios aminoácidos del extremo C-terminal, se incluirán por la frase "región constante de cadena pesada", así como las regiones constantes de cadena pesada con modificaciones de secuencia, por ejemplo 1, 2, 3, 4, 5, 6, 7, 8, 9, ó 10 sustituciones de aminoácido. Las sustituciones de aminoácido pueden hacerse en una o más posiciones que se seleccionan de, por ejemplo (con referencia a la numeración UE de una región constante de inmunoglobulina, por ejemplo una región constante de IgG humana), 228, 233, 234, 235, 236, 237, 238, 239, 241, 248, 249, 250, 252, 254, 255, 256, 258, 265, 267, 268, 269, 270, 272, 276, 278, 280, 283, 285, 286, 289, 290, 292, 293, 294, 295, 296, 297, 298, 301, 303, 305, 307, 308, 309, 311, 312, 315, 318, 320, 322, 324, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 337, 338, 339, 340, 342, 344, 356, 358, 359, 360, 361, 362, 373, 375, 376, 378, 380, 382, 383, 384, 386, 388, 389, 398, 414, 416, 419, 428, 430, 433, 434, 435, 437, 438, y 439.

Por ejemplo, y no a modo de limitación, una región constante de cadena pesada puede modificarse para mostrar una semividua en suero aumentada (en comparación con la misma región constante de cadena pesada, sin la modificación o modificaciones mencionadas) y puede tener una modificación en la posición 250 (por ejemplo E o Q); 250 y 428 (por ejemplo, L o F); 252 (por ejemplo, L/Y/F/W o T), 254 (por ejemplo, S o T), y 256 (por ejemplo, S/R/Q/E/D o T); o una modificación en 428 y/o 433 (por ejemplo, L/R/SI/P/Q o K) y/o 434 (por ejemplo, H/F o Y); o una modificación en 250 y/o 428; o una modificación en 307 ó 308 (por ejemplo, 308F, V308F), y 434. En otro ejemplo, la modificación puede comprender una modificación 428L (por ejemplo, M428L) y 434S (por ejemplo, N434S); una modificación 428L, 259I (por ejemplo, V259I), y una modificación 308F (por ejemplo, V308F); una modificación 433K (por ejemplo, H433K) y una modificación 434 (por ejemplo, 434Y); una modificación 252, 254, y 256 (por ejemplo, 252Y, 254T, y 256E); una modificación 250Q y 428L (por ejemplo, T250Q y M428L); una modificación 307 y/o 308 (por ejemplo, 308F o 308P).

La frase "cadena ligera" incluye una región constante de cadena ligera ( $C_L$ ) de inmunoglobulina de cualquier organismo, y a menos que se especifique otra cosa, incluye cadenas ligeras  $\kappa$  humana y  $\lambda$  humana. Los dominios variables de cadena ligera ( $V_L$ ) típicamente incluyen tres CDR de cadena ligera y cuatro regiones marco (FR), a menos que se especifique de otro modo. En general, una cadena ligera de longitud completa ( $V_L + C_L$ ) incluye, del extremo amino al extremo carboxilo, un dominio  $V_L$  que incluye FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, y una región  $C_L$ . Las cadenas ligeras ( $V_L + C_L$ ) que pueden usarse en esta invención, incluyen aquellas que, por ejemplo, no se unen selectivamente a un primer o segundo epítipo (en el caso de proteínas de unión biespecíficas) selectivamente unido por la proteína de unión (por ejemplo, el o los epítopos selectivamente unidos por el dominio  $V_L$  fusionado con la región  $C_H$ ). Los dominios  $V_L$  que no se unen selectivamente al o los epítopos unidos por el dominio  $V_L$  que está fusionado con la región  $C_H$ , incluyen aquellos que pueden ser identificados mediante el cribado de las cadenas ligeras más comúnmente empleadas en las bibliotecas de anticuerpos existentes (bibliotecas húmedas o *in silico*), donde las cadenas ligeras no interfieren sustancialmente con la afinidad y/o selectividad de los dominios de unión al epítipo de las proteínas de unión. Las cadenas ligeras adecuadas incluyen aquellas que pueden unirse (por sí solas o en combinación con su  $V_L$  equivalente fusionada con la región  $C_H$ ) a un epítipo que es unido específicamente por el dominio  $V_L$  fusionado con la región  $C_H$ .

La frase "rango micromolar" pretende referirse a 1-999 micromolar; la frase "rango nanomolar" pretende referirse a 1-999 nanomolar; la frase "rango picomolar" pretende referirse a 1-999 picomolar.

La frase "animales no humanos" pretende incluir cualquier vertebrado, tal como ciclostomas, peces con esqueleto, peces cartilaginosos tales como tiburones y rayas, anfibios, reptiles, mamíferos y aves. Los animales no humanos adecuados incluyen los mamíferos. Los mamíferos adecuados incluyen primates no humanos, cabras, ovejas, cerdos, perros, vacas y roedores. Los animales no humanos adecuados se seleccionan de la familia de roedores,

incluyendo ratas y ratones. De acuerdo con la invención, los animales no humanos son ratones.

**Ratones, secuencias nucleotídicas y proteínas de unión**

5 Se describen proteínas de unión que están codificadas por elementos de loci de inmunoglobulina, donde las proteínas de unión comprenden regiones constantes de cadena pesada de inmunoglobulina fusionadas con dominios variables de cadena ligera de inmunoglobulina. Además, se describen múltiples estrategias para modificar genéticamente un locus de cadena pesada de inmunoglobulina en un ratón, para codificar proteínas de unión que contengan elementos codificados por loci de cadena ligera de inmunoglobulina. Tales ratones genéticamente  
10 modificados representan una fuente para generar poblaciones únicas de proteínas de unión que tienen una estructura de inmunoglobulina, pero que todavía exhiben una mayor diversidad sobre los anticuerpos tradicionales.

Los aspectos de las proteínas de unión que se describen en el presente documento, incluyen proteínas de unión que están codificadas por loci de inmunoglobulina modificados, que se modifican de tal forma que segmentos génicos  
15 que normalmente (es decir, en un animal de tipo silvestre) codifican dominios variables de cadena ligera de inmunoglobulina (o porciones de los mismos), están unidos operativamente a secuencias de nucleótidos que codifican regiones constantes de cadena pesada. Tras el reordenamiento de los segmentos génicos de cadena ligera, se obtiene una secuencia nucleotídica reordenada que comprende una secuencia que codifica un dominio variable de cadena ligera, fusionado con una secuencia que codifica una región constante de cadena pesada. Esta  
20 secuencia codifica un polipéptido que tiene un dominio variable de cadena ligera de inmunoglobulina fusionado con una región constante de cadena pesada. Por lo tanto, en un aspecto, el polipéptido consiste esencialmente, del extremo N-terminal al extremo C-terminal, en un dominio  $V_L$ , un dominio  $C_{H1}$ , una bisagra, un dominio  $C_{H2}$ , un dominio  $C_{H3}$ , y opcionalmente un dominio  $C_{H4}$ .

25 En los ratones modificados descritos en el presente documento, se producen dichas proteínas de unión que también comprenden una cadena ligera equivalente, donde, en un aspecto, la cadena ligera equivalente se empareja con el polipéptido que se ha descrito anteriormente, para producir una proteína de unión que es similar a un anticuerpo, pero la proteína de unión comprende un dominio  $V_L$ -no un dominio  $V_H$ - fusionado con una región  $C_H$ .

30 En diversos aspectos, los ratones modificados producen proteínas de unión que comprenden un dominio  $V_L$  fusionado con una región  $C_H$  (una cadena pesada híbrida), donde el dominio  $V_L$  de la cadena pesada híbrida exhibe un mayor grado de hipermutación somática. En estos aspectos, la mejora es sobre un dominio  $V_L$  que está fusionado con una región  $C_L$  (una cadena ligera). En algunos aspectos, un dominio  $V_L$  de una cadena pesada híbrida exhibe aproximadamente 1,5, 2, 2,5, 3, 3,5, 4, 4,5 ó 5 veces o más hipermutaciones somáticas, que un  
35 dominio  $V_L$  fusionado con una región  $C_L$ . En algunos aspectos, los ratones modificados en respuesta a un antígeno, exhiben una población de proteínas de unión que comprenden un dominio  $V_L$  de cadena pesada híbrida, donde la población de proteínas de unión muestra un promedio de aproximadamente 1,5, 2, 2,5, 3, 3,5, 4, 4,5, 5 veces o más hipermutaciones somáticas, en el dominio  $V_L$  de la cadena pesada híbrida, de lo que se observa en un ratón de tipo silvestre en respuesta al mismo antígeno. En un aspecto, las hipermutaciones somáticas en el dominio  $V_L$  de la  
40 cadena pesada híbrida, comprenden una o más, o dos o más, N adiciones en una CDR3.

En diversos aspectos, las proteínas de unión comprenden dominios variables codificados por secuencias de cadena ligera de inmunoglobulina que comprenden un número mayor de N adiciones que las observadas en la naturaleza para las cadenas ligeras reordenadas a partir de un locus de cadena ligera endógena, por ejemplo una proteína de  
45 unión que comprende una región constante de cadena pesada de ratón fusionada con un dominio variable derivado de segmentos génicos V de cadena ligera humana y segmentos génicos J humanos (de cadena ligera o pesada), donde los segmentos génicos V humanos y J humanos se reordenan para formar un gen reordenado que comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más, N adiciones.

50 En diversos aspectos, los ratones descritos producen proteínas de unión que en promedio son más pequeñas que los anticuerpos de tipo silvestre (es decir, anticuerpos que tienen un dominio  $V_H$ ), y poseen ventajas asociadas a un tamaño menor. El menor tamaño se debe, al menos en parte, a la ausencia de una secuencia de aminoácidos codificada por un segmento génico  $D_H$ , normalmente presente en un dominio  $V_H$ . El menor tamaño también se puede deber a la formación de una CDR3 derivada, por ejemplo, de un segmento génico  $V_k$  y un segmento génico  
55  $J_k$ .

En otro aspecto, se describe un ratón y un método para proporcionar una población de proteínas de unión que tienen dominios  $V_L$  somáticamente hipermutados, por ejemplo dominios  $V_k$  humanos somáticamente mutados y, por ejemplo, dominios  $V_k$  humanos codificados por genes variables  $\kappa$  reordenados, que comprenden 1-10 o más N

adiciones. En un aspecto, en ausencia de un dominio  $V_H$  para generar diversidad de anticuerpos, un ratón de la invención generará proteínas de unión, por ejemplo, en respuesta a una estimulación con un antígeno, cuyos dominios  $V$  son exclusiva o sustancialmente dominios  $V_L$ . El proceso de selección clonal del ratón, por lo tanto, se limita a seleccionar exclusiva o sustancialmente proteínas de unión que tienen dominios  $V_L$ , en vez de dominios  $V_H$ .

- 5 La hipermutación somática de los dominios  $V_L$  será tan frecuente, o sustancialmente más frecuente (por ejemplo, 2 a 5 veces mayor, o más) que en los ratones de tipo silvestre (los cuales también tienen mutaciones en dominios  $V_L$  con alguna frecuencia). El proceso de selección clonal en un ratón de la invención, generará proteínas de unión de alta afinidad a partir del locus de inmunoglobulina modificado, incluyendo proteínas de unión que se unen específicamente a un epítipo con una afinidad en el rango nanomolar o picomolar. Las secuencias que codifican  
10 tales proteínas de unión, se pueden emplear para producir proteínas de unión terapéuticas que contengan dominios variables humanos y regiones constantes humanas, usando un sistema de expresión apropiado.

- En otros aspectos, se puede producir un ratón donde los loci de inmunoglobulina de cadena pesada y/o de cadena ligera murina están deshabilitados, se vuelven no funcionales, o se inactivan, y se pueden colocar transgenes  
15 completamente humanos o quiméricos humano-murinos, en el ratón, donde al menos uno de los transgenes contiene un locus de cadena pesada modificado (por ejemplo, que tiene segmentos génicos de cadena ligera unidos operativamente a una o más secuencias génicas de cadena pesada). Tales ratones también pueden producir una proteína de unión como se describe en el presente documento.

- 20 En un aspecto, se describe un método para aumentar la diversidad, incluyendo por hipermutación somática o por N adiciones en un dominio  $V_L$ , que comprende colocar un segmento génico  $V$  de cadena ligera no reordenado y un segmento génico  $J$  no reordenado, unidos operativamente a una secuencia génica  $C_H$  murina; exponer al animal a un antígeno de interés; y aislar del animal una secuencia génica  $V(\text{ligera})/J$  reordenada y somáticamente hipermutada, donde dicha secuencia génica  $V(\text{ligera})/J$  reordenada está fusionada con una secuencia nucleotídica  
25 que codifica una región  $C_H$  de inmunoglobulina.

En un aspecto, la cadena pesada de inmunoglobulina fusionada con la  $V_L$  hipermutada, es una IgM; en otro aspecto, una IgG; en otro aspecto, una IgE; en otro aspecto, una IgA.

- 30 En un aspecto, el dominio  $V_L$  de clase conmutada y somáticamente hipermutado contiene aproximadamente de 2 a 5 veces o más de las hipermutaciones somáticas observadas para un anticuerpo de clase conmutada y reordenado que tenga un dominio  $V_L$ , que esté unido operativamente a una región  $C_L$ . En un aspecto, las hipermutaciones somáticas observadas en el dominio  $V_L$  somáticamente hipermutado, son aproximadamente las mismas en número que las observadas en un dominio  $V_H$  expresado a partir de un segmento génico  $V_H$  fusionado con una región  $C_H$ .

- 35 En un aspecto, se describe un método para preparar un dominio  $V_L$  humano de alta afinidad, que comprende exponer un ratón descrito en el presente documento a un antígeno de interés, permitir que el ratón desarrolle una respuesta inmunitaria hacia el antígeno de interés, y aislar un dominio  $V_L$  humano de clase conmutada y somáticamente mutado del ratón, que se une específicamente al antígeno de interés con una alta afinidad.

- 40 En un aspecto, la  $K_D$  de una proteína de unión que comprende el dominio  $V_L$  humano de clase conmutada y somáticamente mutado, está en el rango nanomolar o picomolar.

- 45 En un aspecto, la proteína de unión consiste esencialmente en un polipéptido dimérico, donde el polipéptido consiste esencialmente en la proteína de unión de clase conmutada y somáticamente mutada, que comprende un dominio  $V_L$  humano fusionado con una región  $C_H$  humana.

- 50 En un aspecto, la proteína de unión consiste esencialmente en un polipéptido dimérico y dos cadenas ligeras, donde el polipéptido consiste esencialmente en la proteína de unión de clase conmutada y somáticamente mutada que tiene un dominio  $V_L$  humano fusionado con una región  $C_H$  humana; y donde cada polipéptido del dímero está asociado con una cadena ligera equivalente que comprende un dominio  $V_L$  de cadena ligera equivalente y una región  $C_L$  humana.

- 55 En un aspecto, se describe un método para hipermutar somáticamente una secuencia génica  $V_L$  humana, que comprende colocar un segmento génico  $V_L$  humano y un segmento génico  $J_L$  humano unidos operativamente a una región  $C_H$  de ratón endógena en un locus de inmunoglobulina de cadena pesada de ratón endógeno, exponer el ratón a un antígeno de interés; y obtener del ratón un dominio  $V_L$  humano somáticamente hipermutado, que se una al antígeno de interés.

En un aspecto, el método comprende además obtener a partir del ratón una secuencia génica  $V_L$  de una cadena ligera, que sea equivalente al dominio  $V_L$  humano somáticamente hipermutado que se une al antígeno de interés.

#### Proteínas de Unión $V_L$ con secuencias $D_H$

5

En varios aspectos, los ratones que comprenden un segmento génico V de cadena ligera de inmunoglobulina no reordenado y un segmento génico J (por ejemplo, de cadena ligera o pesada) no reordenado, también comprenden un segmento génico  $D_H$  no reordenado que es capaz de recombinarse con el segmento génico J para formar una secuencia D/J reordenada, que, a su vez, es capaz de reordenarse con el segmento génico V de cadena ligera para formar una secuencia variable reordenada derivada de (a) el segmento génico V de cadena ligera, (b) el segmento génico  $D_H$ , y (c) el segmento génico J (por ejemplo, de cadena ligera o pesada); donde la secuencia variable reordenada está unida operativamente a una región constante de cadena pesada (por ejemplo, que se selecciona de CH1, bisagra, CH2, CH3, y una combinación de las mismas; por ejemplo, unida operativamente a una CH1, una bisagra, una CH2 y una CH3 murina o humana).

15

En diversos aspectos, los ratones que comprenden segmentos génicos V de cadena ligera humana no reordenados y segmentos génicos J que también comprenden un segmento génico D humano, son útiles, por ejemplo, como fuente de una mayor diversidad de secuencias CDR3. Normalmente, las secuencias CDR3 surgen en las cadenas ligeras por recombinación V/J, y en las cadenas pesadas por recombinación V/D/J. Se proporciona diversidad adicional mediante adiciones de nucleótidos que ocurren durante la recombinación (por ejemplo, N adiciones) y también como resultado de hipermutación somática. Las características de unión conferidas por las secuencias CDR3, generalmente están limitadas a aquellas conferidas por la secuencia CDR3 de cadena ligera, la secuencia CDR3 de cadena pesada, y una combinación de las secuencias CDR3 de cadena ligera y pesada, según pueda ser el caso. En los ratones como se describe en el presente documento, sin embargo, está disponible una fuente de diversidad adicional debido a las características de unión conferidas como resultado de una combinación de una primera CDR3 de cadena ligera (en el polipéptido de cadena pesada) y una segunda CDR3 de cadena ligera (en el polipéptido de cadena ligera). Es posible una diversidad adicional cuando la primera CDR3 de cadena ligera puede contener una secuencia derivada de un segmento génico D, como en un ratón como se describe en el presente documento, que comprende un segmento génico V no reordenado procedente de un dominio V de cadena ligera, unido operativamente a un segmento génico D y unido operativamente a un segmento génico J (de cadena ligera o pesada), empleando el diseño de RSS como se indica aquí.

Otra fuente de diversidad son las N adiciones y/o P adiciones que pueden ocurrir en las recombinaciones V(ligera)J o V(ligera)/D/J, que son posibles en los ratones como se describe. Por lo tanto, los ratones descritos en el presente documento, no sólo proporcionan una diferente fuente de diversidad (cadena ligera-cadena ligera), sino también una fuente de diversidad adicional debido a la adición de, por ejemplo 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10, o más, N adiciones en un gen V(ligera)/J reordenado o un gen V(ligera)/D/J reordenado, en un ratón como se describe en el presente documento.

En diversos aspectos, el uso de un segmento génico D unido operativamente a un segmento génico J y un segmento génico V de cadena ligera, proporciona una mayor diversidad. La unión operativa de un segmento génico  $D_H$  en este caso, requerirá que el segmento génico D sea capaz de recombinarse con el segmento J en cuestión. Por lo tanto, se requerirá que el segmento génico D tenga yuxtapuesta una RSS en 3' que coincida con la RSS yuxtapuesta en 5' del segmento génico J, de tal modo que el segmento génico D y el segmento génico J puedan reordenarse. Además, el segmento génico D requerirá una RSS apropiada yuxtapuesta en 5' que coincida con la RSS yuxtapuesta en 3' del segmento génico V, de tal modo que el segmento génico D/J reordenado y el segmento génico V puedan reordenarse, para formar un gen que codifique para un dominio variable.

Una RSS, o secuencia de señal de recombinación, comprende una secuencia heptamérica de ácido nucleico conservada separada por 12 pares de bases (pb) ó 23 pares de bases (pb) de secuencia no conservada, de una secuencia nonamérica de ácido nucleico conservada. Las RSS se usan por las recombinasas para unir segmentos de genes de inmunoglobulina durante el proceso de reordenamiento, siguiendo la regla 12/23. De acuerdo con la regla 12/23, un segmento génico yuxtapuesto con una RSS que tiene un separador de 12 pb (no conservado), se reordena con un segmento génico yuxtapuesto con una RSS que tiene un separador de 23 pb (no conservado); es decir, generalmente no se observan reordenamientos entre segmentos génicos en los que cada uno tiene una RSS con un separador de 12 pb, o donde cada uno tiene una RSS con un separador de 23 pb.

En el caso del locus de cadena ligera  $\lambda$ , los segmentos génicos variables (segmentos génicos  $V_\lambda$ ) están flanqueados en 3' (con respecto a la dirección de la transición de la secuencia V) con una RSS que tiene un separador 23-mer, y



segmentos génicos de unión (segmentos génicos  $J\lambda$ ) están flanqueados en 5' (con respecto a la dirección de la transición de la secuencia J) con una RSS que tiene un separador 12-mer. Por lo tanto, los segmentos génicos  $V\lambda$  y  $J\lambda$  están flanqueados por RSS que son compatibles con la regla 12/23, y por lo tanto son capaces de recombinarse durante el reordenamiento.

5

Sin embargo, en el locus  $\kappa$  de un organismo silvestre, cada segmento génico  $V\kappa$  funcional está flanqueado en 3' por una RSS que tiene un separador 12-mer. Los segmentos génicos  $J\kappa$ , por lo tanto, tienen separadores 23-mer yuxtapuestos en el lado 5' del segmento génico  $J\kappa$ . En el locus de cadena pesada, los segmentos génicos  $V_H$  están yuxtapuestos en 3' por una RSS que tiene un separador 23-mer, seguido de un segmento génico  $D_H$  yuxtapuesto en 10 5' y 3' por un separador 12-mer, y segmentos génicos  $J_H$  cada uno con un segmento 23-mer yuxtapuesto en el lado 5' del segmento génico  $J_H$ . En el locus de la cadena pesada, primero ocurre la recombinación D/J, mediada por la RSS de  $D_H$  en 3' con el separador 12-mer y la RSS de  $J_H$  en 3' con el separador 23-mer, para producir una secuencia D-J reordenada intermedia que tiene una RSS yuxtapuesta en el lado 5' que tiene una RSS con un separador 12-mer. El segmento génico D-J reordenado que tiene la RSS con el separador 12-mer yuxtapuesta en el 15 lado 5', entonces, se reordena con el segmento génico  $V_H$  que tiene la RSS con el separador 23-mer yuxtapuesto en su lado 3', para formar una secuencia V/D/J reordenada.

En un aspecto, se emplea un segmento génico  $V\lambda$ , en el locus de cadena pesada con un segmento génico J que es un segmento génico  $J\lambda$ , donde el segmento génico  $V\lambda$  comprende una RSS yuxtapuesta en el lado 3' de la 20 secuencia  $V\lambda$ , y la RSS comprende un separador 23-mer, y el segmento génico J es un segmento génico  $J\lambda$  con una RSS yuxtapuesta en su lado 5' que tiene un separador 12-mer (por ejemplo, tal como se encuentra en la naturaleza).

En un aspecto, se emplea un segmento génico  $V\lambda$  en el locus de cadena pesada con un segmento génico J que es un segmento génico  $J\kappa$  o  $J_H$ , donde la secuencia  $V\lambda$  tiene yuxtapuesta en su lado 3' una RSS que comprende un 25 separador 23-mer, y el segmento génico  $J\kappa$  o  $J_H$  tiene yuxtapuesta una RSS en su lado 5', que comprende un separador 12-mer.

En un aspecto, se emplea un segmento génico  $V\lambda$  en el locus de cadena pesada con un segmento génico  $D_H$  y un segmento génico J. En un aspecto, el segmento génico  $V\lambda$ , comprende una RSS yuxtapuesta en el lado 3' del 30 segmento génico  $V\lambda$ , con una RSS que tiene un separador 23-mer; el segmento génico  $D_H$  comprende una RSS yuxtapuesta en el lado 5' y en el lado 3' del segmento génico  $D_H$ , que tiene una RSS que tiene un separador 12-mer, y un segmento génico J que tiene una RSS yuxtapuesta en su lado 5' que tiene un separador 23-mer, donde el segmento génico J se selecciona de  $J\lambda$ ,  $J\kappa$ , y  $J_H$ .

35 En un aspecto, el segmento génico  $V\kappa$  se emplea en el locus de cadena pesada con un segmento génico J (sin ningún segmento génico D interviniente), donde el segmento génico  $V\kappa$  tiene una RSS yuxtapuesta en el lado 3' del segmento génico  $V\kappa$ , que comprende una RSS con un separador 12-mer, y el segmento génico J tiene yuxtapuesta en su lado 5', una RSS con un separador 23-mer, y el segmento génico  $J\kappa$  se selecciona de un segmento génico  $J\kappa$ , un segmento génico  $J\lambda$  y un segmento génico  $J_H$ . En un aspecto, el segmento génico V y/o el segmento génico J 40 son humanos.

En un aspecto, el segmento génico  $V\kappa$  se emplea en el locus de cadena pesada con un segmento génico D y un segmento génico J, donde el segmento génico  $V\kappa$  tiene una RSS yuxtapuesta en el lado 3' del segmento génico  $V\kappa$ , que comprende una RSS con un separador 12-mer, el segmento génico D tiene yuxtapuesta en su lado 5' y en su 45 lado 3', una RSS separada 23-mer, y el segmento génico J tiene yuxtapuesta en su lado 5', una RSS separada 12-mer. En un aspecto, el segmento génico J se selecciona de un segmento génico  $J\kappa$ , un segmento génico  $J\lambda$  y un segmento génico  $J_H$ . En un aspecto, el segmento génico V y/o el segmento génico J son humanos.

Un segmento génico  $J\lambda$  con una RSS que tiene un separador 23-mer yuxtapuesta en su extremo 5', o un segmento 50 génico  $J\kappa$  o  $J_H$  con una RSS que tiene un separador 12-mer yuxtapuesta en su extremo 5', se prepara usando cualquiera de los métodos adecuados para preparar secuencias de ácido nucleico que se conocen en la técnica. Un método adecuado para preparar un segmento génico J que tenga una RSS yuxtapuesta en 5', donde la RSS tiene un separador seleccionado (por ejemplo, 12-mer ó 23-mer), es sintetizar químicamente un ácido nucleico que comprende el heptámero, el nonámero, y el separador seleccionado, y fusionarlo con una secuencia de segmento 55 génico J que se sintetiza químicamente o se clona de una fuente adecuada (por ejemplo, una fuente de secuencia humana), y emplear la secuencia del segmento génico J fusionado y una RSS en un vector de direccionamiento para dirigir la RSS-J a un sitio adecuado.

Un segmento génico D con una RSS separada 23 yuxtapuesta en 5' y 3' puede hacerse mediante cualquier método conocido en la técnica. Un método comprende sintetizar químicamente la RSS 23-mer en 5' y la secuencia del segmento génico y la RSS 23-mer en 3', y colocar el segmento génico D flanqueado por RSS en un vector adecuado.

- 5 El vector puede dirigirse para reemplazar uno o más segmentos génicos D de ratón con un segmento génico D humano con secuencias de RSS 12-mer yuxtapuestas en los lados 5' y 3', o dirigirse para insertarse en, por ejemplo, un locus humanizado en una posición entre un segmento génico V humano y un segmento génico J humano o de ratón.
- 10 Se conocen en la técnica nonámeros y heptámeros adecuados para la construcción de RSS (por ejemplo, véanse Janeway's Immunobiology, 7ª ed., Murphy et al., (2008, Garland Science, Taylor & Francis Group, LLC) en la página 148, fig. 4.5). Las secuencias de separador no conservadas adecuadas incluyen, por ejemplo, secuencias de separador observadas en secuencias RSS en loci de inmunoglobulina humana o de ratón.

### 15 **Proteínas de unión biespecíficas**

Las proteínas de unión descritas en el presente documento, y las secuencias de nucleótidos que las codifican, se pueden emplear para preparar proteínas de unión multiespecíficas, por ejemplo proteínas de unión biespecíficas. En este aspecto, un primer polipéptido que consiste esencialmente en un primer dominio  $V_L$  fusionado con una región  $C_H$ , se puede asociar a un segundo polipéptido que consiste esencialmente en un segundo dominio  $V_L$  fusionado con una región  $C_H$ . Cuando el primer dominio  $V_L$  y el segundo dominio  $V_L$  se unen específicamente a un epítipo diferente, se puede preparar una molécula de unión biespecífica empleando los dos dominios  $V_L$ . La región  $C_H$  puede ser igual o diferente. En un aspecto, por ejemplo una de las regiones  $C_H$  puede ser modificada para eliminar un determinante de unión a proteína A, mientras que la otra región constante de cadena pesada no es modificada.

- 20
- 25 Esta disposición particular simplifica el aislamiento de la proteína de unión biespecífica de, por ejemplo, una mezcla de homodímeros (por ejemplo, homodímeros del primer o el segundo polipéptidos).

En un aspecto, los métodos y composiciones que se describen en el presente documento se usan para preparar proteínas de unión biespecíficas. En este aspecto, un primer dominio  $V_L$  que está fusionado a una región  $C_H$  y un segundo dominio  $V_L$  que está fusionado a una región  $C_H$ , se clonan cada uno independientemente en marco con una secuencia de IgG humana del mismo isotipo (por ejemplo, una IgG1, IgG2, IgG3 o IgG4 humana). El primer dominio  $V_L$  se une específicamente a un primer epítipo, y el segundo dominio  $V_L$  se une específicamente a un segundo epítipo. El primer y segundo epítopos pueden estar en antígenos diferentes, o en el mismo antígeno.

- 30
- 35 En un aspecto, el isotipo IgG de la región  $C_H$  fusionada al primer dominio  $V_L$  y el isotipo IgG de la región  $C_H$  fusionada al segundo dominio  $V_L$ , son el mismo isotipo, pero difieren en que un isotipo IgG comprende al menos una sustitución de aminoácido. En un aspecto, la al menos una sustitución de aminoácido hace que la cadena pesada portadora de la sustitución sea incapaz o sustancialmente incapaz de unirse a la proteína A, en comparación con la cadena pesada que carece de la sustitución.

- 40
- 45 En un aspecto, la primera región  $C_H$  comprende un primer dominio  $C_{H3}$  de una IgG humana, que se selecciona de IgG1, IgG2, e IgG4; y la segunda región  $C_H$  comprende un segundo dominio  $C_{H3}$  de una IgG humana que se selecciona de IgG1, IgG2 e IgG4, donde el segundo dominio  $C_{H3}$  comprende una modificación que reduce o elimina la unión del segundo dominio  $C_{H3}$  con la proteína A.

- 50
- 55 En un aspecto, el segundo dominio  $C_{H3}$  comprende una modificación 435R, numerada de acuerdo con el índice UE de Kabat. En otro aspecto, el segundo dominio  $C_{H3}$  comprende además una modificación 436F, numerada de acuerdo con el índice UE de Kabat.

- 60
- 65 En un aspecto, el segundo dominio  $C_{H3}$  es el de una IgG1 humana que comprende una modificación que se selecciona del grupo que consiste en D356E, L358M, N384S, K392N, V397M y V422I, numeradas de acuerdo con el índice UE de Kabat.

- 70
- 75 En un aspecto, el segundo dominio  $C_{H3}$  es el de una IgG2 humana que comprende una modificación que se selecciona del grupo que consiste en N384S, K392N, y V422I, numeradas de acuerdo con el índice UE de Kabat.

En un aspecto, el segundo dominio  $C_{H3}$  es el de una IgG4 humana que comprende una modificación que se selecciona del grupo que consiste en Q355R, N384S, K392N, V397M, R409K, E419Q, y V422I, numeradas de acuerdo con el índice UE de Kabat.

En un aspecto, la proteína de unión comprende regiones  $C_H$  que tienen una o más modificaciones como se menciona en el presente documento, donde la región constante de la proteína de unión es no inmunógena o sustancialmente no inmunógena en un ser humano. En un aspecto específico, las regiones  $C_H$  comprenden 5 secuencias de aminoácidos que no presentan un epítipo inmunógeno en un ser humano. En otro aspecto específico, la proteína de unión comprende una región  $C_H$  que no se encuentra en una cadena pesada humana de tipo silvestre, y la región  $C_H$  no comprende una secuencia que genere un epítipo de células T.

### Ejemplos

10

Los siguientes ejemplos se proporcionan para describir cómo preparar y utilizar los métodos y composiciones que se describen en el presente documento, y no pretenden limitar el alcance de la invención según se define en las reivindicaciones. A menos que se indique otra cosa, la temperatura se indica en grados Celsius, y la presión es atmosférica o casi atmosférica.

15

### Ejemplo I

#### Introducción de segmentos génicos de cadena ligera en un locus de cadena pesada

20 Se prepararon varias construcciones de direccionamiento usando tecnología de ingeniería genética de VELOCIGENE® (véanse, por ejemplo, la Pat. de Estados Unidos n.º 6.586.251 y Valenzuela, D.M., Murphy, A.J., Friendewey, D., Gale, N.W., Economides, A.N., Auerbach, W., Poueymirou, W.T., Adams, N.C., Rojas, J., Yassenchak, J., Chernomorsky, R., Boucher, M., Elsasser, A.L., Esau, L., Zheng, J., Griffiths, J.A., Wang, X., Su, H., Xue, Y., Dominguez, M.G., Noguera, I., Torres, R., Macdonald, L.E., Stewart, A.F., DeChiara, T.M., Yancopoulos, 25 G.D. (2003). High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. Nat Biotechnol 21, 652-659), para modificar bibliotecas de Cromosomas Artificiales Bacterianos (BAC) genómicos murinos. El ADN BAC murino se modificó por recombinación homóloga para inactivar el locus de cadena pesada murino endógeno, mediante la eliminación dirigida a diana de los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$ , para asegurar la inserción de secuencias de genes de cadena ligera  $\kappa$  de línea germinal humana no reordenados (parte superior de la 30 fig. 2).

Brevemente, el locus de cadena pesada de ratón fue eliminado en dos eventos de direccionamiento sucesivos usando una recombinación mediada por recombinasa. El primer evento de direccionamiento incluyó un direccionamiento en el extremo 5' del locus de cadena pesada de ratón, utilizando un vector de direccionamiento que 35 comprende, de 5' a 3', un brazo de homología murino 5', un sitio de reconocimiento de recombinasa, un casete de neomicina, y un brazo de homología 3'. Los brazos de homología 5' y 3' contenían la secuencia 5' del locus de cadena pesada de ratón. El segundo evento de direccionamiento incluyó un direccionamiento en el extremo 3' del locus de cadena pesada de ratón en la región de los segmentos génicos  $J_H$ , utilizando un segundo vector de direccionamiento que contenía, de 5' a 3', un brazo de homología de ratón 5', un sitio de reconocimiento de 40 recombinasa 5', un segundo sitio de reconocimiento de recombinasa, un casete de higromicina, un tercer sitio de reconocimiento de recombinasa, y un brazo de homología de ratón 3'. Los brazos de homología 5' y 3' contenían una secuencia que flanqueaba los segmentos génicos  $J_H$  de ratón y en dirección 5' del potenciador intrónico, y las regiones constantes. Las células ME positivas que contenían un locus de cadena pesada modificado hacia el cual estaban dirigidos ambos vectores de direccionamiento (como se ha descrito anteriormente), se confirmaron por 45 cariotipado. Después, se aisló el ADN de las células ME de doble diana y se sometió a tratamiento con una recombinasa, mediando de esta manera la delección de ADN genómico del locus de cadena pesada de ratón, entre el sitio de reconocimiento de la recombinasa 5' en el primer vector de direccionamiento, y el sitio de reconocimiento de recombinasa 5' en el segundo vector de direccionamiento, dejando un solo sitio de reconocimiento de recombinasa y el casete de higromicina flanqueados por dos sitios de reconocimiento de recombinasa (véase la parte superior de la 50 fig. 2). Por lo tanto, se creó un locus de cadena pesada de ratón modificado que contenía genes  $C_H$  intactos, para insertar progresivamente segmentos génicos de línea germinal  $\kappa$  humana, de una manera precisa mediante el uso de vectores de direccionamiento que se describen a continuación.

Se diseñaron cuatro vectores de direccionamiento por separado, para insertar progresivamente 40 segmentos 55 génicos  $V_\kappa$  humanos y cinco segmentos génicos  $J_\kappa$  humanos, en el locus de cadena pesada de ratón inactivado (descrito anteriormente), empleando las técnicas moleculares estándar reconocidas en la técnica (fig. 2). Los segmentos génicos  $\kappa$  humanos utilizados para el diseño por ingeniería de las cuatro construcciones de direccionamiento, se encuentran en la naturaleza de manera contigua proximal al locus de cadena ligera  $\kappa$  humana de línea germinal (fig. 1B y Tabla 1).

Se diseñó por ingeniería un fragmento genómico humano de ~110.499 pb que contenía los primeros seis segmentos génicos  $V\kappa$  humanos y cinco segmentos génicos  $J\kappa$  humanos, para que incluyera un sitio PI-SceI de 431 pb aguas abajo (3') del segmento génico  $J\kappa 5$  humano. Otro sitio PI-SceI se diseñó por ingeniería en el extremo 5' de un fragmento genómico de ~7.852 pb, que contenía el potenciador intrónico de cadena pesada de ratón, la región de conmutación IgM ( $S\mu$ ) y el gen IgM del locus de cadena pesada de ratón. Este fragmento de ratón se usó como un brazo de homología 3', al ligarlo al fragmento humano de ~110,5 kb, lo que creó una unión 3' que contenía de 5' a 3', la secuencia genómica de ~110,5 kb del locus de cadena ligera  $\kappa$  humana que contenía los primeros seis segmentos génicos  $V\kappa$  consecutivos y cinco segmentos génicos  $J\kappa$ , un sitio PI-SceI, una secuencia de cadena pesada de ratón de ~7.852 pb que contenía el potenciador intrónico de ratón,  $S\mu$  y el gen constante IgM de ratón. Aguas arriba (5') del segmento génico  $V\kappa 1-6$  humano, se encontraba una secuencia  $\kappa$  humana adicional de 3.710 pb, antes del inicio del brazo de homología de ratón 5', que contenía 19.752 pb de ADN genómico de ratón correspondiente a la secuencia 5' del locus de cadena pesada de ratón. Entre el brazo de homología 5' y el inicio de la secuencia humana  $\kappa$ , había un casete de neomicina flanqueado por tres sitios de reconocimiento de recombinasa (véase el Vector de direccionamiento 1, fig. 2). El vector de direccionamiento final para la primera inserción de la secuencia  $\kappa$  humana, de 5' a 3', incluía un brazo de homología 5' que contenía ~20 kb de secuencia genómica de ratón en 5' del locus de la cadena pesada, un primer sitio de reconocimiento de recombinasa (R1), un casete de neomicina, un segundo sitio de reconocimiento de recombinasa (R2), un tercer sitio de reconocimiento de recombinasa (R3), una secuencia  $\kappa$  genómica humana de ~110,5 kb que contenía los primeros seis segmentos génicos  $V\kappa$  humanos consecutivos y cinco segmentos génicos  $J\kappa$  humanos, un sitio PI-SceI, y un brazo de homología 3' que contenía ~8 kb de secuencia genómica de ratón, incluyendo el potenciador intrónico,  $S\mu$  y el gen constante IgM de ratón (véase la fig. 2, Vector de direccionamiento 1). La recombinación homóloga con este vector de direccionamiento, creó un locus de cadena pesada de ratón modificado que contenía seis segmentos génicos  $V\kappa$  humanos y cinco segmentos génicos  $J\kappa$  humanos, unidos operativamente a los genes constantes de cadena pesada de ratón endógenos, que, después de la recombinación, dio lugar a una cadena pesada híbrida (es decir, un dominio  $V\kappa$  humano y una región  $C_H$  de ratón).

Tabla 1

Vector de direccionamiento	Tamaño de la secuencia $\kappa$ humana	Segmentos génicos $\kappa$ humanos añadidos	
		$V_{\kappa}$	$J_{\kappa}$
1	~110,5 kb	4-1, 5-2, 7-3, 2-4, 1-5, 1-6	1 - 5
2	~140 kb	3-7, 1-8, 1-9, 2-10, 3-11, 1-12, 1-13, 2-14, 3-15, 1-16	-
3	~161 kb	1-17, 2-18, 2-19, 3-20, 6-21, 1-22, 1-23, 2-24, 3-25, 2-26, 1-27, 2-28, 2-29, 2-30	-
4	~90 kb	3-31, 1-32, 1-33, 3-34, 1-35, 2-36, 1-37, 2-38, 1-39, 2-40	-

**Introducción de diez segmentos génicos  $V_{\kappa}$  humanos adicionales en un locus de cadena pesada híbrido.** Se diseñó por ingeniería un segundo vector de direccionamiento para la introducción de 10 segmentos génicos  $V_{\kappa}$  humanos adicionales en el locus de cadena pesada de ratón modificado que se ha descrito anteriormente (véase la fig. 2, Vector de direccionamiento 2). Se diseñó por ingeniería un fragmento genómico humano de 140.058 pb que contenía 12 segmentos génicos  $V_{\kappa}$  humanos consecutivos procedentes del locus de cadena ligera  $\kappa$  humano, con un brazo de homología 5' que contenía la secuencia genómica de ratón en 5' del locus de cadena pesada de ratón y un brazo de homología 3' que contenía una secuencia  $\kappa$  genómica humana. Aguas arriba (5') del segmento génico  $V_{\kappa}1-16$  humano, había una secuencia  $\kappa$  humana adicional de 10.170 pb, antes del inicio del brazo de homología de ratón 5', el cual era el mismo brazo de homología 5' utilizado para la construcción del Vector de direccionamiento 1 (véase la fig. 2). Entre el brazo de homología 5' y el inicio de la secuencia  $\kappa$  humana, había un casete de higromicina flanqueado por sitios de reconocimiento de recombinasa. El brazo de homología 3' incluía un solapamiento de 31.165 pb de la secuencia  $\kappa$  genómica humana correspondiente al extremo 5' equivalente del fragmento de ~110,5 kb de la secuencia  $\kappa$  genómica humana del Vector de direccionamiento 1 (fig. 2). El vector de direccionamiento final para la inserción de 10 segmentos génicos  $V_{\kappa}$  humanos adicionales, de 5' a 3', incluía un brazo de homología 5' que contenía ~20 kb de secuencia genómica de ratón en 5' del locus de la cadena pesada, un primer sitio de reconocimiento de recombinasa (R1), un casete de higromicina, un segundo sitio de reconocimiento de recombinasa (R2), y una secuencia  $\kappa$  genómica humana de ~140 kb, que contenía 12 segmentos génicos  $V_{\lambda}$  humanos consecutivos, ~31 kb de los cuales se solapan con el extremo 5' de la secuencia  $\kappa$  humana del Vector de direccionamiento 1 y sirven como brazo de homología 3' para esta construcción de direccionamiento. La recombinación homóloga con este vector de direccionamiento creó un locus de cadena pesada de ratón modificado que contenía 16 segmentos génicos  $V_{\kappa}$  humanos y cinco segmentos génicos  $J_{\kappa}$  humanos, unidos operativamente a los genes constantes de cadena pesada de ratón que, después de la recombinación, dio lugar a la formación de una cadena pesada híbrida.

**Introducción de catorce segmentos génicos  $V_{\kappa}$  humanos adicionales en un locus de cadena pesada híbrida.** Se diseñó por ingeniería un tercer vector de direccionamiento para la introducción de 14 segmentos génicos  $V_{\kappa}$  humanos adicionales en el locus de cadena pesada de ratón modificado que se ha descrito anteriormente (véase la fig. 2, Vector de direccionamiento 3). Se diseñó por ingeniería un fragmento genómico humano de 160.579 pb que contenía 15 segmentos génicos  $V_{\kappa}$  humanos consecutivos con un brazo de homología 5' que contenía la secuencia genómica de ratón en 5' del locus de cadena pesada de ratón y un brazo de homología 3' que contenía la secuencia  $\kappa$  genómica humana. Aguas arriba (5') del segmento génico  $V_{\kappa}2-30$  humano, había una secuencia  $\kappa$  humana adicional de 14.687 pb antes del inicio del brazo de homología murino 5', el cual era el mismo brazo de homología 5' utilizado para los dos vectores de direccionamiento anteriores (descritos anteriormente, véase también la fig. 2). Entre el brazo de homología 5' y el inicio de la secuencia  $\kappa$  humana, había un casete de neomicina flanqueado por sitios de reconocimiento de recombinasa. El brazo de homología 3' incluía un solapamiento de 21.275 pb de la secuencia  $\kappa$  genómica humana, correspondiente al extremo 5' equivalente del fragmento de ~140 kb de la secuencia  $\kappa$  genómica humana del Vector de direccionamiento 2 (fig. 2). El vector de direccionamiento final para la inserción de 14 segmentos génicos  $V_{\kappa}$  humanos adicionales, de 5' a 3', incluía un brazo de homología 5' que contenía ~20 kb de una secuencia genómica de ratón en 5' del locus de la cadena pesada de ratón, un primer sitio de reconocimiento de recombinasa (R1), un casete de neomicina, un segundo sitio de reconocimiento de recombinasa (R2), y ~161 kb de una secuencia  $\kappa$  genómica humana que contenía 15 segmentos génicos  $V_{\kappa}$  humanos, ~21 kb de los cuales se solapan con el extremo 5' de la secuencia  $\kappa$  humana del Vector de direccionamiento 2 y sirven como brazo de homología 3' para esta construcción de direccionamiento. La recombinación homóloga con este vector de

direccionamiento creó un locus de cadena pesada de ratón modificado que contenía 30 segmentos génicos  $V_{\kappa}$  humanos y cinco segmentos génicos  $J_{\kappa}$  humanos, unidos operativamente a los genes constantes de cadena pesada de ratón que, después de la recombinación, da lugar a la formación de una cadena pesada  $\kappa$  quimérica.

5 **Introducción de diez segmentos génicos  $V_{\kappa}$  humanos adicionales en un locus de cadena pesada híbrida.** Se diseñó por ingeniería un cuarto vector de direccionamiento para la introducción de 10 segmentos génicos  $V_{\kappa}$  humanos adicionales en el locus de cadena pesada de ratón modificado que se ha descrito anteriormente (véase la fig. 2, Vector de direccionamiento 4). Se diseñó por ingeniería un fragmento genómico humano de 90.398 pb, que contenía 16 segmentos génicos  $V_{\kappa}$  humanos consecutivos, con un brazo de homología 5' que contenía una  
 10 secuencia genómica de ratón en 5' del locus de cadena pesada de ratón, y un brazo de homología 3' que contenía una secuencia  $\kappa$  genómica humana. Aguas arriba (5') del segmento génico  $V_{\kappa}2-40$  humano, había una secuencia  $\kappa$  humana adicional de 8.484 pb antes del inicio del brazo de homología de ratón 5', el cual era el mismo brazo de homología 5' que el de los vectores de direccionamiento anteriores (anteriormente descritos, véase también la fig. 2). Entre el brazo de homología 5' y el inicio de la secuencia  $\kappa$  humana, había un casete de higromicina flanqueado por  
 15 sitios de reconocimiento de recombinasa. El brazo de homología 3' incluía un solapamiento de 61.615 pb de la secuencia  $\kappa$  genómica humana, que correspondía al extremo 5' equivalente del fragmento de ~160 kb de la secuencia  $\kappa$  genómica humana del Vector de direccionamiento 3 (fig. 2). El vector de direccionamiento final para la inserción de 10 segmentos génicos  $V_{\kappa}$  humanos adicionales, de 5' a 3', incluía un brazo de homología 5' que contenía ~20 kb de una secuencia genómica de ratón en 5' del locus de la cadena pesada de ratón, un primer sitio  
 20 de reconocimiento de recombinasa (R1), un casete de higromicina, un segundo sitio de reconocimiento de recombinasa (R2), y ~90 kb de una secuencia  $\kappa$  genómica humana que contenía 16 segmentos génicos  $V_{\kappa}$  humanos, ~62 kb de los cuales se solapan con el extremo 5' de la secuencia  $\kappa$  humana del Vector de direccionamiento 3, y sirven como el brazo de homología 3 para esta construcción de direccionamiento. La recombinación homóloga con este vector de direccionamiento creó un locus de cadena pesada de ratón modificado  
 25 que contenía 40 segmentos génicos  $V_{\kappa}$  humanos y cinco segmentos génicos  $J_{\kappa}$  humanos, unidos operativamente a los genes constantes de cadena pesada de ratón que, después de la recombinación, da lugar a la formación de una cadena pesada  $\kappa$  quimérica (fig. 3).

Usando un enfoque similar al que se ha descrito anteriormente, se construyeron otras combinaciones de dominios variables de cadena ligera humanos en el contexto de regiones constantes de cadena pesada de ratón. Se pueden  
 30 obtener dominios variables de cadena ligera adicionales a partir de segmentos génicos  $V_{\lambda}$  y  $J_{\lambda}$  humanos (fig. 4A y 4B).

El locus de la cadena ligera  $\lambda$  humana se extiende en más de 1.000 kb y contiene más de 80 genes que codifican  
 35 segmentos variables (V) o de unión (J). Entre los 70 segmentos génicos  $V_{\lambda}$  del locus de cadena ligera  $\lambda$  humana, cualquier fracción entre 30 y 38 parecen ser segmentos génicos funcionales de acuerdo con los informes publicados. Las 70 secuencias  $V_{\lambda}$  se disponen en tres grupos, todos los cuales contienen diferentes números de distintos grupos de familias génicas V (grupos A, B y C). En el locus de la cadena ligera  $\lambda$  humana, más de la mitad de todos los dominios  $V_{\lambda}$  observados están codificados por los segmentos génicos 1-40, 1-44, 2-8, 2-14, y 3-21. Existen siete  
 40 segmentos génicos  $J_{\lambda}$ , sólo cuatro de los cuales se consideran como segmentos génicos  $J_{\lambda}$  generalmente funcionales  $\downarrow$   $J_{\lambda}1$ ,  $J_{\lambda}2$ ,  $J_{\lambda}3$ , y  $J_{\lambda}7$ . En algunos alelos, se ha indicado un quinto par de segmentos génicos  $J_{\lambda}$ - $C_{\lambda}$ , como un pseudogen ( $C_{\lambda}6$ ). La incorporación de múltiples segmentos génicos  $J_{\lambda}$  humanos en un locus de cadena pesada híbrida, como se describe en el presente documento, se construye mediante síntesis *de novo*. De esta manera, un fragmento genómico que contiene múltiples segmentos génicos  $J_{\lambda}$  humanos en configuración de línea  
 45 germinal, se diseña por ingeniería con múltiples segmentos génicos  $V_{\lambda}$  humanos y permite la recombinación V-J normal en el contexto de una región constante de cadena pesada.

El acoplamiento de dominios variables de cadena ligera con regiones constantes de cadena pesada, representa una fuente potencialmente rica de diversidad para generar proteínas de unión  $V_L$  únicas con regiones  $V_L$  humanas en  
 50 animales no humanos. La explotación de esta diversidad del locus de cadena ligera  $\lambda$  humana (o del locus  $\kappa$  humano, como se ha descrito anteriormente) en ratones, da como resultado el diseño por ingeniería de cadenas pesadas híbridas únicas, y da lugar a otra dimensión de proteínas de unión al repertorio inmunitario de animales genéticamente modificados y su posterior uso como una plataforma de próxima generación para la generación de productos terapéuticos.

55 Adicionalmente, los segmentos génicos  $D_H$  y  $J_H$  (o  $J_{\kappa}$ ) humanos pueden incorporarse con cualquiera de los segmentos génicos  $V_{\kappa}$  o  $V_{\lambda}$  humanos para construir loci híbridos novedosos que darán lugar, después de la

recombinación, a nuevos dominios variables diseñados por ingeniería (fig. 5A y 5B). En este último caso, las combinaciones diseñadas por ingeniería de los segmentos génicos que normalmente no están contenidos en un único locus, requerirá una atención específica a las secuencias de señal de recombinación (RSS) que están asociadas a los respectivos segmentos génicos, de tal modo que se pueda lograr una recombinación normal cuando éstos se recombinen en un solo locus. Por ejemplo, se sabe que la recombinación V(D)J está guiada por secuencias de ADN no codificantes conservadas, conocidas como secuencias heptaméricas y nonaméricas que se encuentran adyacentes a cada segmento génico en la localización precisa en la que se produce la recombinación. Entre estas secuencias de ADN no codificantes, hay regiones separadoras no conservadas de 12 ó 23 pares de bases (pb) de longitud. Generalmente, la recombinación sólo ocurre en segmentos génicos localizados en el mismo cromosoma y los segmentos génicos flanqueados, por un separador de 12 pb pueden ser unidos a un segmento génico flanqueado por un separador de 23 pb, es decir, la regla 12/23, aunque se ha observado la unión de dos segmentos génicos  $D_H$  (cada uno flanqueado por separadores de 12 pb) en una pequeña proporción de anticuerpos. Para permitir la recombinación entre segmentos génicos que normalmente no tienen separadores compatibles (por ejemplo,  $V_K$  y un  $D_H$  o  $D_H$  y  $J_L$ ), se sintetizaron separadores únicos y compatibles en localizaciones adyacentes a los segmentos génicos deseados para la construcción de cadenas pesadas híbridas únicas que permiten una exitosa recombinación para formar cadenas pesadas únicas que contienen regiones variables de cadena ligera.

Por lo tanto, el uso de la estrategia que se ha señalado anteriormente para la incorporación de segmentos génicos de cadena ligera  $\kappa$  humanos en un locus de cadena pesada endógeno permite el uso de otras combinaciones de segmentos génicos de cadena ligera  $\lambda$  humana, así como segmentos génicos de cadena pesada humana específicos (por ejemplo,  $D_H$  y  $J_H$ ) y combinaciones de los mismos.

### Ejemplo II

#### 25 Identificación de células ME dirigidas portadoras de segmentos génicos de cadena ligera humana en un locus de cadena pesada endógeno

El ADN BAC dirigido que se preparó en los Ejemplos anteriores, se utilizó para electroporar células ME murinas para crear células ME modificadas para la generación de ratones quiméricos que expresen proteínas de unión  $V_L$  (es decir, segmentos génicos de cadena ligera  $\kappa$  humanos unidos operativamente a regiones constantes de cadena pesada de ratón). Las células ME que contenían una inserción de segmentos génicos de cadena ligera  $\kappa$  humana no reordenados se identificaron mediante el ensayo de PCR cuantitativa, TAQMAN® (Lie y Petropoulos, 1998. Curr. Opin. Biotechnology 9: 43-48). Se diseñaron conjuntos de cebadores específicos y sondas para la inserción de secuencias  $\kappa$  humanas y los casetes de selección asociados, la pérdida de secuencias de cadena pesada de ratón y la retención de secuencias de ratón que flanqueaban el locus de cadena pesada endógeno.

Las células ME portadoras de los segmentos génicos de cadena ligera  $\kappa$  humanos pueden transfectarse con una construcción que exprese una recombinasa con el fin de eliminar cualquier casete de selección no deseado que se haya introducido por la inserción de la construcción de direccionamiento que contiene los segmentos génicos  $\kappa$  humanos. Opcionalmente, el casete de selección puede eliminarse al aparear los animales con ratones que expresen la recombinasa (por ejemplo, documento US 6.774.279). Opcionalmente, el casete de selección se retiene en el ratón.

### Ejemplo III

#### 45 Generación y análisis de ratones que expresan proteínas de unión $V_L$

Las células ME dirigidas que se han descrito anteriormente se usaron como células ME donantes y se introdujeron en embriones de ratón en estadio de 8 células mediante el método VELOCIMOUSE® (véanse por ejemplo, Pat. de Estados Unidos n.º 7.294.754 y Poueymirou, W.T., Auerbach, W., Friendewey, D., Hickey, J.F., Escaravage, J.M., Esau, L., Dore, A.T., Stevens, S., Adams, N.C., Dominguez, M.G., Gale, N.W., Yancopoulos, G.D., DeChiara, T.M., Valenzuela, D.M. (2007). F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. Nat Biotechnol 25, 91-99). Se identificaron segmentos génicos  $\kappa$  humanos que portaban independientemente VELOCIMICE® (ratones F0 totalmente derivados de la célula ME donante) en el locus de cadena pesada de ratón mediante una genotipificación utilizando una modificación del ensayo de alelos (Valenzuela et al., *supra*), que detectó la presencia de los segmentos génicos  $\kappa$  humanos únicos en el locus de cadena pesada endógena (*supra*). Las crías se genotifican, y se selecciona una cría heterocigota para el locus del gen de cadena pesada híbrido para caracterizar la expresión de proteínas de unión  $V_L$ .

**Citometría de Flujo.** La introducción de segmentos génicos de cadena ligera  $\kappa$  humana en locus de cadena pesada de ratón se realizó en una línea ME F1 (F1H4; Valenzuela et al. 2007, *supra*) derivada de embriones heterocigotos 129S6/SvEvTac y C57BL/6NTac que contenían además un reemplazo *in situ* de los segmentos génicos de cadena ligera  $\kappa$  de ratón por segmentos génicos de cadena ligera  $\kappa$  humana (documento US 6.596.541). Los segmentos génicos variables de línea germinal de cadena ligera  $\kappa$  humana se dirigen al alelo 129S6, que es portador del haplotipo IgM<sup>a</sup>, mientras que el alelo C576BL/6N de ratón no modificado es portador del haplotipo IgM<sup>b</sup>. Estas formas alélicas de IgM pueden distinguirse por citometría de flujo, utilizando anticuerpos específicos para los polimorfismos encontrados en los alelos IgM<sup>a</sup> o IgM<sup>b</sup>. Los ratones heterocigotos portadores de segmentos génicos de cadena ligera  $\kappa$  humana en el locus de cadena pesada endógeno, como se describió en el Ejemplo I, se evaluaron con respecto a la expresión de proteínas de unión V<sub>L</sub> humanas usando citometría de flujo.

Brevemente, se extrajo sangre de grupos de ratones ( $n = 6$  por grupo) y se mezcló usando portaobjetos de vidrio. Se usaron ratones C57BL/6 y Balb/c como grupos de control. Tras la lisis de glóbulos rojos (RBC) con tampón de lisis ACK (Lonza Walkersville), las células se resuspendieron en tampón de tinción BD Pharmingen FACS y se bloquearon con anti-CD16/32 de ratón (BD Pharmingen). Los linfocitos se tiñeron con anti-IgM<sup>b</sup> de ratón-FITC (BD Pharmingen), anti-IgM<sup>a</sup> de ratón-PE (BD Pharmingen), anti-CD19 de ratón (Clon 1D3; BD Biosciences), y anti-CD3 de ratón (17A2; BIOLEGEND®) seguido de fijación con BD CYTOFIX™, todo de acuerdo con las instrucciones del fabricante. Los sedimentos celulares finales se resuspendieron en tampón de tinción y se analizaron usando un software BD FACSCALIBUR™ y BD CELLQUEST PRO™. La Tabla 2 expone los valores porcentuales medios para la expresión de células B (CD19<sup>+</sup>), células T (CD3<sup>+</sup>), cadena pesada híbrida (CD19<sup>+</sup>IgM<sup>hi</sup>) y cadena pesada de tipo silvestre (CD19<sup>+</sup>IgM<sup>hi</sup>) observada en grupos de animales portadores de cada modificación genética.

En un experimento similar, el contenido de células B de los compartimentos del bazo, sangre y médula ósea de ratones homocigotos para seis segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos, unidos operativamente a la región constante de cadena pesada de ratón (que se describen en el Ejemplo I, fig. 2), se analizó con respecto a la progresión a través del desarrollo de células B, usando citometría de flujo de diversos marcadores de superficie celular.

Brevemente, dos grupos ( $n = 3$  cada uno, hembras de 8 semanas de edad) de ratones de tipo silvestre y homocigotos para seis segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos, unidos operativamente a la región constante de cadena pesada de ratón, se sacrificaron y se les extrajo sangre, el bazo y la médula ósea. La sangre se recogió en tubos microtainer con EDTA (BD Biosciences). La médula ósea se recogió de los fémures, mediante un lavado con medio RPMI completo (medio RPMI complementado con suero fetal de ternera, piruvato sódico, HEPES, 2-mercaptoetanol, aminoácidos no esenciales y gentamicina). Los RBC de preparaciones del bazo y de médula ósea se lisaron con tampón de lisis ACK (Lonza Walkersville), seguido de un lavado con medio RPMI completo.

Se incubaron células ( $1 \times 10^6$ ) con anti-CD16/CD32 de ratón (2.4G2, BD) sobre hielo durante diez minutos, seguido de marcado con el siguiente cóctel de anticuerpos durante treinta minutos sobre hielo: anti-CD43 de ratón-FITC (1B11, BIOLEGEND®), PE-ckit (2B8, BIOLEGEND®), PeCy7-IgM (II/41, EBIOSCIENCE®), PerCP-Cy5.5-IgD (11-26c.2a, BIOLEGEND®), APC-eFluor 780-B220 (RA3-6B2, EBIOSCIENCE®), APC-CD19 (MB19-1, EBIOSCIENCE®). Médula ósea: células B inmaduras (B220<sup>int</sup>IgM<sup>+</sup>), células B maduras (B220<sup>hi</sup>IgM<sup>+</sup>), células pro-B (CD19<sup>+</sup>ckit<sup>+</sup>CD43<sup>-</sup>), células pre-B (CD19<sup>+</sup>ckit<sup>+</sup>CD43<sup>-</sup>), células pre-B (CD19<sup>+</sup>CD43<sup>int</sup>IgM<sup>hi</sup>), células B inmaduras (CD19<sup>+</sup>CD43<sup>-</sup>IgM<sup>hi</sup>). Sangre y bazo: células B (CD19<sup>+</sup>), células B maduras (CD19<sup>+</sup>IgM<sup>int</sup>IgD<sup>hi</sup>), células B de transición/inmaduras (CD19<sup>+</sup>IgM<sup>hi</sup>IgD<sup>int</sup>).

Después de la tinción, las células se lavaron y se fijaron en formaldehído al 2%. La adquisición de datos se realizó en un citómetro de flujo LSRII y se analizó con el software FLOWJO™ (Tree Star, Inc.). Las fig. 6A, 6B y 6C muestran los resultados para el compartimento esplénico. Las fig. 7A-7G muestran los resultados para el compartimento de médula ósea. Los resultados obtenidos para el compartimento de sangre de cada grupo de ratones demostraron resultados similares en comparación con el compartimento esplénico de cada grupo (datos no mostrados).

En un experimento similar, se analizó el contenido de células B de los compartimentos del bazo, sangre y médula ósea de ratones homocigotos para treinta segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos, unidos operativamente a la región constante de cadena pesada de ratón (descritos en el Ejemplo I, fig. 2), con respecto a la progresión a través del desarrollo de células B usando citometría de flujo de diversos marcadores de superficie celular.



Brevemente, dos grupos de ratones (N = 3 cada uno, hembras de 6 semanas de edad) que contenían un locus de cadena pesada de tipo silvestre y un reemplazo de los segmentos génicos V $\kappa$  y J $\kappa$  endógenos por segmentos génicos V $\kappa$  y J $\kappa$  humanos (TS), y ratones homocigotos para treinta segmentos génicos hV $\kappa$  y cinco segmentos génicos J $\kappa$ , y un reemplazo de los segmentos génicos V $\kappa$  y J $\kappa$  endógenos por segmentos génicos V $\kappa$  y J $\kappa$  humanos (30hV $\kappa$ -5hJ $\kappa$  HO) se sacrificaron y se obtuvieron los bazos y médula ósea. Se prepararon la médula ósea y los esplenocitos para una tinción con diversos marcadores de superficie celular (como se ha descrito anteriormente).

Se incubaron las células (1 x 10<sup>6</sup>) con anti-CD16/CD32 de ratón (2.4G2, BD Biosciences) sobre hielo durante diez minutos, seguido de marcado con paneles de médula ósea o esplenocitos durante treinta minutos sobre hielo. Panel de médula ósea: anti-CD43 de ratón-FITC (1B11, BIOLEGEND®), PE-ckit (2B8, BIOLEGEND®), PeCy7-IgM (II/41, EBIOSCIENCE®), APC-CD19 (MB19-1, EBIOSCIENCE®). Panel de médula ósea y bazo: FITC-Ig $\kappa$  anti-ratón (187.1 BD Biosciences), PE-Ig $\lambda$  (RML-42, BIOLEGEND®), PeCy7-IgM (II/41, EBIOSCIENCE®), PerCP-Cy5.5-IgD (11-26c.2a, BIOLEGEND®), Pacific Blue-CD3 (17A2, BIOLEGEND®), APC-B220 (RA3-6B2, EBIOSCIENCE®), APC-H7-CD19 (ID3, BD). Médula ósea: células B inmaduras (B220<sup>int</sup>IgM<sup>+</sup>), células B maduras (B220<sup>hi</sup>IgM<sup>+</sup>), células pro-B (CD19<sup>+</sup>ckit<sup>+</sup>CD43<sup>+</sup>), células pre-B (CD19<sup>+</sup>ckit-CD43-), células B Ig $\kappa$ <sup>+</sup> inmaduras (B220<sup>int</sup>IgM<sup>+</sup>Ig $\kappa$ <sup>+</sup>Ig $\lambda$ <sup>-</sup>), células B Ig $\lambda$ <sup>+</sup> inmaduras (B220<sup>int</sup>IgM<sup>+</sup>Ig $\kappa$ <sup>-</sup>Ig $\lambda$ <sup>+</sup>), células B Ig $\kappa$ <sup>+</sup> maduras (B220<sup>hi</sup>IgM<sup>+</sup>Ig $\kappa$ <sup>+</sup>Ig $\lambda$ <sup>-</sup>), células B Ig $\lambda$ <sup>+</sup> maduras (B220<sup>hi</sup>IgM<sup>+</sup>Ig $\kappa$ <sup>-</sup>Ig $\lambda$ <sup>+</sup>). Bazo: células B (CD19<sup>+</sup>), células B maduras (CD19<sup>+</sup>IgD<sup>hi</sup>IgM<sup>int</sup>), células B de transición/inmaduras (CD19<sup>+</sup>IgD<sup>int</sup>IgM<sup>hi</sup>). Médula ósea y bazo: células B Ig $\kappa$ <sup>+</sup> (CD19<sup>+</sup>Ig $\kappa$ <sup>+</sup>Ig $\lambda$ <sup>-</sup>), células B Ig $\lambda$ <sup>+</sup> (CD19<sup>+</sup>Ig $\kappa$ <sup>-</sup>Ig $\lambda$ <sup>+</sup>).

Después de la tinción, las células se lavaron y se fijaron con formaldehído al 2%. La adquisición de datos se realizó en un citómetro de flujo LSRII y se analizó con el software FLOWJO™ (Tree Star, Inc.). Los resultados demostraron patrones de tinción y poblaciones de células similares para los tres compartimentos en comparación con los ratones homocigotos para seis segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos (descritos anteriormente). Sin embargo, estos ratones demostraron una pérdida de expresión de cadena ligera  $\lambda$  endógena, tanto en el compartimiento esplénico como en el de médula ósea (fig. 8A y 8B, respectivamente), a pesar de que el locus de cadena ligera  $\lambda$  endógeno está intacto en estos ratones. Esto podría reflejar una incapacidad de los dominios de cadena ligera  $\kappa$  humanos reordenados, en el contexto de regiones constantes de cadena pesada, para emparejarse o asociarse con dominios de cadena ligera  $\lambda$  murinos, lo que conduce a la delección de células Ig $\lambda$ <sup>+</sup>.

**Expresión de isotipo.** Se determinó la inmunoglobulina M (IgM) y la inmunoglobulina G1 (IgG1) total y de superficie (es decir, unida a la membrana), en ratones homocigotos para loci humanos de genes variables de cadena pesada y de cadena ligera  $\kappa$  (VELCOIMMUNE® Humanized Mice, véase el documento US 7.105.348) y ratones homocigotos para seis segmentos génicos V $\kappa$  humanos y 5 segmentos génicos J $\kappa$  humanos, diseñados por ingeniería en el locus de cadena pesada endógeno ((6hV $\kappa$ -5hJ $\kappa$  HO), mediante un ensayo de PCR cuantitativa utilizando sondas TAQMAN® (como se ha descrito anteriormente en el Ejemplo II).

Brevemente, las células B CD19<sup>+</sup> se purificaron a partir de los bazos de grupos de ratones (n = 3 a 4 ratones por grupo), utilizando microperlas CD19 de ratón (Miltenyi Biotec), de acuerdo con las instrucciones del fabricante. El ARN total se purificó utilizando el Mini kit RNEASY™ (Qiagen). El ARN genómico se eliminó usando un tratamiento en columna de ADNasa sin ARNasa (Qiagen). Se sometieron a transcripción inversa aproximadamente 200 ng de ARNm para obtener ADNc usando el kit de síntesis de ADNc de primera cadena (Invitrogen) y después se amplificó con la mezcla máster de PCR universal TAQMAN® (Applied Biosystems), usando el sistema de detección de secuencia ABI 7900 (Applied Biosystems). Se emplearon combinaciones únicas de cebador/sonda para determinar específicamente la expresión de las formas total, de superficie (es decir, transmembrana) y secretadas de los isotipos IgM e IgG1 (Tabla 3). La expresión relativa se normalizó con respecto a la región constante  $\kappa$  de ratón (mC $\kappa$ ).

Tabla 2

Genotipo de ratón	% de CD3	% de CD19	% de IgM <sup>a</sup>	% de IgM <sup>b</sup>
C57BL/6	22	63	0	100
Balb/c	11	60	100	0
6hV $\kappa$ -5hJ $\kappa$ HET	43	30	7	85
16hV $\kappa$ -5hJ $\kappa$ HET	33	41	7	81

50

Tabla 3

Isotipo	Secuencia (5'-3')	SEQ ID NOs:
IgM de superficie	sentido: GAGAGGACCG TGGACAAGTC	1
	antisentido: TGACGGTGGT GCTGTAGAAG	2
	sonda: ATGCTGAGGA GGAAGGCTTT GAGAACCT	3
IgM total	sentido: GCTCGTGAGC AACTGAACCT	4
	antisentido: GCCACTGCAC ACTGATGTC	5
	sonda: AGTCAGCCAC AGTCACCTGC CTG	6
IgG1 de superficie	sentido: GCCTGCACAA CCACCATAC	7
	antisentido: GAGCAGGAAG AGGCTGATGA AG	8
	sonda: AGAAGAGCCT CTCCCCTCT CCTGG	9
IgG1 total	sentido: CAGCCAGCGG AGAACTACAA G	10
	antisentido: GCCTCCCAGT TGCTCTTCTG	11
	sonda: AACACTCAGC CCATCATGGA CACA	12
C $\kappa$	sentido: TGAGCAGCAC CCTCACGTT	13
	antisentido: GTGGCCTCAC AGGTATAGCT GTT	14
	sonda: ACCAAGGACG AGTATGAA	15

Los resultados del ensayo de PCR TAQMAN® cuantitativo demostraron una disminución de la IgM total y de la IgG1 total. Sin embargo, la relación de las formas secretada frente a superficial de IgM e IgG 1 pareció normal en 5 comparación con los ratones humanizados VELCOIMMUNE® (datos no mostrados).

**Análisis del uso del segmento génico  $\kappa$  humano y la unión V $\kappa$ -J $\kappa$ .** Ratones sin tratar homocigotos para treinta segmentos génicos hV $\kappa$  y cinco segmentos génicos J $\kappa$  y un reemplazo de los segmentos génicos V $\kappa$  y J $\kappa$  endógenos por segmentos génicos V $\kappa$  y J $\kappa$  humanos (30hV $\kappa$ -5hJ $\kappa$  HO) se analizaron con respecto a 10 reordenamientos V $\kappa$ - J $\kappa$  humanas únicas en la cadena pesada de ratón (IgG), mediante una reacción en cadena de polimerasa de transcripción reversa (RT-PCR) usando ARN aislado de esplenocitos.

Brevemente, se recogieron los bazos y se perfundieron con 10 ml de RPMI-1640 (Sigma) con HI-FBS al 5% en 15 bolsas desechables estériles. Cada bolsa que contenía un único bazo se colocó en un STOMACHER™ (Seward) y se homogeneizó a una configuración media durante 30 segundos. Los bazos homogeneizados se filtraron utilizando un filtro celular de 0,7  $\mu$ m y después se sedimentaron con una centrifuga (1.000 rpm durante 10 minutos) y los RBC se lisaron en BD PHARM LYSE™ (BD Biosciences) durante tres minutos. Los esplenocitos se diluyeron con RPMI-1640 y se centrifugaron de nuevo, seguido de resuspensión en 1 ml de PBS (Irvine Scientific). El ARN se aisló de los 20 esplenocitos sedimentados usando técnicas convencionales conocidas en la técnica.

La RT-PCR se realizó en ARN de esplenocitos usando cebadores específicos para segmentos génicos hV $\kappa$  humanos y la IgG de ratón. El cebador de IgG de ratón se diseñó de tal forma que fuera capaz de amplificar ARN obtenido de todos los isotipos de IgG de ratón. Los productos de las PCR se purificaron en gel y se clonaron en el vector pCR2.1-TOPO TA (Invitrogen) y se secuenciaron con cebadores M13 directos (GTAAAACGAC GGCCAG; 25 SEQ ID NO: 16) y M13 inversos (CAGGAAACAG CTATGAC; SEQ ID NO: 17) localizados en el vector en localizaciones que flanqueaban el sitio de clonación. El uso de los segmentos génicos V $\kappa$  y J $\kappa$  humanos entre doce clones seleccionados se muestra en la Tabla 4. La fig. 9 muestra la secuencia nucleotídica de la unión hV $\kappa$ -hJ $\kappa$ -mIgG para los doce clones RT-PCR seleccionados.

30 Como se muestra en este Ejemplo, los ratones homocigotos para seis segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos, u homocigotos para treinta segmentos génicos V $\kappa$  humanos y cinco segmentos génicos J $\kappa$  humanos, unidos operativamente a la región constante de cadena pesada de ratón, demostraron la expresión de dominios variables de cadena ligera humanos a partir de un locus de cadena pesada modificado que contenía segmentos génicos variables de cadena ligera en su configuración de línea germinal. Se observó en estos 35 ratones la progresión a través de los diversos estadios del desarrollo de células B, indicando múltiples eventos de recombinación productiva que involucraron a los segmentos génicos variables de cadena ligera de un locus de cadena pesada endógeno y la expresión de tales cadenas pesadas híbridas (es decir, la región variable de cadena ligera humana unida a una región constante de cadena pesada), como parte del repertorio de anticuerpos.

Tabla 4

Clon	Cadena pesada híbrida			SEQ ID NO:
	V $\kappa$	J $\kappa$	C $_H$	
1E	1-5	4	IgG2A/C	18
1G	1-9	4	IgG2A/C	19
1A	1-16	5	IgG3	20
2E	1-12	2	IgG1	21
1C	1-27	4	IgG2A/C	22
2H	2-28	1	IgG1	23
3D	3-11	4	IgG1	24
3A	3-20	4	IgG2A/C	25
4B	4-1	5	IgG2A/C	26
4C	4-1	2	IgG3	27
5A	5-2	2	IgG2A/C	28
5D	5-2	1	IgG1	29

**Ejemplo IV**

**Propagación de ratones que expresan proteínas de unión V $_L$**

5

Para crear una nueva generación de proteínas de unión V $_L$ , los ratones portadores de los segmentos génicos  $\kappa$  humanos no reordenados se pueden aparear con otro ratón que contiene una delección del otro alelo de cadena pesada endógeno. De esta manera, la progenie obtenida expresará solamente cadenas pesadas híbridas como las descritas en el Ejemplo I. La cría se realizó mediante técnicas estándar reconocidas en la técnica y, como alternativa, por empresas comerciales, por ejemplo, The Jackson Laboratory. Las cepas de ratones portadores de un locus de cadena pesada híbrida se criaron con respecto a la presencia de las cadenas pesadas híbridas únicas y la ausencia de cadenas pesadas murinas tradicionales.

10

Como alternativa, los ratones portadores de segmentos génicos  $\kappa$  humanos no reordenados en el locus de la cadena pesada de ratón pueden optimizarse al aparearlos con otros ratones que contienen una o más delecciones de los loci de cadena ligera murina ( $\kappa$  y  $\lambda$ ). De esta manera, la progenie obtenida expresará únicamente anticuerpos únicamente con cadena pesada  $\kappa$  humana únicos, como los descritos en el Ejemplo I. De manera similar, la cría se realizó mediante técnicas estándar reconocidas en la técnica y, como alternativa, por empresas comerciales, por ejemplo, The Jackson Laboratory. Las cepas de ratón portadoras de un locus de cadena pesada híbrida y una o más delecciones de los loci de cadena ligera de ratón se criaron con respecto a la presencia de las cadenas pesadas híbridas únicas que contenían dominios de cadena ligera  $\kappa$  humanos y dominios constantes de cadena pesada de ratón, y la ausencia de cadenas ligeras de ratón endógenas.

15

20

Los ratones portadores de un locus de cadena pesada híbrido no reordenado también se aparean con ratones que contienen un reemplazo del locus génico variable de cadena ligera  $\kappa$  de ratón endógeno por el locus génico variable de cadena ligera  $\kappa$  humano (véase el documento US 6.596.541, Regeneron Pharmaceuticals, The VELOCIMMUNE® Humanized Mouse Technology). El paquete VELOCIMMUNE® Humanized Mouse incluye, en parte, tener un genoma que comprende regiones variables de cadena ligera  $\kappa$  humanas unidas operativamente a loci de región constante variable de cadena ligera  $\kappa$  de ratón endógeno, de tal modo que el ratón produzca anticuerpos que comprendan un dominio variable de cadena ligera  $\kappa$  humano y un dominio constante de cadena pesada de ratón, en respuesta a una estimulación antigénica. El ADN que codifica las regiones variables de las cadenas ligeras de los anticuerpos puede aislarse y unirse operativamente a ADN que codifica las regiones constantes de cadena ligera humana. Después, el ADN puede expresarse en una célula capaz de expresar la cadena ligera completamente humana del anticuerpo. Después de un programa de cría y apareamiento adecuado, se obtuvieron

30

ratones portadores de un reemplazo de la cadena ligera  $\kappa$  de ratón endógena por el locus de cadena ligera  $\kappa$  humana, y un locus de cadena pesada híbrida no reordenado. Las proteínas de unión  $V_L$  únicas que contienen dominios  $V\kappa$  humanos somáticamente mutados se pueden aislar después de una inmunización con un antígeno de interés.

5

### Ejemplo V

#### Generación de proteínas de unión $V_L$

10 Después del apareamiento de ratones que contienen el locus de cadena pesada híbrida no reordenado con diversas cepas deseadas que contienen modificaciones y deleciones de otros loci de Ig endógenos (como se describe en el Ejemplo IV), los ratones seleccionados pueden inmunizarse con un antígeno de interés.

15 Generalmente, un ratón humanizado VELOCIMMUNE® que contiene al menos un locus de cadena pesada híbrida se expone a un antígeno, y las células (tales como las células B) se recuperan del animal (por ejemplo, del bazo o de los ganglios linfáticos). Las células pueden fusionarse con una línea celular de mieloma, para preparar líneas celulares de hibridoma inmortales, y tales líneas celulares de hibridoma se criban y se seleccionan para identificar líneas de células de hibridoma que produzcan anticuerpos que contengan cadenas pesadas híbridas específicas para el antígeno utilizado para la inmunización. El ADN que codifica las regiones  $V\kappa$  humanas de las cadenas pesadas híbridas puede aislarse y unirse a las regiones constantes deseables, por ejemplo, de cadena pesada y/o de cadena ligera. Debido a la presencia de segmentos génicos  $V\kappa$  humanos fusionados a regiones constantes de cadena pesada de ratón, se produce un repertorio único de moléculas similares a anticuerpos y la diversidad del repertorio de inmunoglobulinas se aumenta drásticamente como resultado del formato de anticuerpo único creado. Esto contiene un nivel añadido de diversidad al repertorio específico contra el antígeno después de la inmunización.

25 Las secuencias de anticuerpo clonadas resultantes, posteriormente pueden producirse en una célula, tal como una célula CHO. Como alternativa, el ADN que codifica las proteínas de unión  $V_L$  específicas de antígeno o los dominios variables, se puede aislar directamente de los linfocitos específicos de antígeno (por ejemplo, células B).

30 Inicialmente, se aíslan proteínas de unión  $V_L$  de alta afinidad que tienen una región  $V\kappa$  humana y una región constante murina. Como se ha descrito anteriormente, las proteínas de unión  $V_L$  se caracterizan y seleccionan para características deseables, incluyendo la afinidad, selectividad, epítipo, etc. Las regiones constantes murinas se reemplazan por una región constante humana deseada, para generar proteínas de unión  $V_L$  completamente humanas únicas, que contienen dominios  $V\kappa$  humanos somáticamente mutados procedentes de un locus de cadena pesada híbrida no reordenado de la invención. Las regiones constantes humanas adecuadas incluyen, por ejemplo, IgG1 o IgG4 de tipo silvestre o modificadas o, como alternativa,  $C\kappa$  o  $C\lambda$ .

40 Cohortes separadas de ratones que contenían un reemplazo del locus de cadena pesada de ratón endógeno por seis segmentos génicos  $V\kappa$  humanos y cinco segmentos génicos  $J\kappa$  humanos (como se describe en el Ejemplo I) y un reemplazo de los segmentos génicos  $V\kappa$  y  $J\kappa$  endógenos por segmentos génicos  $V\kappa$  y  $J\kappa$  humanos, se inmunizaron con una proteína receptora de superficie celular humana (antígeno X). El antígeno X se administra directamente en la almohadilla plantar de las patas traseras de los ratones con seis inyecciones consecutivas, cada 3-4 días. De dos a tres microgramos del antígeno X se mezclan con 10  $\mu\text{g}$  de oligonucleótido CpG (Cat. n.º tIrl-modn - oligonucleótido ODN1826; InVivogen, San Diego, CA) y 25  $\mu\text{g}$  de Adju-Phos (adyuvante de gel de fosfato de aluminio Cat. n.º H-71639-250; Brenntag Biosector, Frederikssund, Dinamarca) antes de la inyección. Se administran 45 un total de seis inyecciones antes del refuerzo final del antígeno, que se proporciona 3-5 días antes del sacrificio. Se recogen las extracciones de sangre después de la 4ª y 6ª inyección, y la respuesta inmunitaria de anticuerpos se monitoriza mediante un inmunoensayo específico de antígeno estándar.

50 Cuando se logra una respuesta inmunitaria deseada, los esplenocitos se recogen y se fusionan con células de mieloma murino para conservar su viabilidad y para formar líneas celulares de hibridoma. Las líneas celulares de hibridoma se criban y se seleccionan para identificar líneas celulares que produzcan proteínas de unión  $V_L$  específicas de antígeno X. Con el uso de esta técnica, se obtienen varias proteínas de unión  $V_L$  anti-antígeno X específicas (es decir, proteínas de unión que poseían dominios  $V\kappa$  humanos en el contexto de dominios constantes de cadena pesada y ligera murinos).

55

Como alternativa, se aíslan proteínas de unión  $V_L$  anti-antígeno X directamente de las células B positivas para el antígeno, sin fusión con células de mieloma, como se describe en el documento U.S. 2007/0280945A1. Con el uso de este método, se obtuvieron varias proteínas de unión  $V_L$  anti-antígeno X completamente humanas (es decir,

anticuerpos que poseían dominios V $\kappa$  humanos y dominios constantes humanos).

**Uso del segmento génico  $\kappa$  humano.** Para analizar la estructura de las proteínas de unión V<sub>L</sub> anti-antígeno X producidas, los ácidos nucleicos que codifican los dominios V $\kappa$  humanos (tanto de las cadenas pesadas como de las ligeras de la proteína de unión V<sub>L</sub>), se clonaron y se secuenciaron usando métodos adaptados a partir de los descritos en el documento US 2007/0280945A1 (*supra*). A partir de las secuencias de ácido nucleico y de las secuencias de aminoácidos predichas de los anticuerpos, se identificó el uso de los genes para la región variable de cadena pesada híbrida de las proteínas de unión V<sub>L</sub> seleccionadas obtenidas de ratones inmunizados (descrito anteriormente). La Tabla 5 presenta el uso de los genes de los segmentos génicos V $\kappa$  y J $\kappa$  humanos procedentes de proteínas de unión V<sub>L</sub> anti-antígeno X seleccionadas, lo que demuestra que los ratones de acuerdo con la invención, generan proteínas de unión V<sub>L</sub> específicas de antígeno, a partir de una diversidad de segmentos génicos V $\kappa$  y J $\kappa$  humanos, debido a una diversidad de reordenamientos en los loci de cadena pesada y los loci de cadena ligera  $\kappa$  endógenos, tanto aquellos que contenían segmentos génicos V $\kappa$  como aquellos contenían segmentos génicos J $\kappa$  humanos no reordenados. Los segmentos génicos V $\kappa$  humanos se reordenaron con una diversidad de segmentos J $\kappa$  humanos, para producir proteínas de unión V<sub>L</sub> específicas de antígeno únicas.

Tabla 5

Anticuerpo	Cadena pesada híbrida		Cadena ligera	
	V $\kappa$	J $\kappa$	V $\kappa$	J $\kappa$
A	4-1	3	3-20	1
B	4-1	3	3-20	1
C	4-1	3	3-20	1
D	4-1	3	3-20	1
E	4-1	3	3-20	1
F	4-1	3	3-20	1
G	4-1	3	3-20	1
H	4-1	3	3-20	1
I	4-1	3	3-20	1
J	1-5	3	1-33	3
K	4-1	3	3-20	1
L	4-1	3	1-9	3
M	4-1	1	1-33	4
N	4-1	1	1-33	3
O	1-5	1	1-9	2
P	1-5	3	1-16	4
Q	4-1	3	3-20	1
R	4-1	3	3-20	1
S	1-5	1	1-9	2
T	1-5	1	1-9	2
U	5-2	2	1-9	3
V	1-5	2	1-9	2
W	4-1	1	1-33	4

**Ensayo inmunoabsorbente ligando a enzimas (ELISA).** Las proteínas de unión V<sub>L</sub> humanas producidas contra el

antígeno X se ensayaron para determinar su capacidad de bloquear la unión del ligando natural del antígeno X (ligando Y), en un ensayo de ELISA.

Brevemente, se recubrieron placas de 96 pocillos con el ligando Y a una concentración de 2 µg/ml, diluido en PBS y se incubó durante una noche seguido de cuatro lavados con PBS con Tween-20 al 0,05%. Después, la placa se bloqueó con PBS (Irvine Scientific, Santa Ana, CA) que contenía BSA al 0,5% (p/v) (Sigma-Aldrich Corp., St. Louis, MO) durante una hora a temperatura ambiente. En una placa separada, los sobrenadantes que contenían proteínas de unión  $V_L$  anti-antígeno X se diluyeron 1:10 en tampón. Se usó un sobrenadante simulado con los mismos componentes de las proteínas de unión  $V_L$  como control negativo. El dominio extracelular (ECD) del antígeno X se conjugó con la porción Fc de la IgG2a murina (antígeno X-mFc). Se añadió antígeno X-mFc a una concentración final de 0,150 nM y se incubó durante una hora a temperatura ambiente. Después, la mezcla proteína de unión  $V_L$ /antígeno X-mFc se añadió a la placa que contenía el Ligando Y, y se incubó durante una hora a temperatura ambiente. La detección del antígeno X-mFc unido al ligando Y se determinó con peroxidasa de rábano picante (HRP) conjugada con un anticuerpo anti-Penta-His (Qiagen, Valencia, CA) y se reveló mediante una respuesta colorimétrica estándar, usando el sustrato tetrametilbenzidina (TMB) (BD Biosciences, San Jose, CA) neutralizado con ácido sulfúrico. La absorbancia se leyó a una DO450 durante 0,1 s. La absorbancia de fondo de una muestra sin el antígeno X, se restó de todas las muestras. El porcentaje de bloqueo se calculó para >250 (tres placas de 96 pocillos) proteínas de unión  $V_L$  específicas de antígeno X, dividiendo el MFI restado del fondo de cada muestra por el valor del control negativo ajustado, multiplicando por 100 y restando el valor resultante de 100.

Los resultados mostraron que varias proteínas de unión  $V_L$  aisladas de ratones inmunizados con el antígeno X, se unieron específicamente al dominio extracelular del antígeno X fusionado a la porción Fc de la IgG2a murina (datos no mostrados).

**25 Determinación de afinidad.** Se determinaron las constantes de disociación en equilibrio ( $K_D$ ) para los sobrenadantes de proteína de unión  $V_L$  específicas contra el antígeno X seleccionados mediante SPR (Resonancia de plasmón superficial), usando un instrumento BIACORE™ T100 (GE Healthcare). Todos los datos se obtuvieron utilizando HBS-EP (HEPES 10 mM, NaCl 150 mM, EDTA 0,3 mM, tensioactivo P20 al 0,05%, pH 7,4) tanto como los tampones de realización y de muestra, a 25°C.

Brevemente, se capturaron las proteínas de unión  $V_L$  a partir de muestras de sobrenadante en bruto en una superficie de chip sensor CM5 previamente derivatizada con una alta densidad de anticuerpos anti-Fc humano usando la química de acoplamiento amina estándar. Durante la etapa de captura, los sobrenadantes se inyectaron por la superficie del anti-Fc humano a un caudal de 3 µl/min, durante un total de 3 minutos. La etapa de captura se siguió de una inyección del tampón de realización o analito a una concentración de 100 nM durante 2 minutos, a un caudal de 35 µl/min. La disociación del antígeno de la proteína de unión  $V_L$  capturada se monitorizó durante 6 minutos. La proteína de unión  $V_L$  capturada se retiró mediante una breve inyección de glicina 10 mM, pH 1,5. Todos los sensogramas fueron referenciados por duplicado, restando los sensogramas de las inyecciones de tampón de los sensogramas de analitos, eliminando así los artefactos causados por la disociación de la proteína de unión  $V_L$  de la superficie de captura. Los datos de unión para cada proteína de unión  $V_L$ , se ajustaron a un modelo de unión 1:1 con un transporte de masa usando el software de evaluación BIAcore T100 v2.1.

Las afinidades de unión de treinta y cuatro proteínas de unión  $V_L$  seleccionadas variaron, presentando todas una  $K_D$  en el rango nanomolar (1,5 a 130 nM). Además, aproximadamente el 70% de las proteínas de unión  $V_L$  seleccionadas (23 de 34), demostró una afinidad nanomolar de un único dígito. Las mediciones de  $T^{1/2}$  para estas proteínas de unión  $V_L$  seleccionadas, demostraron un rango de aproximadamente 0,2 a 66 minutos. De las treinta y cuatro proteínas de unión  $V_L$ , seis mostraron una afinidad mayor de 3 nM para el antígeno X (1,53, 2,23, 2,58, 2,59, 2,79 y 2,84). Los datos de afinidad son consistentes con las proteínas de unión  $V_L$  resultantes de la asociación combinatoria de dominios variables de cadena ligera humana reordenados unidos a regiones constantes de cadena pesada y ligera (descritos en la Tabla 4), que son de alta afinidad, seleccionadas por clonación, y somáticamente mutadas. Las proteínas de unión  $V_L$  generadas por los ratones descritos en el presente documento comprenden una colección de diversas proteínas de unión únicas y de alta afinidad, que muestran especificidad para uno o más epítopos del antígeno X.

En otro experimento, las proteínas de unión  $V_L$  humanas seleccionadas producidas contra el antígeno X, se ensayaron para determinar su capacidad de bloquear la unión del ligando natural del antígeno X (ligando Y) a antígeno X en un ensayo basado en perlas LUMINEX® (datos no mostrados). Los resultados demostraron que, además de unirse específicamente al dominio extracelular del antígeno X con afinidades en el rango nanomolar (descrito anteriormente), las proteínas de unión  $V_L$  seleccionadas también fueron capaces de unirse al antígeno X de

mono cinomolgus (*Macaca fascicularis*).

**LISTADO DE SECUENCIAS**

	<110> Macdonald, Lynn Gurer, Cagan 5 Hosiawa, Karolina A. Murphy, Andrew J.	
	<120> R ratones que producen proteínas de unión V <sub>L</sub>	
10	<130> 1200A-WO	
	<140> Pendiente de asignación <141>	
15	<150> 61/369.909 <151> 02-08-2010	
	<160> 29	
20	<170> FastSEQ para Windows Versión 4.0	
	<210> 1 <211> 20	
25	<212> ADN <213> Secuencia artificial	
	<220> <223> sintética	
30	<400> 1 gagaggaccg tggacaagtc	20
	<210> 2 <211> 20 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintética	
40	<400> 2 tgacggtggt gctgtagaag	20
	<210> 3 <211> 28 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintética	
50	<400> 3 atgctgagga ggaaggctt gagaacct	28
55	<210> 4 <211> 20 <212> ADN <213> Secuencia artificial	



	<220>	
	<223> sintética	
5	<400> 4 gctcgtgagc aactgaacct	20
	<210> 5	
	<211> 19	
10	<212> ADN <213> Secuencia artificial	
	<220>	
	<223> sintética	
15	<400> 5 gccactgcac actgatgc	19
	<210> 6	
20	<211> 23 <212> ADN <213> Secuencia artificial	
	<220>	
25	<223> sintética	
	<400> 6 agtcagccac agtcacctgc ctg	23
30	<210> 7 <211> 19 <212> ADN <213> Secuencia artificial	
35	<220> <223> sintética	
	<400> 7 gcctgcacaa ccaccatac	19
40	<210> 8 <211> 22 <212> ADN <213> Secuencia artificial	
45	<220> <223> sintética	
	<400> 8 gagcaggaag aggctgatga ag	22
50	<210> 9 <211> 25 <212> ADN <213> Secuencia artificial	
55	<220> <223> sintética	

	<400> 9 agaagagcct ctcccactct cctgg	25
	<210> 10	
5	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
10	<223> sintética	
	<400> 10 cagccagcgg agaactacaa g	21
15	<210> 11	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> sintética	
	<400> 11 gcctcccagt tgcttctctg	20
25	<210> 12	
	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> sintética	
	<400> 12 aacactcagc ccatcatgga caca	24
35	<210> 13	
	<211> 19	
	<212> ADN	
40	<213> Secuencia artificial	
	<220>	
	<223> sintética	
45	<400> 13 tgagcagcac cctcacgtt	19
	<210> 14	
	<211> 23	
50	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> sintética	
55	<400> 14 gtggcctcac aggtatagct gtt	23
	<210> 15	

# ES 2 612 459 T3

	<211> 18 <212> ADN <213> Secuencia artificial	
5	<220> <223> sintética	
	<400> 15 accaaggacg agtatgaa	18
10	<210> 16 <211> 16 <212> ADN <213> Secuencia artificial	
15	<220> <223> sintética	
	<400> 16 20 gtaaaacgac ggccag	16
	<210> 17 <211> 17 <212> ADN 25 <213> Secuencia artificial	
	<220> <223> sintética	
30	<400> 17 caggaaacag ctatgac	17
	<210> 18 <211> 145 35 <212> ADN <213> Secuencia artificial	
	<220> <223> sintética	
40	<400> 18 tctgggacag aattcactct caccatcagc agcctgcagc ctgatgatt tgcaacttat tactgccaac agtataatac cctcacttcc ggcggagggga ccaaggtgga gatcaaaccc aaaacaacag ccccatcggg ctatc	60 120 145
45	<210> 19 <211> 151 <212> ADN <213> Secuencia artificial	
50	<220> <223> sintética	
	<400> 19 55 tctgggacag aatccactct cacaatcagc agcctgcagc ctgaagatt tgcaacttat tactgtcaac agcttaatag ttacccttcc acttccggcg gagggaccaa ggtggagatc aaacccaaaa caacagcccc atcgggtctat c	60 120 151
	<210> 20	

	<211> 151	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> sintética	
	<400> 20	
	tctgggacag atttactct caccatcagc agcctgcagc ctgaagattt tgcaacttat	60
10	tactgccaac agtataatag ttaccctccc accttcggcc aagggacacg actggagatt	120
	aaacctacaa caacagcccc atctgtctat c	151
	<210> 21	
	<211> 153	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> sintética	
20	<400> 21	
	tctgggacag atttactct caccatcagc agcctgcagc ctgaagattt tgcaacttac	60
	tattgtcaac aggctaacag ttcccgtac acttttggcc aggggaccaa gctggagatc	120
	aaacccaaaa cgacaccccc atctgtctat cca	153
25	<210> 22	
	<211> 151	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> sintética	
	<400> 22	
35	tctgggacag atttactct caccatcagc agcctgcagc ctgaagatgt tgcaacttat	60
	tactgtcaaa agtataacag tgcccctcac acttctggcg gagggaccaa ggtggagatc	120
	aaacccaaaa caacagcccc atcgttctat c	151
	<210> 23	
40	<211> 153	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
45	<223> sintética	
	<400> 23	
	tcaggcacag attttactt gaaaatcagc agagtggagg ctgaggatgt tggggtttat	60
	tactgcatgc aagctctaca aatttcgtgg acgittcggcc aagggaccaa ggtggaaatc	120
50	aaacccaaaa cgacaccccc atctgtctat cca	153
	<210> 24	
	<211> 150	
	<212> ADN	
55	<213> Secuencia artificial	
	<220>	
	<223> sintética	

# ES 2 612 459 T3

	<400> 24		
	tctgggacag acttactct caccatcagc agcctagagc ctgaagatt tgcagttat		60
	tactgtcagc agcgtagccc ccgttact ttcggcggag ggaccaaggt ggagatcaaa		120
	cccaaaacga caccacctc tgtctatcca		150
5	<210> 25		
	<211> 148		
	<212> ADN		
	<213> Secuencia artificial		
10	<220>		
	<223> sintética		
	<400> 25		
15	tctgggacag acttactct caccatcagc agactggagc ctgaagatt tgcctgtat		60
	tactgtcagc agtatgtag ctactcact ttcggcggag ggaccaaggt ggagatcaaa		120
	cccaaaacaa cagccccatc ggtctatc		148
	<210> 26		
20	<211> 151		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
25	<223> sintética		
	<400> 26		
	tctgggacag atttactct caccatcagc agcctgcagg ctgaagatgt ggcagttat		60
	tactgtcagc aatattatag tactccgatc accttcggcc aaggacacg actggagatt		120
30	aaacccaaaa caacagcccc atcgtctat c		151
	<210> 27		
	<211> 151		
	<212> ADN		
35	<213> Secuencia artificial		
	<220>		
	<223> sintética		
40	<400> 27		
	tctgggacag atttactct caccatcagc agcctgcagg ctgaagatgt ggcagttat		60
	tactgtcagc aatattatag tactggccc actttggcc agggaccaa gctggagatc		120
	aaacctacaa caacagcccc atcgtctat c		151
45	<210> 28		
	<211> 151		
	<212> ADN		
	<213> Secuencia artificial		
50	<220>		
	<223> sintética		
	<400> 28		
	tatggaacag atttaccct cacaattaat aacattgaat gtgaggatgc tgcattatc		60
55	ttctgtctac aacatgataa ttcccgtac actttggcc agggaccaa gctggagatc		120
	aaacccaaaa caacagcccc atcgtctat c		151
	<210> 29		
	<211> 147		

# ES 2 612 459 T3

<212> ADN

<213> Secuencia artificial

<220>

5 <223> sintética

<400> 29

tatggaacag atttaccct cacaattaat aacatagaat ctgaggatgc tgcatttac

60

ttctgtctac aacatgataa ttggacgttc ggccaagga ccaaggtgga aatcaaacc

120

10 aaaacgacac ccccatctgt ctatcca

147

## REIVINDICACIONES

1. Un ratón que comprende en su línea germinal un primer segmento génico variable de cadena ligera kappa ( $V_{\kappa}$ ) humano no reordenado y un segmento génico de unión de cadena ligera kappa ( $J_{\kappa}$ ) humano no reordenado unidos operativamente a la región constante de cadena pesada ( $C_H$ ) de ratón endógena en el locus de cadena pesada de ratón endógeno, donde el primer segmento génico  $V_{\kappa}$  humano no reordenado y el segmento génico  $J_{\kappa}$  humano no reordenado reemplazan todos los segmentos génicos variables de cadena pesada ( $V_H$ ) de ratón endógenos funcionales, todos los segmentos génicos de diversidad ( $D_H$ ) de ratón endógenos funcionales y todos los segmentos génicos de unión de cadena pesada ( $J_H$ ) de ratón endógenos funcionales, donde el primer segmento génico  $V_{\kappa}$  humano no reordenado y el segmento génico  $J_{\kappa}$  humano no reordenado participan en el reordenamiento para formar una secuencia  $V_{\kappa}/J_{\kappa}$  reordenada unida operativamente a la región constante de cadena pesada de ratón endógena en el ratón, y donde el ratón comprende adicionalmente en su línea germinal un segundo segmento génico variable de cadena ligera ( $V_L$ ) humano y un segmento génico J de cadena ligera ( $J_L$ ) humano unidos operativamente a un gen constante de cadena ligera ( $C_L$ ) de ratón.
2. El ratón de la reivindicación 1, que comprende una célula B que comprende en su genoma una secuencia de ácido nucleico de región variable de cadena ligera kappa de inmunoglobulina humana reordenada unida operativamente a la región constante de cadena pesada de ratón endógena en un locus de cadena pesada de inmunoglobulina de ratón endógeno.
3. El ratón de la reivindicación 1, donde el segundo segmento génico  $V_L$  humano es un segmento génico  $V_{\kappa}$  humano.
4. El ratón de la reivindicación 3, donde el gen constante de cadena ligera es un gen constante de cadena ligera  $\kappa$ .
5. El ratón de la reivindicación 1, donde el segundo segmento génico  $V_L$  humano es un segmento génico  $V_{\lambda}$  humano.
6. El ratón de la reivindicación 5, donde el gen constante de cadena ligera es un gen constante de cadena ligera  $\lambda$ .
7. El ratón de la reivindicación 1 que expresa una proteína de unión a antígeno, donde la proteína de unión a antígeno comprende un primer polipéptido que comprende un primer dominio variable de cadena ligera kappa humano fusionado con una región constante de cadena pesada de inmunoglobulina de ratón, y un segundo polipéptido que comprende un segundo dominio variable de cadena ligera humano fusionado con una región constante de cadena ligera de inmunoglobulina de ratón.
8. El ratón de la reivindicación 7, donde el segundo dominio variable de cadena ligera humano se selecciona de un dominio variable  $V_{\kappa}$  humano y un dominio variable  $V_{\lambda}$  humano.
9. El ratón de la reivindicación 1, que es homocigoto o heterocigoto para dicho locus de cadena pesada endógeno, que comprende un primer segmento génico  $V_{\kappa}$  humano no reordenado y un segmento génico  $J_{\kappa}$  humano no reordenado unidos operativamente con la región constante de cadena pesada de ratón endógena en el locus de cadena pesada de ratón endógeno, donde el primer segmento génico  $V_{\kappa}$  humano no reordenado y el segmento génico  $J_{\kappa}$  humano no reordenado reemplazan todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  de ratón endógenos funcionales.
10. Uso de un ratón de acuerdo con cualquiera de las reivindicaciones anteriores para producir una proteína de unión a antígeno que comprende un primer polipéptido que comprende un primer dominio variable de cadena ligera kappa humano fusionado con una región constante de cadena pesada de inmunoglobulina de ratón, y un segundo polipéptido que comprende un segundo dominio variable de cadena ligera humano fusionado con una región constante de cadena ligera de inmunoglobulina de ratón.
11. Uso de un ratón de acuerdo con cualquiera de las reivindicaciones 1-9 para producir un anticuerpo que comprende un primer polipéptido que comprende un primer dominio variable de cadena ligera kappa humano fusionado con una región constante de cadena pesada de inmunoglobulina de ratón, y un segundo polipéptido que comprende un segundo dominio variable de cadena ligera humano fusionado con una región constante de cadena ligera de inmunoglobulina de ratón.

12. Uso de un ratón de acuerdo con una cualquiera de las reivindicaciones 1-9 para producir un hibridoma o un cuadroma para la producción de un anticuerpo como se define en la reivindicación 11.
13. Uso de acuerdo con la reivindicación 11, donde el anticuerpo es un anticuerpo biespecífico.
- 5 14. Una célula de ratón que comprende en su genoma un primer segmento génico  $V_{\kappa}$  humano no reordenado y un segmento génico  $J_{\kappa}$  humano no reordenado unidos operativamente con la región constante de cadena pesada endógena en el locus de cadena pesada de ratón endógeno, donde el primer segmento génico  $V_{\kappa}$  humano no reordenado y el segmento génico  $J_{\kappa}$  humano no reordenado reemplazan todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  endógenos funcionales, y donde la célula de ratón comprende adicionalmente un segundo segmento génico  $V_L$  humano y un segmento génico  $J_L$  humano unidos operativamente a un gen constante de cadena ligera de ratón.
- 10 15. Una célula de acuerdo con la reivindicación 14 que se obtiene a partir de un ratón de acuerdo con una cualquiera de las reivindicaciones 1 a 9.
- 15 16. Una célula de acuerdo con la reivindicación 14 ó 15, donde la célula es una célula ME.
17. Un tejido obtenido de un ratón de acuerdo con una cualquiera de las reivindicaciones 1-9, que comprende células como se define en la reivindicación 14.
- 20 18. El tejido de la reivindicación 17, que comprende una célula B que comprende una secuencia de ácido nucleico de región variable de cadena ligera kappa de inmunoglobulina humana reordenada unida operativamente a la región constante de cadena pesada de ratón en el locus de cadena pesada de ratón endógeno y una secuencia de ácido nucleico de región variable de cadena ligera de inmunoglobulina humana reordenada unida operativamente a un gen de la región constante de cadena ligera de ratón.
- 25 19. Un embrión de ratón que comprende una célula que comprende en su genoma un primer segmento génico  $V_{\kappa}$  humano no reordenado y un segmento génico  $J_{\kappa}$  humano no reordenado unidos operativamente con la región constante de cadena pesada endógena en el locus de cadena pesada de ratón endógeno, donde el primer segmento génico  $V_{\kappa}$  humano no reordenado y el segmento génico  $J_{\kappa}$  humano no reordenado reemplazan todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  endógenos funcionales, y donde la célula comprende adicionalmente un segundo segmento génico  $V_L$  humano y un segmento génico  $J_L$  humano unidos operativamente a un gen constante de cadena ligera de ratón.
- 30 20. Un método para preparar un ratón genéticamente modificado de acuerdo con una cualquiera de las reivindicaciones 1 a 9, que comprende reemplazar en un locus de cadena pesada endógeno del ratón todos los segmentos génicos  $V_H$ ,  $D_H$  y  $J_H$  funcionales del ratón con un primer segmento génico  $V_{\kappa}$  humano no reordenado y un segmento génico  $J_{\kappa}$  humano no reordenado, para unir operativamente así el primer segmento génico  $V_{\kappa}$  humano no reordenado y el segmento génico  $J_{\kappa}$  humano no reordenado a la región constante de cadena pesada endógena, y también insertar en la línea germinal del ratón un segundo segmento génico  $V_L$  humano y un segmento génico  $J_L$  humano unidos operativamente a un gen de región constante de cadena ligera de ratón.
- 35 40 21. Una proteína de unión a antígeno obtenible a partir de un ratón de acuerdo con una cualquiera de las reivindicaciones 1 a 9, donde la proteína de unión a antígeno comprende un dominio variable de cadena ligera de inmunoglobulina humano fusionado a un dominio constante de cadena ligera de ratón y un dominio variable de cadena ligera kappa de inmunoglobulina humano fusionado a un dominio constante de cadena pesada de ratón.
- 45 22. Un método para preparar una proteína de unión a antígeno, donde dicho método comprende obtener una secuencia nucleotídica que codifica un dominio  $V_K$  de un gen que codifica un dominio  $V_K$  fusionado a una región  $C_H$  de una célula de un ratón de acuerdo con una cualquiera de las reivindicaciones 1 a 9, clonar la secuencia nucleotídica que codifica el dominio  $V_K$  en marco con un gen que codifica una región  $C_H$  humana para formar una secuencia de proteína de unión a antígeno humana, y expresar la secuencia de proteína de unión a antígeno humana en una célula adecuada.
- 50 23. Un método para preparar una proteína de unión a antígeno que comprende un dominio  $V_K$  humano, donde dicho método comprende exponer un ratón de acuerdo con una cualquiera de las reivindicaciones 1 a 9 a un antígeno de interés, permitir que el ratón desarrolle una respuesta inmune al antígeno de interés, y aislar dicha proteína de unión a antígeno, o aislar el dominio  $V_K$  humano de dicha proteína de unión a antígeno.
- 55



24. Un método para obtener una secuencia génica  $V_K$  humana, donde dicho método comprende exponer un ratón de acuerdo con una cualquiera de las reivindicaciones 1 a 9 a un antígeno de interés, y aislar de dicho ratón una secuencia génica  $V_K$  humana reordenada, donde la secuencia génica  $V_K$  humana reordenada se fusiona con una  
5 secuencia nucleotídica que codifica una región  $C_H$  en dicho ratón.

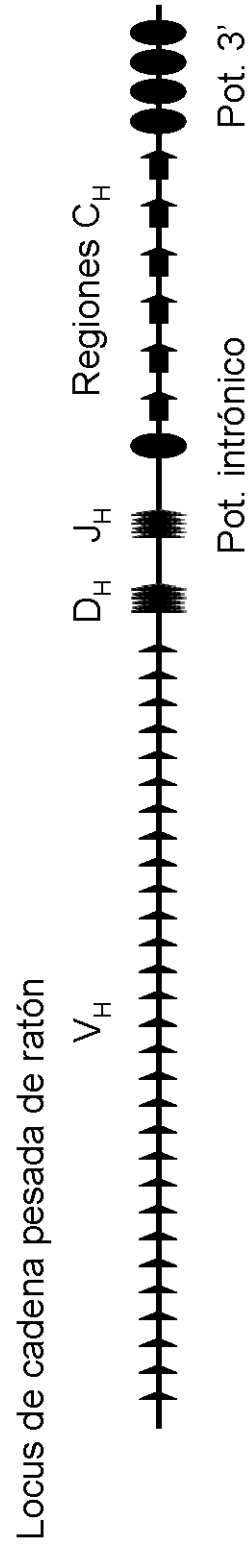


FIG. 1A

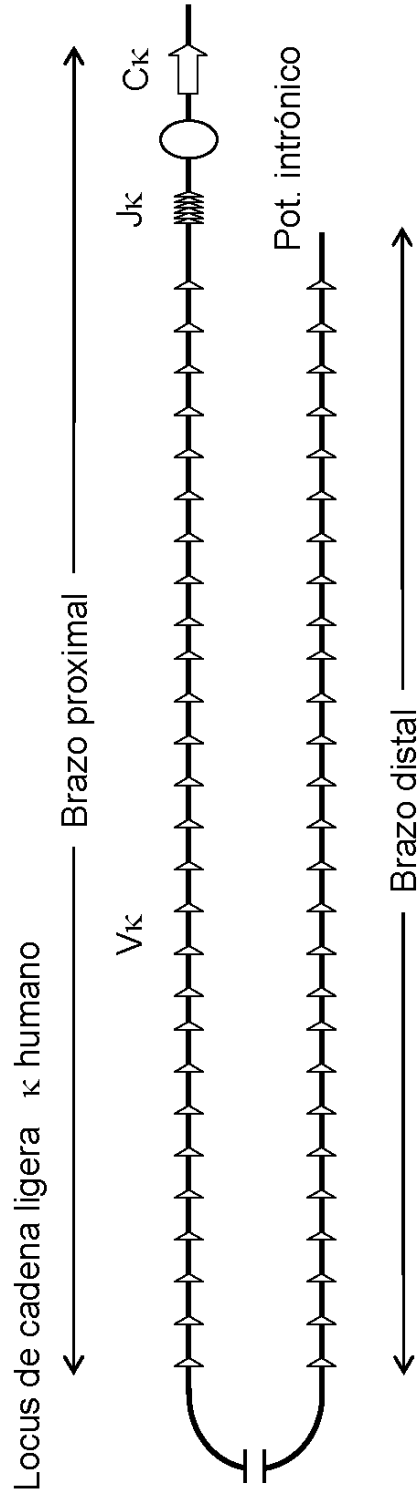


FIG. 1B

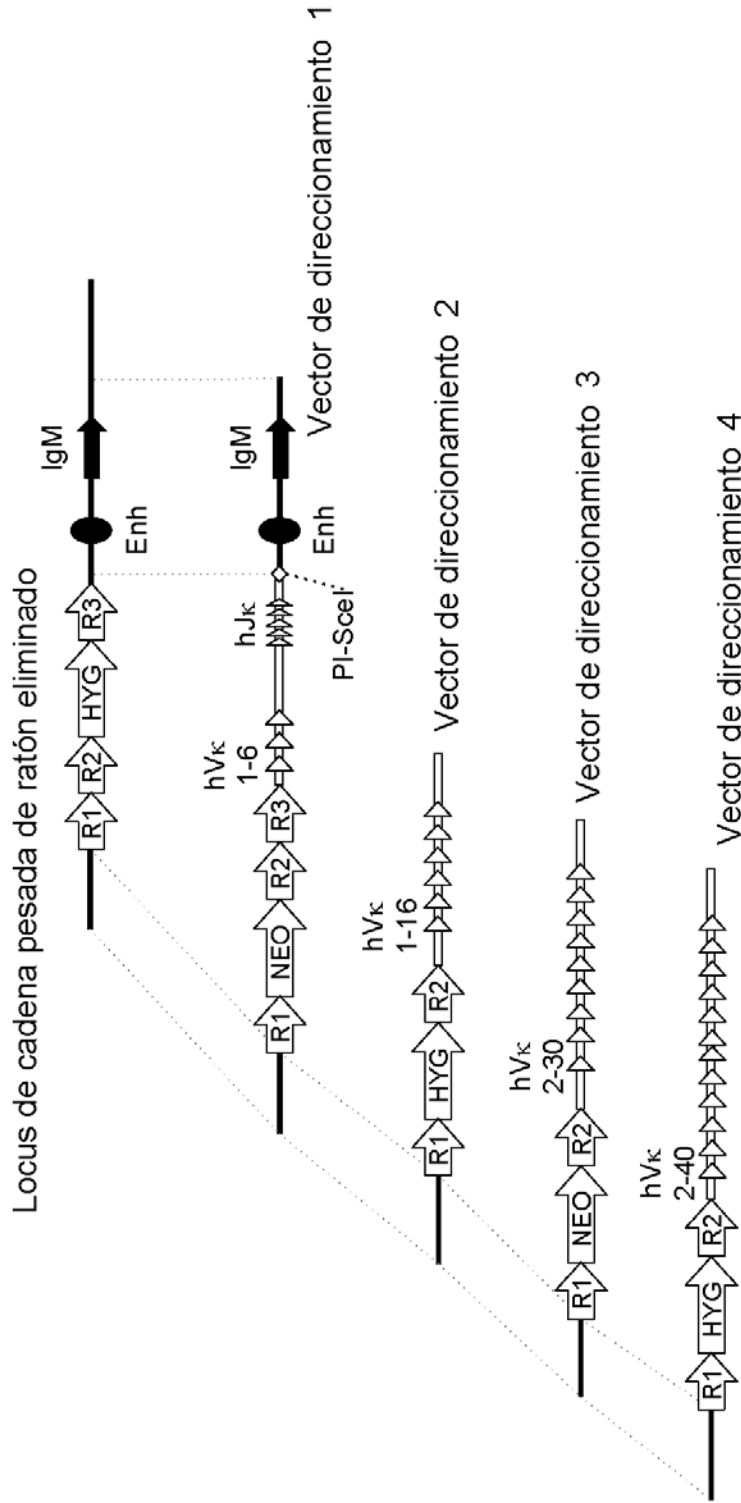


FIG. 2

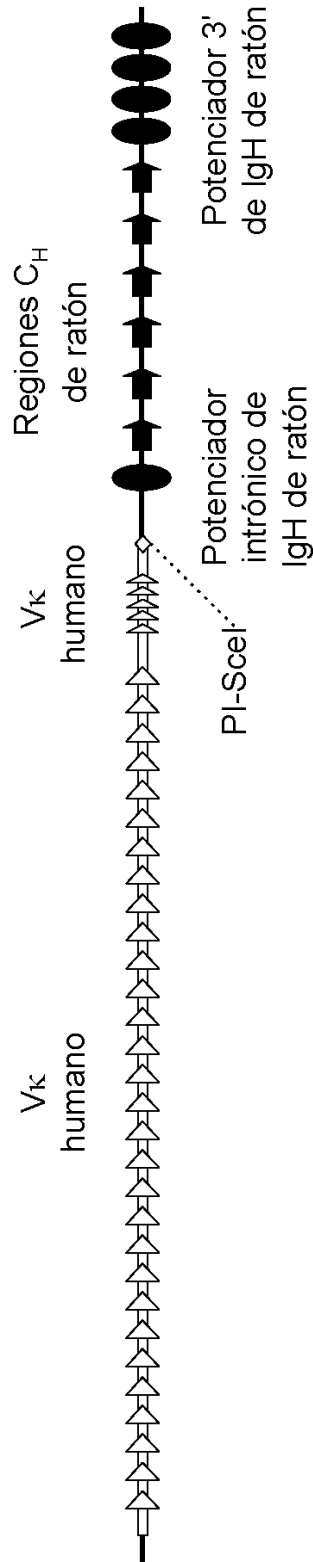


FIG. 3

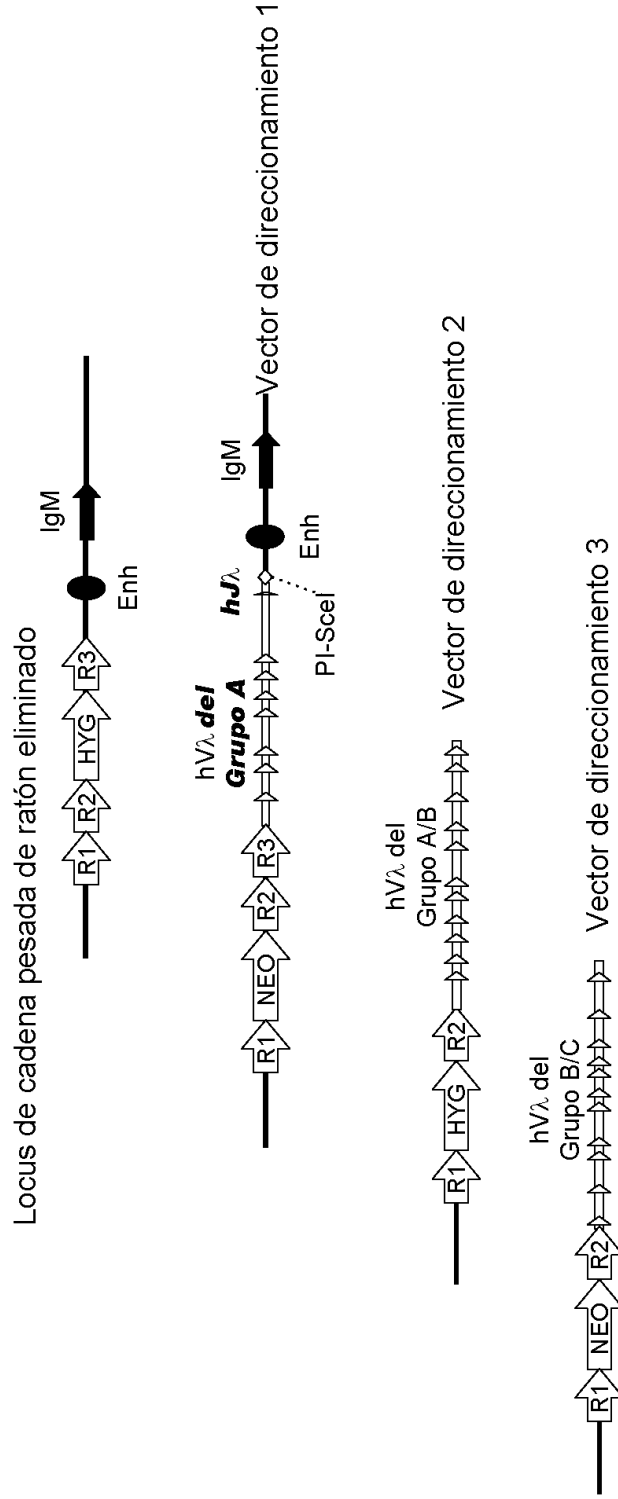


FIG. 4A

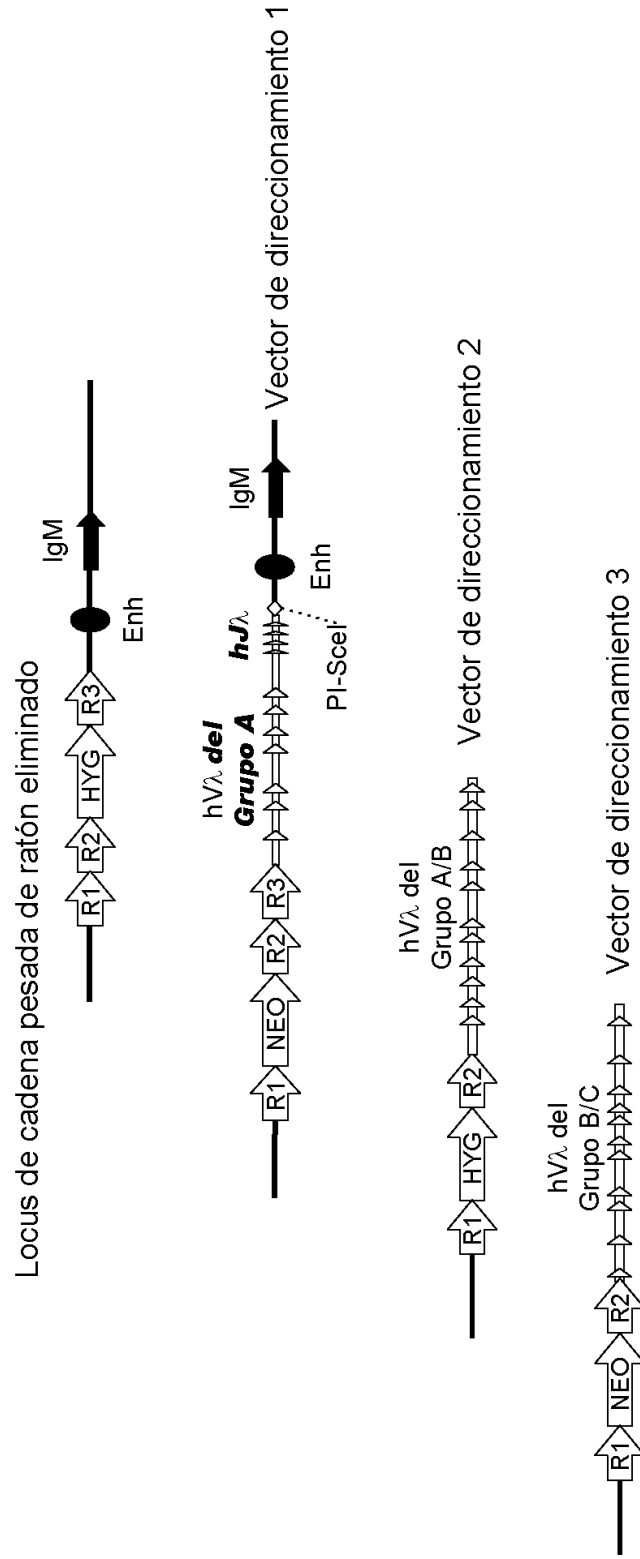


FIG. 4B

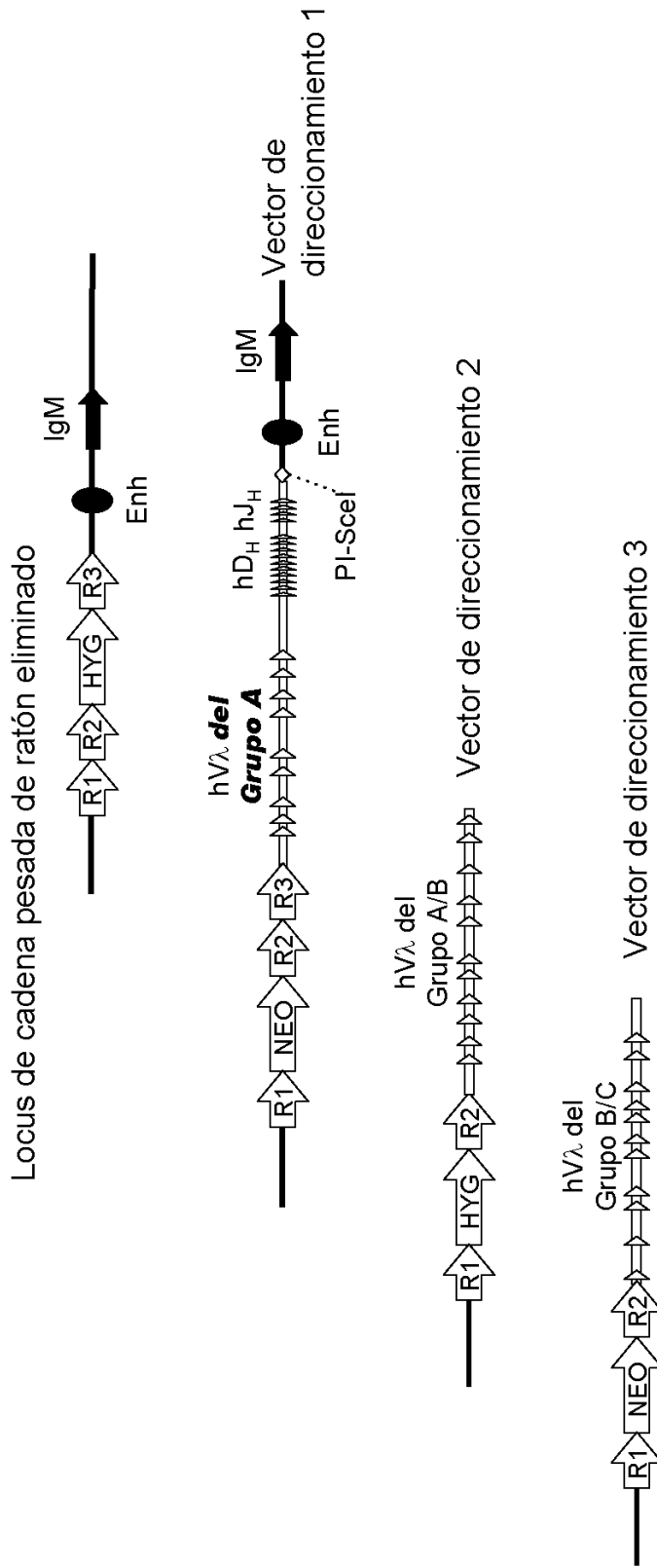


FIG. 5A



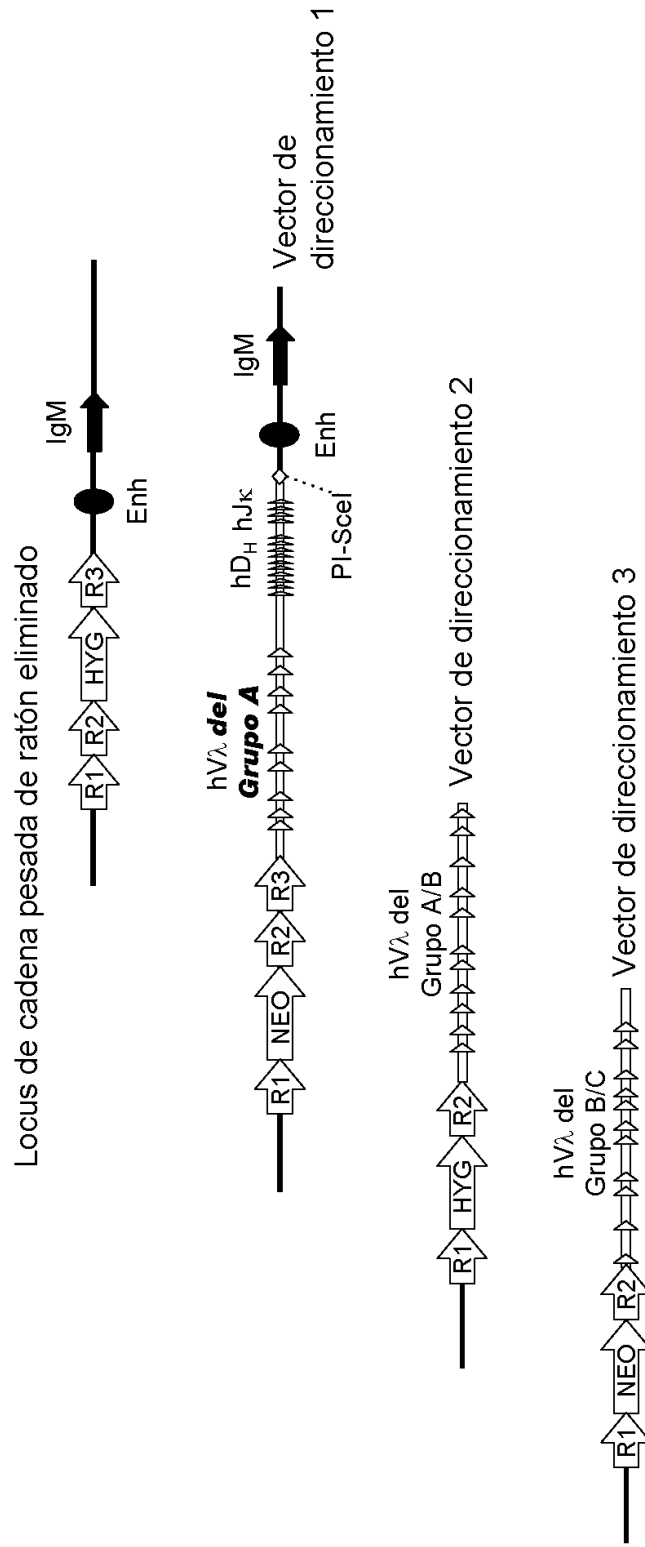


FIG. 5B

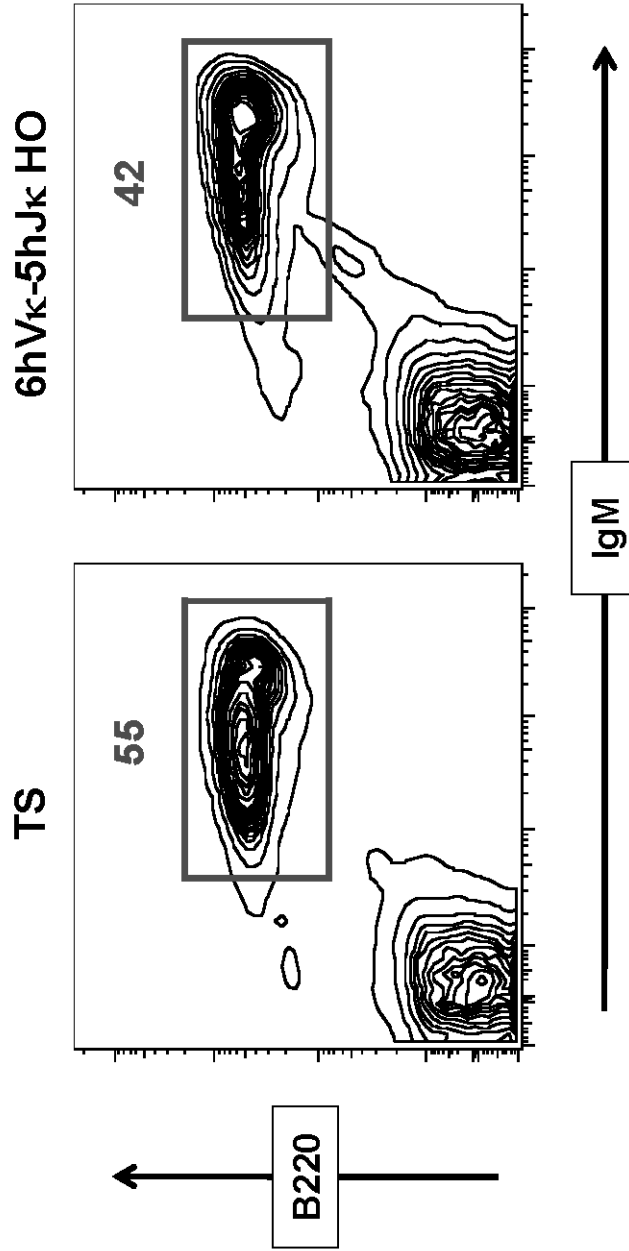


FIG. 6A

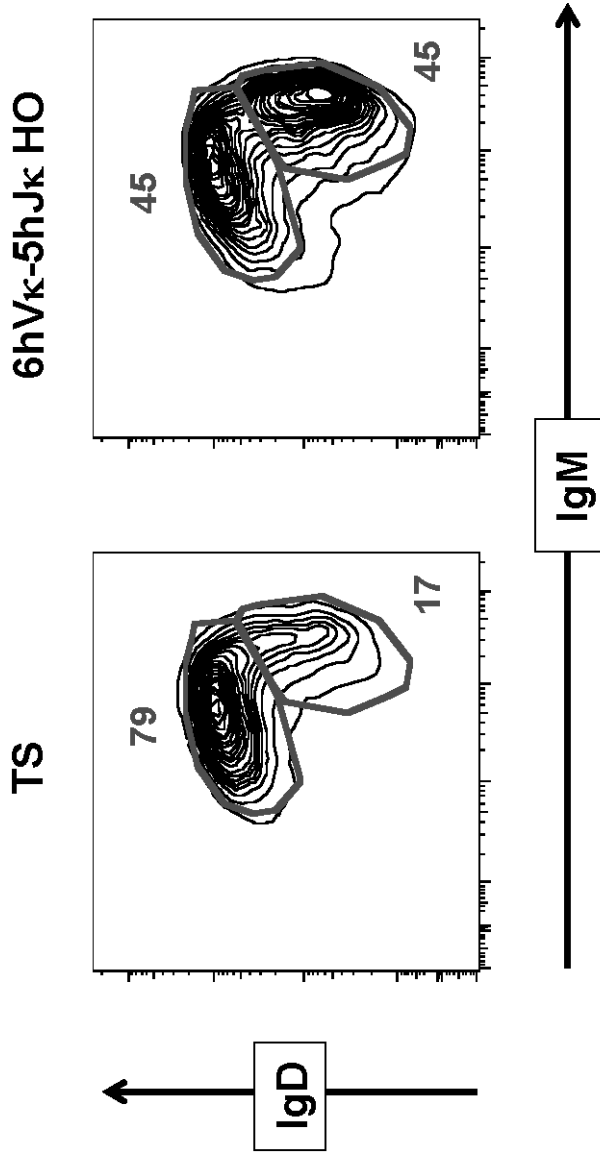


FIG. 6B

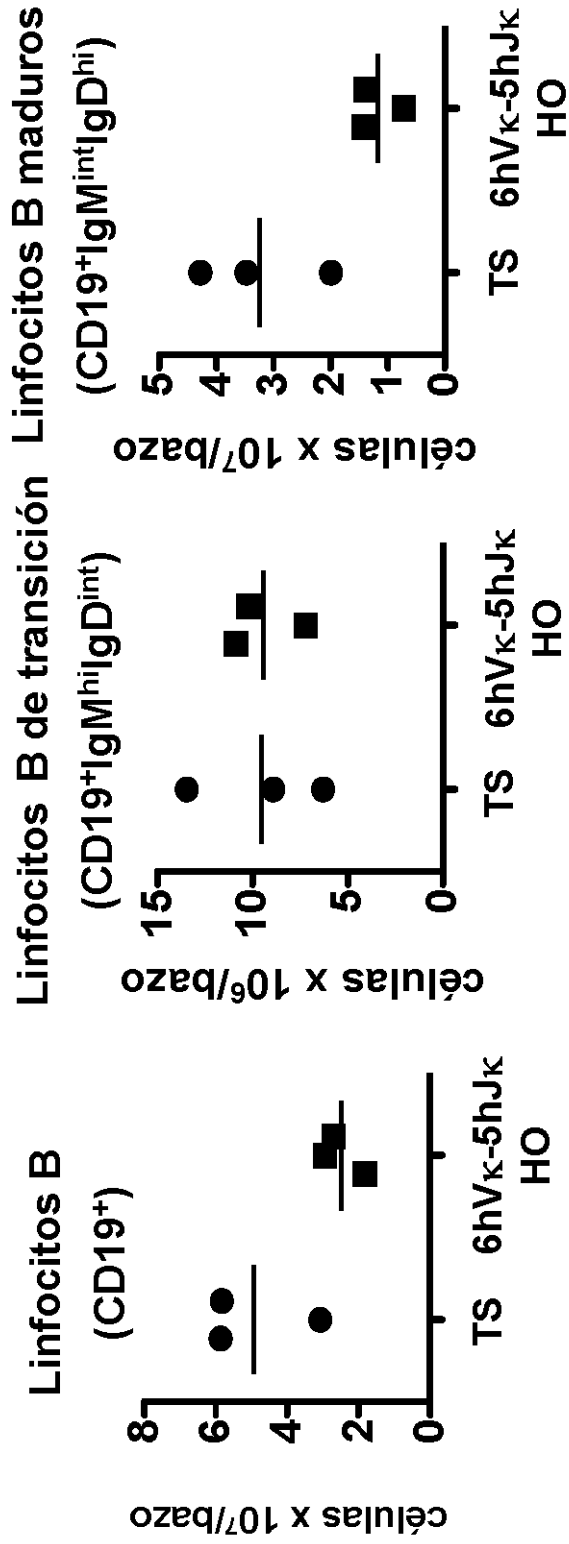


FIG. 6C

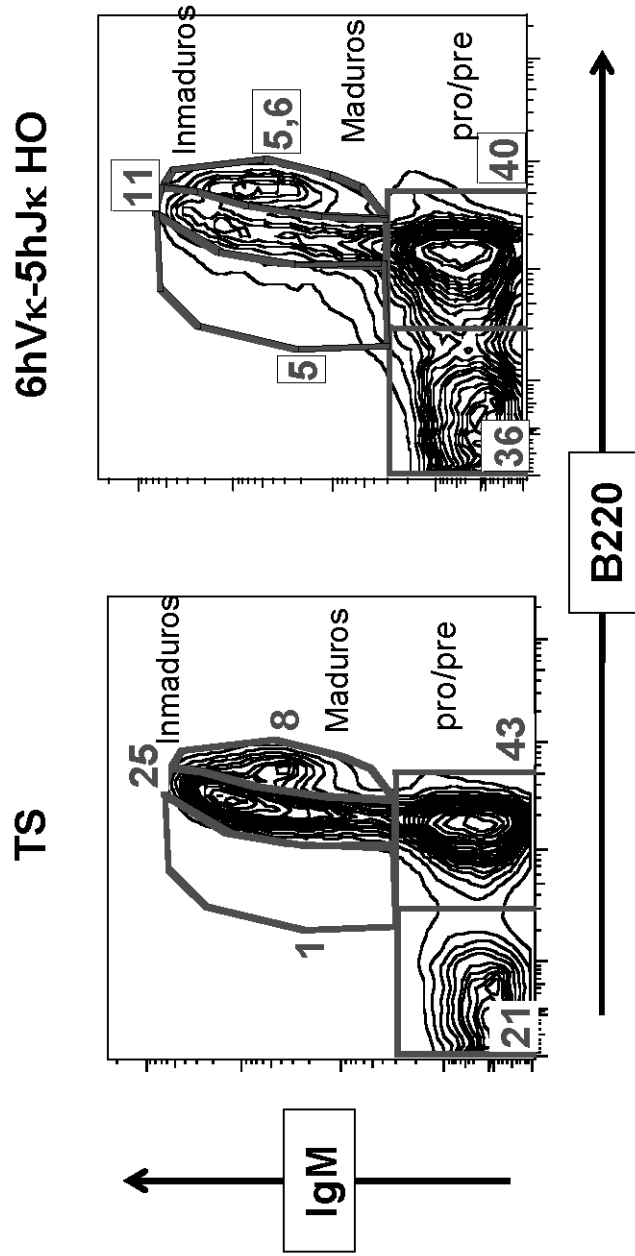


FIG. 7A

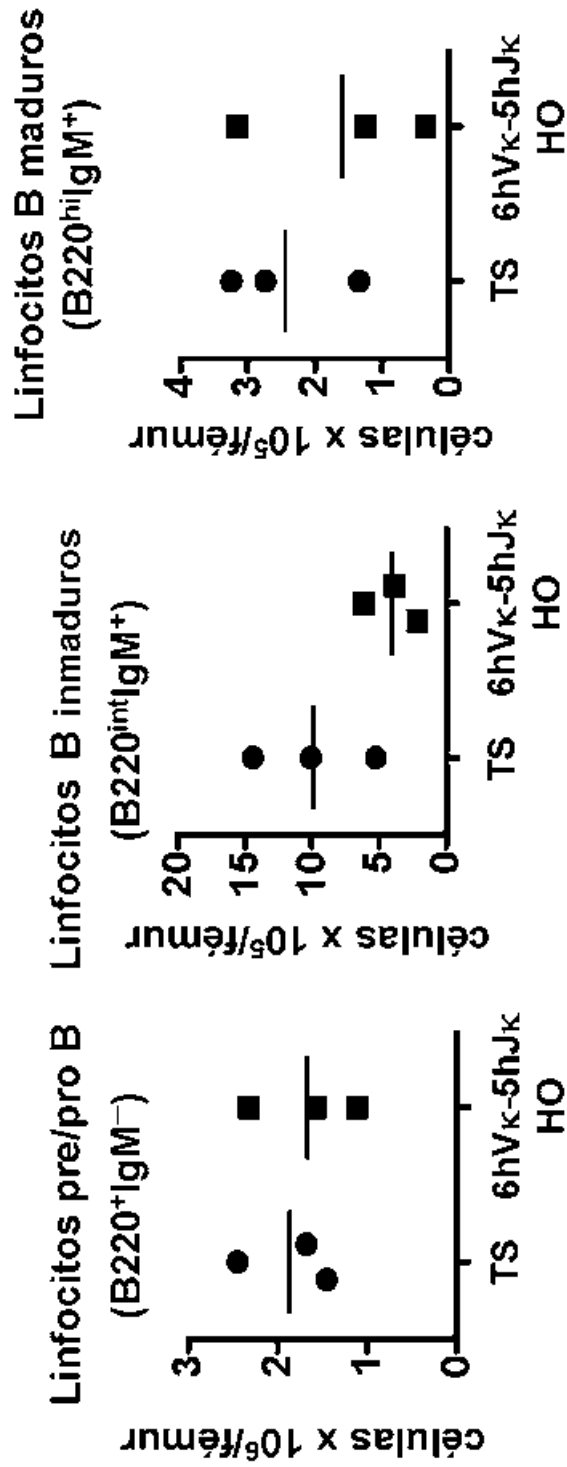


FIG. 7B

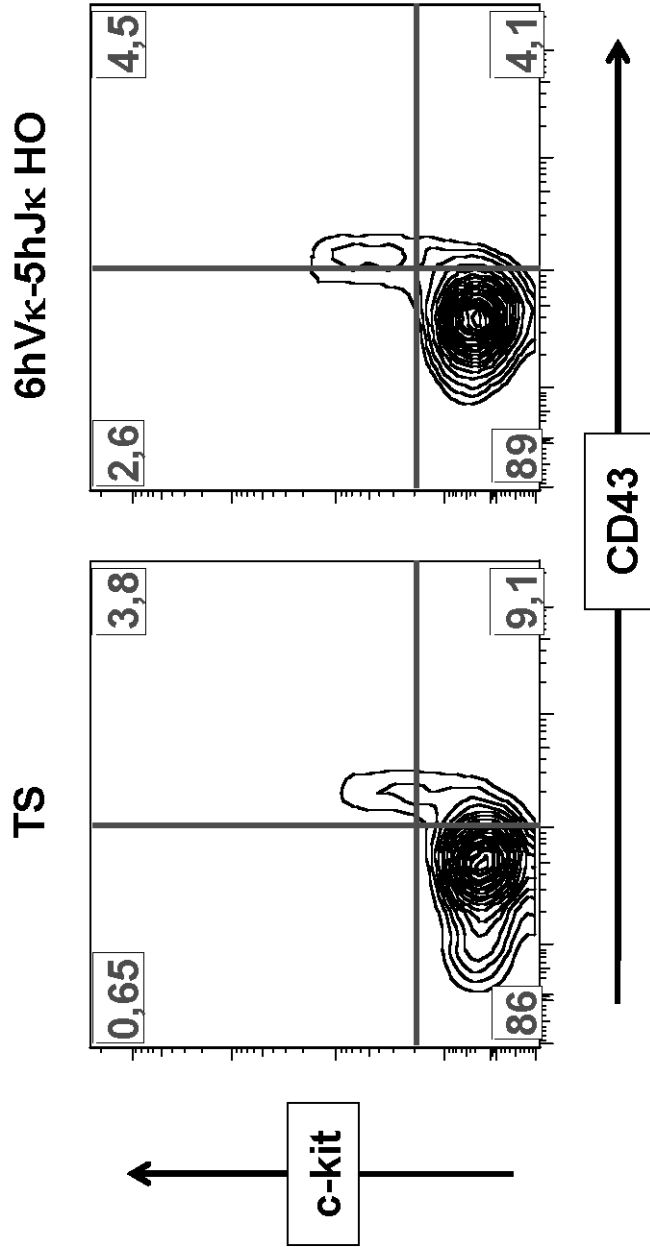


FIG. 7C

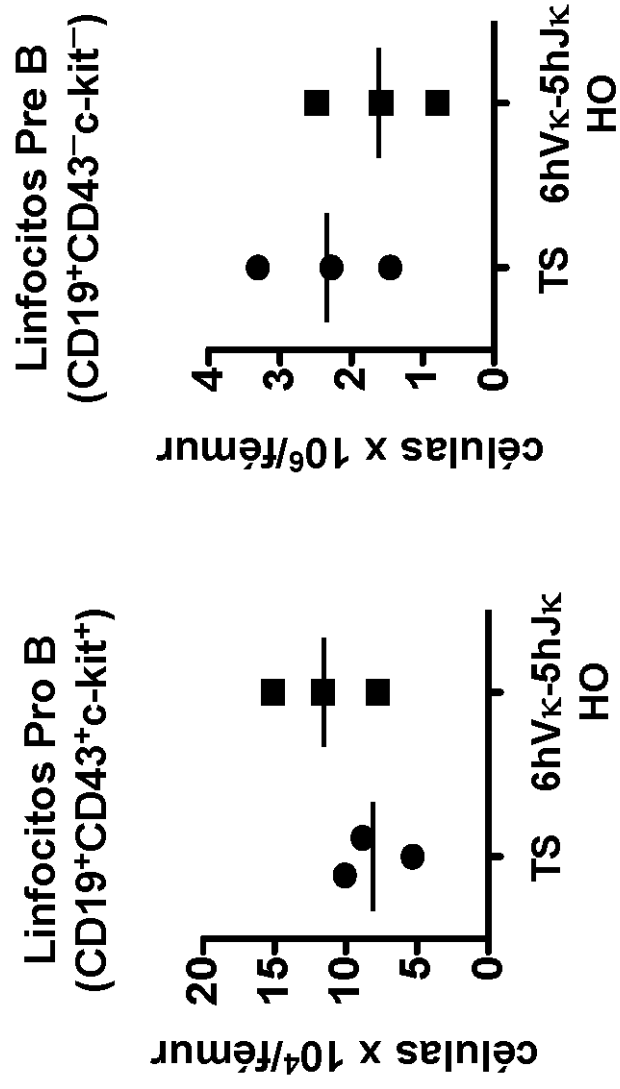


FIG. 7D



6hVκ-5hJκ HO

TS

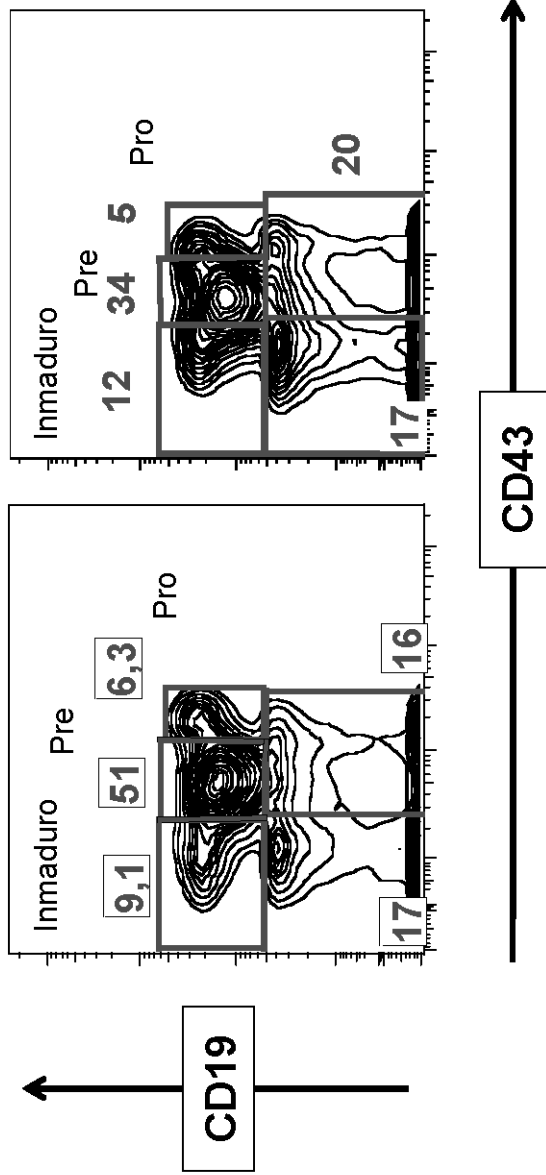


FIG. 7E

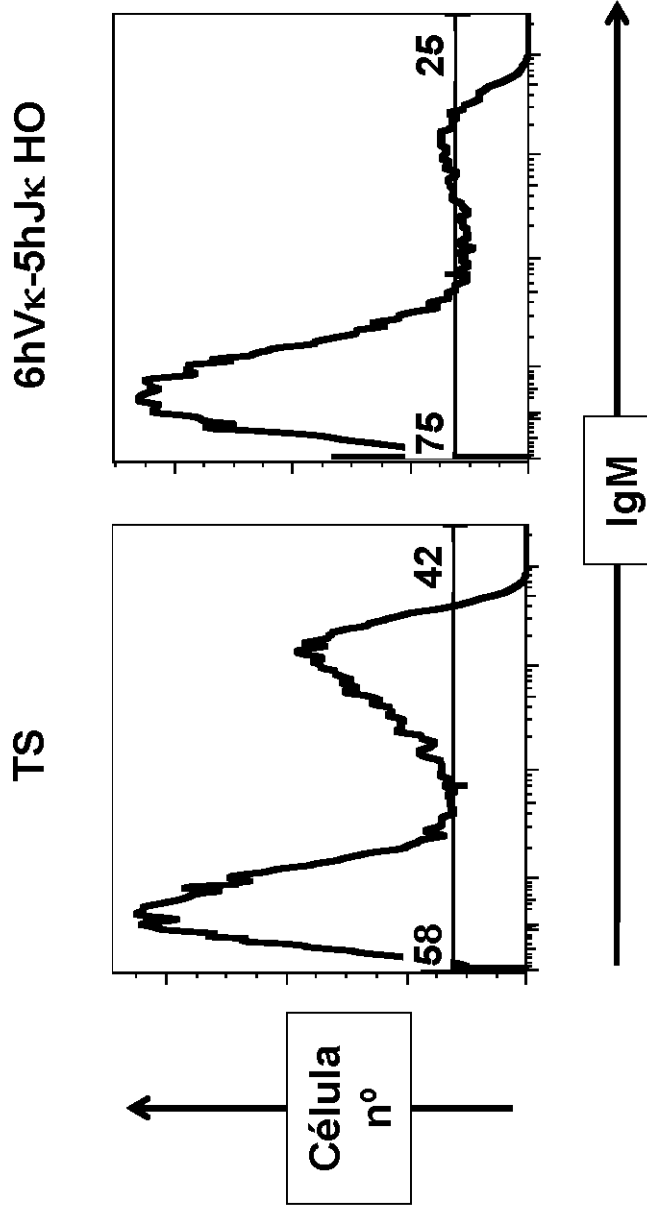


FIG. 7F

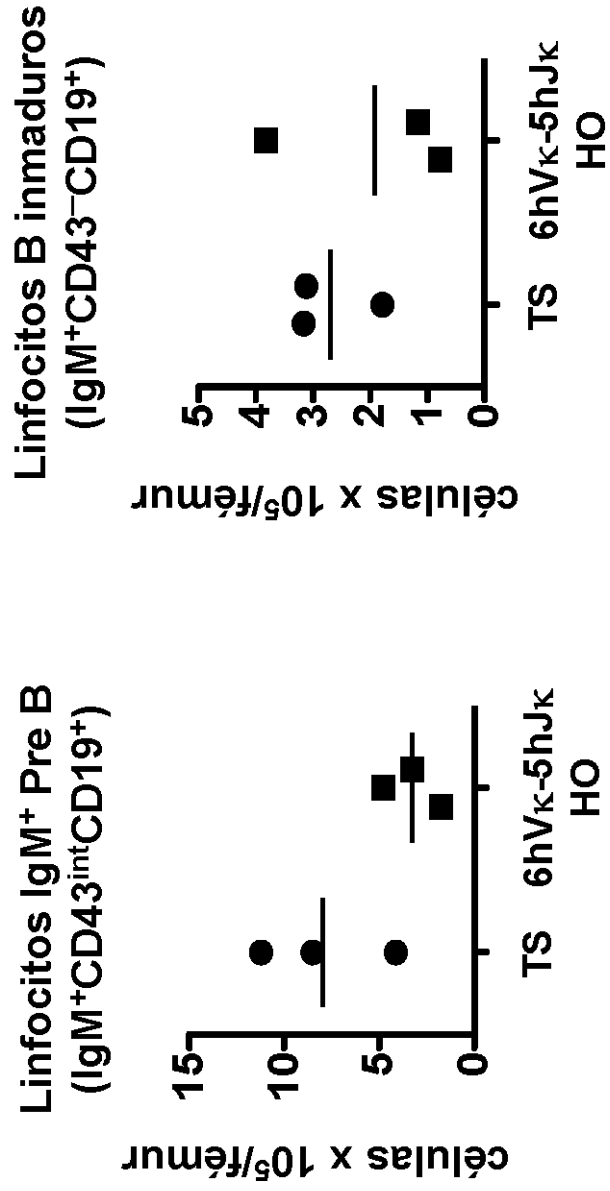


FIG. 7G

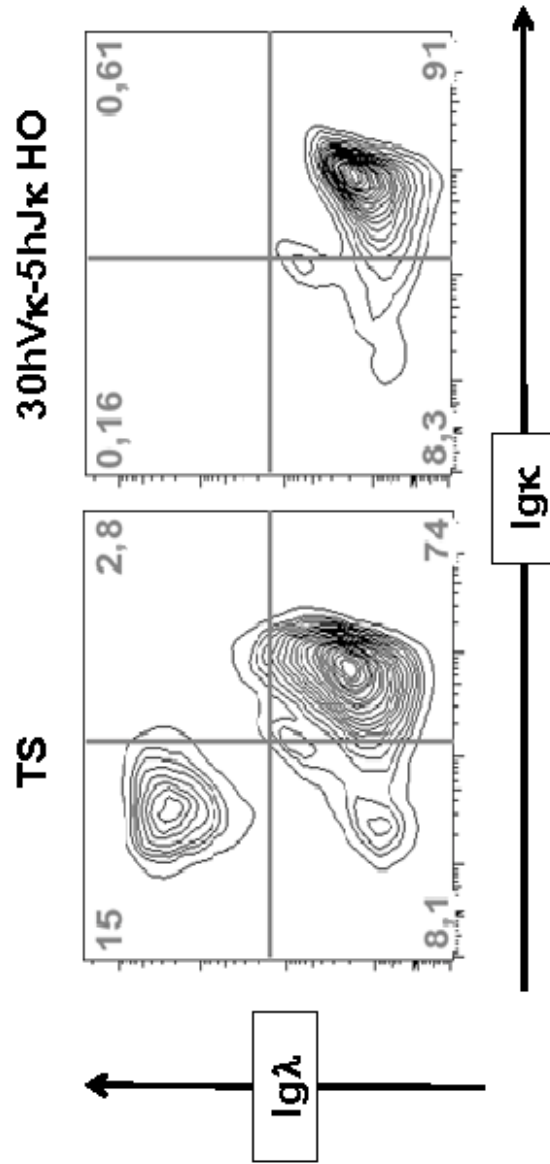


FIG. 8A

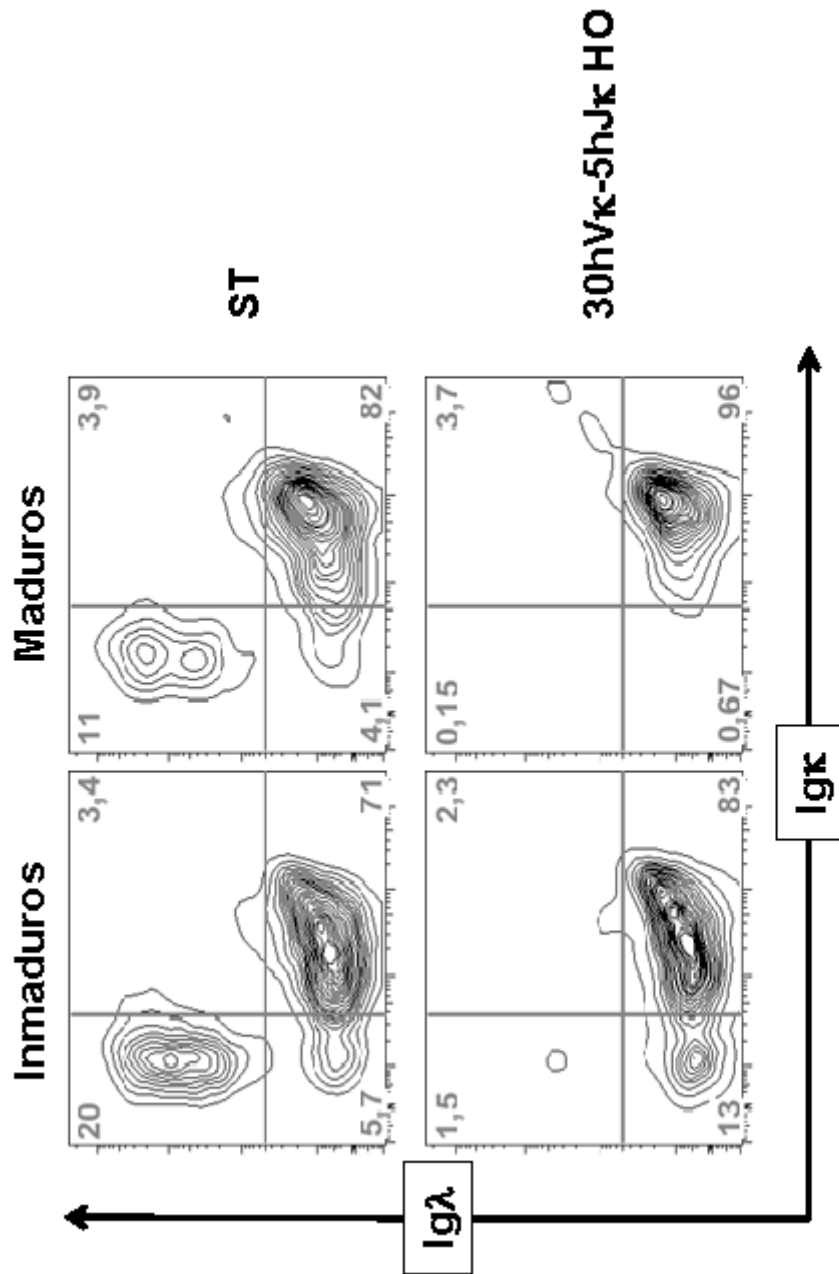


FIG. 8B

<b>Vκ-Jκ humana reordenada</b>	<b>IgG1 de ratón</b>
2E TGTC AACAGGCTAACAGTTTCCCGTACACTTTTGGCCAGGGACCAAGCTGGAGATCAAAC.CCAAAAACGACACACCCCATCTGTCTATCCA 2H TGCAATGCAAGCTCTACAAAATTCGACGTTCCGCCAAGGGACCAAGGTGGAATCAAAC.CCAAAAACGACACACCCCATCTGTCTATCCA 3D TGT CAGCAGCGTAGCccccgtt... TCACTTTCGGCGGAGGACCAAGGTGGAGATCAAAC.CCAAAAACGACACACCCCATCTGTCTATCCA 5D TGTCACAACATGATAATT.....GGACGTTCCGCCAAGGGACCAAGGTGGAATCAAAC.CCAAAAACGACACACCCCATCTGTCTATCCA	
<b>Vκ-Jκ humana reordenada</b>	<b>IgG2A/2C de ratón</b>
1E TGCC AACAGTATAATAcc.....TCACTTTCGGCGGAGGGACCAAGGTGGAGATCAAAC.CCAAAAACACACGCCCCATCGGTCTATC 1G TGTCAACAGCTTAATAGTTACCCCTTCACTTTCGGCGGAGGGACCAAGGTGGAGATCAAAC.CCAAAAACACAGCCCCATCGGTCTATC 1C TGTCAAAAAGTATAACAGTGGCCCCCaCACTTTCGGCGGAGGGACCAAGGTGGAGATCAAAC.CCAAAAACACAGCCCCATCGGTCTATC 3A TGT CAGCAGTATGGTAGCTCA...CTCACTTTCGGCGGAGGGACCAAGGTGGAGATCAAAC.CCAAAAACACAGCCCCATCGGTCTATC 4B TGT CAGCAATATTATAGTACTCCGATCACCTTCGGCCAAAGGGACACGACTGGAGATTAAC.CCAAAAACACAGCCCCATCGGTCTATC 5A TGTCTACAACATGATAATTTCCCGTACACTTTTGGCCAGGGACCAAGCTGGAGATCAAAC.CCAAAAACACAGCCCCATCGGTCTATC	
<b>Vκ-Jκ humana reordenada</b>	<b>IgG3 de ratón</b>
1A TGCC AACAGTATAATAGTTACCCCTCCACCTTCGGCCAAAGGGACACGACTGGAGATTAAC.CTACAACAACAGCCCCATCTGTCTATC 4C TGT CAGCAATATTATAGTACTggccccCACTTTGGCCAGGGACCAAGCTGGAGATCAAAC.CTACAACAACAGCCCCATCTGTCTATC	

FIG. 9