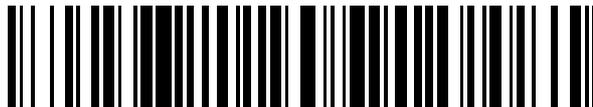


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 551 871**

51 Int. Cl.:

C07K 16/00 (2006.01)

C07K 16/22 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.01.2011 E 11701406 (8)**

97 Fecha y número de publicación de la concesión europea: **12.08.2015 EP 2528944**

54 Título: **Genotecas combinatorias de anticuerpos de roedor**

30 Prioridad:

29.01.2010 EP 10152164

29.01.2010 US 299380 P

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

24.11.2015

73 Titular/es:

**MORPHOSYS AG (100.0%)
Lena-Christ-Strasse 48
82152 Martinsried/Planegg, DE**

72 Inventor/es:

**PRASSLER, JOSEF;
OTT, DAVID;
THIEL, STEFANIE;
STARK, YVONNE;
KECK, UTE;
PIETZONKA, THOMAS y
EBERSBACH, HILMAR**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 551 871 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

Genotecas combinatorias de anticuerpos de roedor

Campo de la invención

5 La presente invención se refiere y proporciona genotecas murinas, tales como genotecas combinatorias de anticuerpos que son adecuadas para la selección de anticuerpos murinos completos. La invención también se refiere a secuencias de ácidos nucleicos sintéticos que codifican anticuerpos individuales o colecciones de tales anticuerpos murinos, es decir, secuencias de ácido nucleico que codifican genotecas de anticuerpos murinos. Se proporcionan métodos para la generación y el uso de tales genotecas. En particular, la invención se refiere a la preparación de una genoteca de genes de anticuerpos obtenidos a partir de ratones mediante el uso de homólogos sintéticos de
10 consenso y/o homólogos muy similares a secuencias de la línea germinal que incluyen el repertorio estructural de anticuerpos codificados en el genoma.

Antecedentes de la invención

15 En los últimos años, muchos anticuerpos humanizados o completamente humanos se han sometido a ensayos clínicos, y varios de ellos han sido aprobados para una aplicación terapéutica (Hoogenboom, H.R. (2005), *Nat Biotechnol* 23, 1105-1116; Carter, P.J. (2006). *Nat Rev Immunol* 6, 343-357). Aunque no es probable que los anticuerpos humanos seleccionados normalmente mediante presentación en fagos (Hoogenboom, 2005) sean inmunogénicos y, por lo tanto, que sean neutralizados rápidamente en humanos, por lo general son víctimas de un rápido aclaramiento en los animales de experimentación debido a la inducción de anticuerpos anti-humanos. El rápido aclaramiento de anticuerpos humanos impide estudios preclínicos a largo plazo de validación de dianas, de mecanismos de acción y
20 ADME (PK/PD) en roedores. Por estas razones, es muy deseable disponer de anticuerpos que no sean inmunogénicos en los modelos animales respectivos. Más preferiblemente, tales anticuerpos se obtienen completamente a partir de las especies animales respectivas. Por ejemplo, un anticuerpo usado en experimentos con animales murinos es idealmente un anticuerpo totalmente murino. Del mismo modo, un anticuerpo usado en experimentos animales con ratas es de forma ideal un anticuerpo totalmente de rata.

25 Para la generación de anticuerpos humanos se han utilizado genotecas combinatorias sintéticas de anticuerpos, por ejemplo, la genoteca HuCAL GOLD descrita por Rothe et al., 2008, *J Mol Biol* 376, 1182-1200. La selección mediante presentación en fagos de ligandos de alta afinidad, bien expresados, selectivos de una diana, se demostró en esta genoteca totalmente humana (Steidl et al. 2008. *Mol Immunol* 46(1):135-44). Las genotecas combinatorias sintéticas ofrecen varias ventajas sobre las genotecas obtenidas a partir de fuentes de linfocitos B no activados. Las secuencias de genes presentes en la naturaleza que codifican regiones estructurales (del inglés, "framework") de anticuerpos, se pueden seleccionar y diseñar fácilmente para que tengan un alto rendimiento en la producción de anticuerpos en hospedadores procariontes o eucariotes, por ejemplo, mediante optimización de codones. Además, la naturaleza sintética de las genotecas permite la aplicación de otras características, por ejemplo, un diseño modular de las regiones estructurales de anticuerpos con sitios de restricción únicos. Tales sitios de restricción introducidos intencionalmente en posiciones adecuadas, permiten procesos de optimización posteriores, tales como la maduración del anticuerpo; por ejemplo, agrupaciones de ligandos se pueden optimizar con rapidez sin tener un conocimiento de las secuencias particulares, utilizando casetes de regiones determinantes de complementariedad (CDRs) diversificados, construidos previamente. Sobre todo, este proceso de optimización está dirigido por una afinidad hacia el antígeno más elevada y por eso conocida como maduración por afinidad (Steidl et al. 2008. *Mol Immunol* 46(1):135-44).

40 Una metodología similar no se consideró que fuera posible para las genotecas de anticuerpos de roedor. Numerosos informes muestran la creencia de la comunidad científica de que, por ejemplo, las genotecas de anticuerpos de ratón no se pueden producir con una complejidad suficiente y/o con niveles de expresión razonablemente útiles. Históricamente, las genotecas de anticuerpos de ratón se obtenían mediante amplificación por PCR de una agrupación de genes VH a partir de bazos de ratones inmunizados (Ward et al., 1989, *Nature* 341, 544-546). Sin embargo, tales genotecas estaban cargadas de problemas, tales como por ejemplo, una expresión débil de los miembros de la genoteca.

50 Solo recientemente, se ha informado sobre la primera genoteca sintética de anticuerpos de ratón (Cobaugh et al., 2008, *Journal of Molecular Biology* 378, 622-633). Pero en este caso, la genoteca estaba enfocada hacia ligandos peptídicos, basándose en una región estructural única obtenida a partir del anticuerpo 26-10, y altamente diversificada solamente en HCDR3.

55 El documento WO 06/030238 describe el "sorprendente descubrimiento de que una región estructural de anticuerpo basada en una cadena pesada VH14 murina y una cadena ligera VK2 murina [...] se expresa de forma soluble en un microorganismo y es estable [...]". Véanse las líneas 15-28 en la página 3 del documento WO 06/030238. Esto resume todos los intentos fallidos anteriores en la técnica para elaborar una genoteca de murino similar o de otros roedores. En particular, el documento WO 06/030238 solo informó sobre la expresión soluble de una y solo una pareja VH/VL particular. En consecuencia, hasta el momento no se ha informado sobre otras genotecas de roedor que hayan tenido éxito. Además, no se ha descrito ninguna genoteca de anticuerpos de roedor que comprenda más de una pareja VH/VL específica.

A pesar de estas referencias desalentadoras, los presentes inventores deseaban la generación de genotecas de anticuerpos de naturaleza roedora, tal como una naturaleza totalmente murina y totalmente de rata, respectivamente, y/o colectivamente. Tales genotecas podrían resultar adecuadas para la selección de anticuerpos completamente murinos y de rata, con propiedades biofísicas deseadas y especificidad de la diana. Uno de los principales objetivos era construir una genoteca combinatoria sintética de anticuerpos totalmente murina (en lo sucesivo HuCAL MOUSE) y una genoteca combinatoria de anticuerpos totalmente de rata (en lo sucesivo HuCAL RAT), que comprendieran preferiblemente la mayor cantidad posible de regiones estructurales VL/VH diferentes.

El documento WO 01/92291 describe un método para el corte y empalme en trans de ARN mediado por una fusión covalente intracelular de transcritos de dos genes diferentes. En el Ejemplo 9, una genoteca de ratón de genes dTS (doble corte y empalme en trans) se construye mediante comparación y los exones líder de 41 genes funcionales VH de ratón. Sin embargo, no se generaron genotecas de anticuerpos de ratón.

De Jaeger et al (FEBS Letters (1997) 403, 116-22) utilizan una genoteca murina de presentación en fagos de cadenas sencillas para el aislamiento de ligantes para a una enzima procedente de *Petunia hybrida*. Sin embargo, esta genoteca no es una genoteca sintética. Además, la genoteca de Jaeger et al. se obtiene a partir de ratones inmunizados, es decir, la genoteca se generó a partir de material de ácido nucleico que se había expuesto previamente al antígeno.

Sommavilla et al (J Immunol Methods (2010) 353, 31-43) informan sobre el diseño y la construcción de una genoteca de anticuerpos de ratón sin activar. Al igual que otros intentos, las genotecas de Sommvilla et al. solo contienen un gen VH y VL de la línea germinal. Estas genotecas de un "solo recipiente" se pueden utilizar p. ej., en modelos de ratón singénico.

Okamoto et al (Biochem Biophys Res Comm (2004) 323, 583-91) describen una genoteca de anticuerpos murinos. Sin embargo, esta genoteca se generó mediante la transcripción inversa de un ARNm procedente de células de bazo y de médula ósea de ratones y, por lo tanto, no es una genoteca de anticuerpos totalmente sintética.

Basándose en los conocimientos de los inventores, no se ha descrito previamente ni una genoteca sintética de anticuerpos de roedor ni ninguna otra genoteca de anticuerpos de roedor fiable con una composición diversa de VH/VL, que comprenda miembros de más de una familia de la línea germinal de VH y/o de VL. Uno de los logros de la presente descripción se basa en la observación de que diferentes regiones estructurales fuerzan conformaciones distintas en las CDRs, y por lo tanto afectan a los tipos de estructuras de anticuerpo capaces de unirse al antígeno. Este es uno de los principales problemas en la generación de anticuerpos de roedor, y una de las razones por las que hasta ahora no se había generado ninguna genoteca de anticuerpos de roedor como se contempla en la presente invención. Por primera vez ha sido posible superar las desventajas de la técnica anterior. Un análisis racional de las secuencias de anticuerpos de roedor presentes en la naturaleza y un diseño sofisticado de las genotecas, ha conducido a las primeras genotecas sintéticas de anticuerpos de roedor que son muy útiles en la investigación biomédica.

Sumario de la invención

Puesto que se espera que las regiones estructurales codificadas por secuencias de la línea germinal murina no sean inmunogénicas en ratones, hemos realizado búsquedas en la base de datos IMGT (Giudicelli, et al. (1997), Nucleic Acids Res 25, 206-211) de secuencias de la línea germinal que representan las clases más abundantes de anticuerpos murinos. Sometimos a ensayo seis secuencias representativas de la línea germinal de VL kappa murina (3 para la cepa de ratón C57BL/6 y 3 para la cepa de ratón Balb/c) y cinco secuencias representativas de la línea germinal de VH murina (además de una variante de una secuencia de la línea germinal de mVH1, denominada mVH1_RV). Además de las 36 combinaciones posibles de mVL-mVH, identificamos cinco combinaciones que mostraban las siguientes propiedades ventajosas: (i) se muestran bien en la punta de fagos filamentosos en formato Fab, (ii) se expresan en grandes cantidades y en forma soluble en formato Fab, así como en formatos IgG, (iii) son termodinámicamente estables y (iv) se identifican con frecuencia en la base de datos de anticuerpos de Kabat.

En las cinco combinaciones más favorables de mVL-mVH, las regiones L-CDR3 y H-CDR3 se reemplazaron por casetes de genotecas altamente diversificadas de L-CDR3 y H-CDR3, respectivamente, consiguiendo de este modo una diversidad total de la genoteca superior a 1×10^{10} . Para los fines de maduración por afinidad de los ligantes obtenidos a partir de las rondas iniciales de inmunoadsorción, los casetes de maduración de L-CDR3 y H-CDR2 se sintetizaron y se clonaron en un vector auxiliar.

Además de la modularidad, se implementaron otras características del sistema HuCAL GOLD (Rothe y col., 2008, J. Mol Biol 376, 1182-1200), así como, por ejemplo, la tecnología de CysDisplay[®] (véase el documento WO 00/105950). Los ligantes seleccionados a partir de la genoteca sintética de anticuerpos de roedor se pueden utilizar, por ejemplo, para la validación de dianas en el ratón. Además, anticuerpos murinos o de rata seleccionados que presentan un modo de unión similar al de sus homólogos específicos de una diana obtenidos a partir de HuCAL GOLD, se pueden usar como sustitutos no inmunogénicos en experimentos respectivos con animales, por ejemplo, para estudios de toxicología y farmacología.

La presente descripción se refiere en general a genotecas sintéticas de anticuerpos de roedor, preferiblemente a genotecas que comprenden miembros de al menos dos familias de la línea germinal de VH y/o al menos dos familias

de la línea germinal de VL. Específicamente, la presente descripción proporciona genotecas sintéticas de anticuerpos de ratón y de rata, aunque en la descripción se consideran las genotecas de otras especies de roedor.

5 La presente descripción también proporciona genotecas sintéticas de anticuerpos de roedor que comprenden más del 50% del repertorio natural de VH de ratón y/o más del 50% del repertorio natural de VL de ratón. Dichas genotecas incluyen, además, más del 50% de las longitudes de H-CDR3 presentes en la naturaleza del repertorio de H-CDR3 murino.

La presente descripción también proporciona genotecas sintéticas de anticuerpos de roedor en las que se presentan de manera eficaz esencialmente todas las combinaciones de VH-VL de dicha genoteca.

10 La presente descripción también proporciona genotecas sintéticas de anticuerpos de roedor en las que esencialmente todas las combinaciones de VH-VL se expresan bien en *E. coli* en formato Fab.

La presente descripción también proporciona genotecas sintéticas de anticuerpos de roedor en las que esencialmente todas las combinaciones de VH-VL se expresan bien en un sistema de mamífero en formato IgG.

La presente descripción también proporciona genotecas sintéticas de anticuerpos de roedor en las que todas las combinaciones de VH-VL son térmicamente estables.

15 La presente descripción también proporciona colecciones de moléculas de ácido nucleico que codifican los anticuerpos de dichas genotecas sintéticas de anticuerpos de roedor.

La presente descripción también proporciona vectores que codifican dichas moléculas de ácido nucleico.

La presente descripción también proporciona células hospedadoras recombinantes que comprenden dichas moléculas o vectores de ácido nucleico.

20 La presente descripción también proporciona métodos para aislar anticuerpos específicos de un antígeno, comprendiendo dicho método las etapas de:

(a) poner en contacto las genotecas sintéticas de anticuerpos de roedor de la presente descripción con un antígeno;

(b) eliminar aquellos miembros de la genoteca que no se unen (o que no sean específicos) al antígeno; y

25 (c) recuperar aquellos miembros de la genoteca unidos (o específicos) al antígeno.

La presente descripción también proporciona un anticuerpo aislado a partir de una genoteca contemplada por la presente descripción, que incluye un cultivo aislado modificado, tal como un anticuerpo humanizado. La presente descripción también proporciona una variante de anticuerpo obtenida a partir de un anticuerpo aislado a partir de una genoteca contemplada por la presente descripción, incluyendo variantes de anticuerpos, en donde dicha variante de anticuerpo es un anticuerpo humanizado.

Muy específicamente, la presente invención se refiere a una genoteca completamente sintética de anticuerpos murinos que comprende regiones estructurales codificadas por los genes de la línea germinal de VH IGHV1-72*01 (VH1), IGHV2-2*01 (VH2) y IGHV5-9*04 (VH5) y al menos dos de los siguientes genes de la línea germinal de VL-kappa: IGKV1-117*01 (Vk1), IGKV3-12*01 (Vk3) y IGHV3-4*01 (Vk3).

35 La selección de las regiones estructurales fue realizada para optimizar la posibilidad de obtener anticuerpos que poseen propiedades biofísicas favorables y que carecen de los fallos de los anticuerpos obtenidos a partir de genotecas sintéticas que no han sido sometidas a maduración *in vivo*. Tales propiedades biofísicas favorables y deseadas incluyen una mayor estabilidad, mayores niveles de expresión y una tendencia baja a la agregación.

Breve descripción de los dibujos

40 La **Figura 1** muestra una alineación de la región estructural 3 de mVH1 y la variante mVH1_RV. Esta variante comprende las mutaciones K77R y T78V. En esta variante se forma un puente salino de estabilización entre R77 y D100, que añade una alta estabilidad extra.

La **Figura 2** muestra la región estructural 4 para las cadenas pesadas y la genoteca de ratón.

La **Figura 3** muestra la región estructural 4 para las cadenas ligeras y la genoteca de ratón.

45 La **Figura 4** muestra la distribución natural de aminoácidos de las secuencias de HCDR3 murinas para secuencias de HCDR3 con una longitud de 7, 8, 9, 10, 11 y 12 aminoácidos. Las distribuciones de aminoácidos aislados están alineadas mediante la agrupación del segmento D (que conduce a una alineación derecha centrada).

La **Figura 5** resume todas las longitudes en un diseño de HCDR3 murina. La media ponderada se calculó para cada posición.

La **Figura 6** muestra el diseño de HCDR3 murina después de implementar todas las modificaciones necesarias.

La **Figura 7** muestra la distribución natural de aminoácidos de las secuencias de LCDR3 murinas con una longitud de 8 aminoácidos.

La **Figura 8** muestra el diseño de LCDR3 murina después de implementar todas las modificaciones necesarias.

5 La **Figura 9** muestra la distribución natural de aminoácidos de las secuencias murinas de HCDR2 para los tres genes maestros mVH seleccionados, mVH1_RV, mVH2 y mVH5. Con el fin de conservar la prevalencia de las secuencias de la línea germinal, los residuos alternativos presentes en los anticuerpos reordenados pero no en las secuencias de la línea germinal, se representaron intencionalmente de forma insuficiente en el diseño.

10 La **Figura 10** resume la distribución de aminoácidos dependiente de la posición de la HCDR2 de los tres genes maestros de mVH seleccionados.

La **Figura 11** muestra el diseño de HCDR2 murina después de implementar todas las modificaciones necesarias.

La **Figura 12** muestra los genes maestros murinos en formato Fab en fagos filamentosos. Se sometieron a ensayo las 36 combinaciones de VL-VH murinas y todas las combinaciones analizadas, con excepción de todas las combinaciones de mVL-B7, se presentaban bien en los fagos.

15 La **Figura 13** representa la expresión soluble de los genes maestros murinos en formato Fab. Todas las combinaciones de VL-VH murinas sometidas a ensayo, con la excepción de todas las combinaciones de mVL-B7, se expresaban bien en *E. coli*.

La **Figura 14** muestra que la expresión soluble en *E. coli* y la presentación en fagos filamentosos se correlaciona muy bien.

20 Las **Figuras 15 y 16** muestran la expresión de IgGs (IgG1 e IgG2a) murinas. Esencialmente todas las combinaciones sometidas a ensayo mostraban niveles de expresión muy por encima de 10 mg/L.

La **Figura 17** muestra un resumen de los rendimientos de la expresión de todas las combinaciones de regiones estructurales de ratón sometidas a ensayo, después de lavar en IMAC con y sin isopropanol en un código de color para los criterios de selección.

25 La **Figura 18** ilustra la estabilidad térmica de genes maestros murinos en formato Fab. La estabilidad térmica de todas las combinaciones de VL-VH murinas sometidas a ensayo, incluyendo las combinaciones de mVLB7, era razonablemente alta.

La **Figura 19** muestra el diseño de los casetes de HCDR3 por separado para las secuencias de HCDR3 con longitudes de 7, 8, 9, 10, 11 y 12 aminoácidos.

30 La **Figura 20** muestra la representación de los aminoácidos reales encontrados en la HCDR3 de la genoteca murina de la presente descripción, en comparación con la aparición prevista de estos residuos. "T" en la segunda fila se refiere a mezclas de trinucleótidos utilizadas para la síntesis de los oligonucleótidos diversificados para el casete de CDR. Las letras "F" y "P" indican la incidencia encontrada y planeada de residuos, respectivamente.

35 La **Figura 21** muestra la representación de los aminoácidos reales encontrados en la HCDR3 de la genoteca murina de la presente descripción, en comparación con la incidencia prevista de estos residuos. "T" y "W" en la segunda fila se refieren a mezclas de trinucleótidos y nucleótidos oscilantes (del inglés "wobble"), utilizados respectivamente, para la síntesis de los oligonucleótidos diversificados para el casete de CDR. Las letras "F" y "P" indican una incidencia encontrada y planeada de residuos, respectivamente.

40 La **Figura 22** muestra la representación de los aminoácidos reales encontrados en el casete de maduración de HCDR2 de la subgenoteca murina de VH1 RV de la presente descripción, en comparación con la incidencia prevista de estos residuos. "MMT" y "W" en la segunda fila se refieren a mezclas de trinucleótidos y nucleótidos oscilantes, respectivamente, utilizados para la síntesis de los oligonucleótidos diversificados para el casete de maduración. Las letras "F" y "P" indican una incidencia encontrada y planeada de residuos, respectivamente.

45 La **Figura 23** muestra la representación de los aminoácidos reales encontrados en el casete de maduración de HCDR2 de la subgenoteca murina de VH2 de la presente descripción, en comparación con la incidencia prevista de estos residuos. "MMT" y "W" en la segunda fila se refieren a mezclas de trinucleótidos y nucleótidos oscilantes, respectivamente, utilizados para la síntesis de los oligonucleótidos diversificados para el casete de maduración. Las letras "F" y "P" indican una incidencia encontrada y planeada de residuos, respectivamente.

50 La **Figura 24** muestra la representación de los aminoácidos reales encontrados en el casete de maduración de HCDR2 de la subgenoteca murina de VH5 de la presente descripción, en comparación con la incidencia prevista de estos residuos. "MMT" y "W" en la segunda fila se refieren a mezclas de trinucleótidos y nucleótidos oscilantes, respectivamente, utilizados para la síntesis de los oligonucleótidos diversificados para el casete de maduración. Las

letras "F" y "P" indican una incidencia encontrada y planeada de residuos, respectivamente.

5 La **Figura 25** muestra la representación de los aminoácidos reales encontrados en el casete de maduración de HCDR2 de la genoteca de rata de la presente descripción, en comparación con la incidencia prevista de estos residuos. "MMT" y "W" en la segunda fila se refieren a mezclas de trinucleótidos y nucleótidos oscilantes, respectivamente, utilizados para la síntesis de los oligonucleótidos diversificados para el casete de maduración. Las letras "F" y "P" indican una incidencia encontrada y planeada de residuos, respectivamente.

La **Figura 26** muestra la distribución de la longitud de HCDR3 planeada en comparación con la distribución de la longitud de HCDR3 encontrada realmente en la genoteca murina de la presente descripción.

10 Las **Figuras 27 a 31** representan una visión general de todos los genes maestros sintetizados de la genoteca murina de la presente descripción. Kabat indica la posición del residuo de aminoácido en las secuencias de VL o VH. mVL indica dominio variable murino en la cadena ligera (exclusivamente del tipo kappa), y mVH indica dominio variable murino en la cadena pesada. Los genes Vk de la línea germinal murina codifican mVL partiendo de la posición 1 (en el extremo N-terminal de la región estructural 1) hasta 95 (en el extremo C-terminal de L-CDR3), y los genes Jk de la línea germinal murina codifican el residuo muy C-terminal en L-CDR3 (posición 96 constante) y toda la región estructural 4 en mVL (posiciones 97 - 108). Hemos elegido Jk2 murino, ya que Tyr es el residuo más frecuente en la posición 96 de los anticuerpos reordenados que se encuentra en la base de datos IMGT.

La **Figura 32** muestra la diversidad y la abundancia de diferentes combinaciones de regiones estructurales aisladas a partir de una prueba de inmunoabsorción con la genoteca murina de la presente descripción.

20 La **Figura 33** muestra un resumen de los rendimientos obtenidos con fragmentos Fab específicos de mM-GSF después de la expresión y la purificación a partir de células bacterianas.

La **Figura 34** muestra los resultados de la determinación de las constantes de afinidad para algunos fragmentos Fab específicos de mM-CSF tal y como se determinan en un ensayo ELISA.

La **Figura 35** muestra el diseño de los dos grupos de genes maestros VL de rata.

La **Figura 36** muestra la región estructural 4 del gen maestro VH de rata.

25 La **Figura 37** muestra la región estructural 4 del gen maestro VL de rata.

La **Figura 38** muestra los genes maestros de rata en formato Fab sobre fagos filamentosos. Las cuatro combinaciones de regiones estructurales rVL-rVH sometidas a ensayo se presentan claramente sobre los fagos.

30 La **Figura 39** representa la expresión soluble de los genes maestros de rata en formato Fab. Las combinaciones de rVL1 se expresaban particularmente bien. Se muestran los resultados de dos experimentos individuales (marcados como "NVS" y "MOR"), así como la media de estos dos experimentos.

La **Figura 40** muestra la expresión de IgG2b de rata. No hay mucha diferencia entre las cuatro combinaciones de rata sometidas a ensayo.

La **Figura 41** muestra la expresión de IgG2c de rata. La mejor expresión se observa con las combinaciones de rVH5.

35 La **Figura 42** muestra un resumen de los rendimientos de la expresión de todas las combinaciones de regiones estructurales de rata Fab sometidas a ensayo después de lavar sobre IMAC con y sin isopropanol, con un código de color para los criterios de selección.

La **Figura 43** ilustra la estabilidad térmica de los genes maestros murinos en formato Fab e IgG. La estabilidad térmica de todas las combinaciones de VL-VH de rata sometidas a ensayo era razonablemente elevada.

La **Figura 44** muestra la variabilidad introducida en la HCDR2 de la genoteca de rata HuCAL sin activar.

40 **Figura 45** muestra la especificidad de cuatro fragmentos Fab de rata hacia GM-CSF. Los cuatro fragmentos Fab eran altamente específicos.

La **Figura 46** muestra que los ligantes aislados a partir de la genoteca murina de la presente descripción conservan su actividad funcional después de la conversión desde formato Fab a formato de inmunoglobulina de longitud completa.

45 **Descripción detallada de la invención**

A menos que se defina de otro modo, todos los términos técnicos y científicos utilizados en esta memoria tienen el mismo significado que el comúnmente entendido por los que tienen conocimientos ordinarios en la técnica a la que pertenece esta invención.

Los términos "alrededor de" o "aproximadamente" en el contexto de valores e intervalos numéricos, se refieren a

valores o intervalos que se aproximan o están cerca de los valores o intervalos citados, de modo que la invención se puede llevar a cabo según lo previsto, tal como tener un número o un porcentaje deseado de homología de secuencia, como es evidente para la persona experta partiendo de las enseñanzas contenidas en el presente documento. Esto es debido, al menos en parte, a las condiciones de cultivo diferentes y a la variabilidad de los sistemas biológicos. Por lo tanto, estos términos abarcan valores más allá de los que se originan por un error sistemático. Estos términos hacen explícito lo que está implícito.

Todos los intervalos establecidos en este documento en el resumen y la descripción de la memoria descriptiva incluyen todos los números o valores entre los mismos o próximos a los números del intervalo. Los intervalos de la descripción se denominan expresamente y exponen todos los valores enteros, decimales y fraccionarios en el intervalo. El término "aproximadamente" se puede utilizar para describir un intervalo.

El término "anticuerpo" tal y como se usa en este documento, incluye anticuerpos completos y cualquier fragmento que se une a antígeno (es decir, "porción que se une a antígeno") o cadenas sencillas del mismo. Un "anticuerpo" de origen natural es una glicoproteína que comprende al menos dos cadenas pesadas (H) y dos cadenas ligeras (L) interconectadas por enlaces disulfuro. Cada cadena pesada está compuesta de una región variable de la cadena pesada (abreviada en este documento como VH) y una región constante de la cadena pesada. La región constante de la cadena pesada se compone de tres dominios, CH1, CH2 y CH3. Cada cadena ligera está compuesta por una región variable de la cadena ligera (abreviada en este documento como VL) y una región constante de la cadena ligera. La región constante de la cadena ligera está compuesta por un dominio, CL. Las regiones VH y VL se pueden subdividir adicionalmente en regiones de hipervariabilidad, denominadas regiones determinantes de complementariedad (CDR), intercaladas entre regiones que están más conservadas, denominadas regiones estructurales (FR). Cada VH y VL se compone de tres CDRs y cuatro FRs dispuestas desde el extremo amino-terminal al extremo carboxi-terminal en el siguiente orden: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. Las regiones variables de las cadenas pesada y ligera contienen un dominio de unión que interacciona con un antígeno. Las regiones constantes de los anticuerpos pueden mediar en la unión de la inmunoglobulina a tejidos del hospedador o a factores, incluyendo diversas células del sistema inmune (por ejemplo, células efectoras) y el primer componente (C1 q) del sistema del complemento clásico.

Las expresiones "porción que se une a antígeno" o "fragmento" de un anticuerpo, se utilizan de manera equivalente en la presente solicitud. Estos términos se refieren a uno o varios fragmentos de un anticuerpo intacto que conservan la capacidad de unirse específicamente a un antígeno dado. Las funciones de unión a un antígeno de un anticuerpo pueden ser realizadas por fragmentos de un anticuerpo intacto. Ejemplos de fragmentos de unión incluidos dentro de la expresión "porción que se une a antígeno" de un anticuerpo, incluyen un fragmento Fab, un fragmento monovalente que consiste en los dominios VL, VH, CL y CH1; un fragmento F(ab)₂, un fragmento bivalente que comprende dos fragmentos Fab unidos por un puente disulfuro en la región bisagra; un fragmento Fd que consiste en los dominios VH y CH1; un fragmento Fv que consiste en los dominios VL y VH de un único brazo de un anticuerpo; un fragmento de anticuerpo de dominio único (dAb) (Ward et al., 1989 Nature 341: 544-546), que consiste en un dominio VH; y una región determinante de complementariedad aislada (CDR). Las porciones o fragmentos preferidos de anticuerpos que se unen a antígenos son los fragmentos Fab.

Además, aunque los dos dominios del fragmento Fv, VL y VH, están codificados por genes distintos, se pueden unir, usando métodos recombinantes, a través de un enlazador peptídico artificial que les permite formar una cadena proteica sencilla en la que las parejas de las regiones VL y VH forman moléculas monovalentes (conocidas como Fv de cadena sencilla (scFv); véase, por ejemplo, Bird et al., 1988 Science 242:423-426; y Huston et al., 1988 Proc. Natl. Acad. Sci. 85:5879-5883). Tales anticuerpos de cadena sencilla incluyen una o varias "porciones que se unen a antígeno" de un anticuerpo. Estos fragmentos de anticuerpo se obtienen usando técnicas convencionales conocidas por los expertos en la técnica, y los fragmentos se escrutan en busca de una utilidad de la misma manera que los anticuerpos intactos.

Las porciones que se unen a antígeno también se pueden incorporar en anticuerpos de dominio sencillo, maxicuerpos, minicuerpos, intracuerpos, diacuerpos, triacuerpos, tetracuerpos, v-NAR y bis-scFv (véase, por ejemplo, Hollinger y Hudson, 2005, Nature Biotechnology, 23, 9, 1126-1136). Las porciones de anticuerpos que se unen a antígeno se pueden injertar en estructuras basadas en polipéptidos, tales como fibronectina de tipo III (Fn3) (véase el documento de patente de EE.UU. n° 6.703.199, que describe monocuerpos polipeptídicos de fibronectina).

Las porciones que se unen a antígeno se pueden incorporar en moléculas de cadena sencilla que comprenden una pareja de segmentos Fv en tándem (VH-CH1-VH-CH1) que, junto con los polipéptidos de la cadena ligera complementaria, forman una pareja de regiones que se unen a antígeno (Zapata et al., 1995 Protein Eng. 8(10):1057-1062 y el documento de patente de EE.UU. n° 5.641.870).

Tal como se utiliza en esta memoria, el término "afinidad" se refiere a la fuerza de la interacción entre un anticuerpo y un antígeno en los sitios antigénicos individuales. Dentro de cada sitio antigénico, la región variable del "brazo" del anticuerpo interacciona a través de fuerzas no covalentes débiles con un antígeno en numerosos sitios; cuantas más interacciones, más fuerte es la afinidad.

Tal como se utiliza en esta memoria, la expresión "afinidad elevada" para un anticuerpo, tal como un anticuerpo IgG,

se refiere a un anticuerpo que tiene una K_D de 10^{-8} M o menor, 10^{-9} M o menor, o 10^{-10} M, o 10^{-11} M o menor hacia un antígeno diana. Sin embargo, una unión de "afinidad elevada" puede variar para otros isotipos de anticuerpo. Por ejemplo, unión con "afinidad elevada" hacia un isotipo IgM se refiere a un anticuerpo que tiene una K_D de 10^{-7} M o menor, o 10^{-8} M o menor.

- 5 La expresión "especificidad de la unión", tal como se usa en esta memoria, se refiere a la capacidad de un sitio de combinación de un anticuerpo individual para reaccionar con solamente un determinante antigénico. El sitio de combinación del anticuerpo se encuentra en la porción Fab de la molécula y está construido a partir de las regiones hipervariables de las cadenas pesadas y ligeras. La afinidad de la unión de un anticuerpo es la fuerza de la reacción entre un determinante antigénico aislado y un sitio de combinación aislado en el anticuerpo. Es la suma de las fuerzas de atracción y repulsión que operan entre el determinante antigénico y el sitio de combinación del anticuerpo.

- 10 Unión específica entre dos entidades significa una unión con una constante de equilibrio (K_A) de al menos $1 \times 10^7 M^{-1}$, $10^8 M^{-1}$, $10^9 M^{-1}$, $10^{10} M^{-1}$ o $10^{11} M^{-1}$. La expresión "se une específicamente (o selectivamente)" a, se refiere a una reacción de unión que es determinante de la presencia de un cognado en una población heterogénea de proteínas y otros productos biológicos. Además de la constante de equilibrio (K_A) indicada anteriormente, un anticuerpo de la descripción también tiene típicamente una constante de tasa de disociación (K_d) de aproximadamente $1 \times 10^2 s^{-1}$, $1 \times 10^3 s^{-1}$, $1 \times 10^4 s^{-1}$, $1 \times 10^4 s^{-1}$, o inferior, y se une a un antígeno diana con una afinidad que es al menos dos veces mayor que su afinidad de unión hacia un antígeno no específico. Las expresiones "un anticuerpo que reconozca un antígeno" y "un anticuerpo específico de un antígeno" se usan indistintamente en el presente documento junto con la expresión "un anticuerpo que se une específicamente a un antígeno".

- 20 Las expresiones "bloqueo cruzado", "bloqueo de forma cruzada" y "bloqueando de forma cruzada" se utilizan indistintamente en este documento para referirse a la capacidad de un anticuerpo u otro agente de unión para interferir en la unión de otros anticuerpos o agentes de unión con la misma diana, en un ensayo convencional de unión competitiva.

- 25 La capacidad o el grado con el que un anticuerpo u otro agente de unión es capaz de interferir en la unión de otro anticuerpo o molécula de unión con la misma diana y, por tanto, si se puede decir bloqueo cruzado de acuerdo con la descripción, se puede determinar usando ensayos de unión competitiva convencionales. Un ensayo adecuado implica el uso de la tecnología Biacore (por ejemplo, utilizando el aparato Biacore 3000 (Biacore, Uppsala, Suecia)), que puede medir el alcance de las interacciones utilizando la tecnología de resonancia de plasmón superficial. Otro ensayo para medir el bloqueo cruzado utiliza una metodología basada en ELISA.

- 30 El término "epítipo" significa un determinante proteico capaz de unirse específicamente a un anticuerpo. Los epítipos consisten habitualmente en agrupaciones superficiales químicamente activas de moléculas tales como aminoácidos o cadenas laterales de azúcar y normalmente tienen características estructurales tridimensionales específicas, así como características de carga específicas. Los epítipos conformacionales y no conformacionales se distinguen en que la unión del primero pero no del último se pierde en presencia de disolventes desnaturantes.

- 35 La expresión "anticuerpo quimérico" es una molécula de anticuerpo en la que (a) la región constante, o una porción de la misma, se altera, se sustituye o se intercambia de modo que el sitio de unión al antígeno (región variable) está unido a una región constante de una clase diferente o alterada, una función y/o especie efectora o una molécula completamente diferente que confiere nuevas propiedades al anticuerpo quimérico, por ejemplo, una enzima, toxina, hormona, factor de crecimiento, fármaco, etc.; o (b) la región variable, o una porción de la misma, está alterada, sustituida o se intercambia con una región variable que tiene una especificidad antigénica diferente o alterada. Por ejemplo, un anticuerpo de ratón se puede modificar mediante la sustitución de su región constante con la región constante de una inmunoglobulina humana. Debido a la sustitución con una región constante humana, el anticuerpo quimérico puede conservar su especificidad para el reconocimiento del antígeno, a la vez que tiene una inmunogenicidad reducida en humanos en comparación con el anticuerpo de ratón original.

- 45 Un anticuerpo "humanizado" es un anticuerpo que conserva la reactividad de un anticuerpo no humano a la vez que es menos inmunogénico en seres humanos. Esto se puede lograr, por ejemplo, conservando las regiones CDRs no humanas y sustituyendo las partes restantes del anticuerpo con sus homólogas humanas (es decir, la región constante, así como las porciones de región estructural de la región variable). Véanse, por ejemplo, Morrison et al., Proc. Natl. Acad. Sci. USA, 81:6851-6855, 1984; Morrison y Oi, Adv. Immunol., 44:65-92, 1988; Verhoeven et al., Science, 239:1534-1536, 1988; Padlan, Molec. Immun., 28:489-498, 1991; y Padlan, Molec. Immun., 31:169-217, 1994. Otros ejemplos de tecnología con modificación genética humana incluyen, pero no se limitan a la tecnología Xoma descrita en el documento US 5.766.886.

- 55 La expresión "anticuerpo humano", tal y como se emplea en esta memoria, se entiende que incluye anticuerpos que tienen regiones variables en las que tanto las regiones estructurales como las regiones CDRs se obtienen a partir de secuencias de origen humano. Además, si el anticuerpo contiene una región constante, la región constante también se obtiene a partir de tales secuencias humanas, por ejemplo, secuencias de la línea germinal humana, o versiones mutadas de las secuencias de la línea germinal humana. Los anticuerpos humanos de la descripción pueden incluir residuos de aminoácidos no codificados por secuencias humanas (por ejemplo, mutaciones introducidas por mutagénesis aleatoria o específica del sitio *in vitro* o por mutación somática *in vivo*).

- La expresión "anticuerpo de roedor", tal y como se emplea en esta memoria, se entiende que incluye anticuerpos que tienen regiones variables en las que tanto las regiones estructurales como las CDRs se obtienen a partir de secuencias con origen de roedor. Por ejemplo, tanto las regiones estructurales como las CDRs se pueden obtener a partir de secuencias con origen de ratón, o tanto las regiones estructurales como las CDRs se pueden obtener a partir de secuencias con origen de rata. Además, si el anticuerpo contiene una región constante, la región constante también se obtiene a partir de tales secuencias de roedor, por ejemplo, secuencias de la línea germinal de roedor, o versiones mutadas de secuencias de la línea germinal de roedor. Los anticuerpos de roedor de la descripción pueden incluir residuos de aminoácidos no codificados por secuencias de roedor (por ejemplo, mutaciones introducidas *in vitro* mediante mutagénesis aleatorias o específicas del sitio o por mutación somática *in vivo*).
- El término "roedor" es conocido en la técnica e incluye las siguientes especies: ratón, rata, ardilla, ardilla listada, tuza, puercoespín, castor, hámster, jerbo, conejillo de indias, degú, chinchilla, perro de las praderas y marmota. En ciertos aspectos de la presente invención, dicho ratón se selecciona a partir de las cepas de laboratorio Balb/c y C57BL/6.
- La expresión "anticuerpo aislado" se refiere a un anticuerpo que está sustancialmente exento de otros anticuerpos antigénicamente diferentes. Un anticuerpo aislado que se une específicamente a un antígeno, sin embargo, puede tener reactividad cruzada con otros antígenos. Además, un anticuerpo aislado puede estar sustancialmente exento de otro material celular y/o de productos químicos.
- El término "isotipo" se refiere a la clase de anticuerpo (por ejemplo, IgM, IgE, IgG tal como IgG1 o IgG4) que está proporcionada por los genes de la región constante de la cadena pesada. El isotipo también incluye versiones modificadas de una de estas clases, cuando se realizan modificaciones para alterar la función de Fc, por ejemplo, para aumentar o reducir las funciones efectoras o de unión a los receptores de Fc.
- El término "Kasoc" o "Ka", tal y como se emplea en esta memoria, se entiende que se refiere a la tasa de asociación de una interacción anticuerpo-antígeno particular, mientras que el término "Kdis" o "Kd", tal y como se emplea en esta memoria, se entiende que se refiere a la tasa de disociación de una interacción anticuerpo-antígeno particular. El término "K_D", tal y como se emplea en esta memoria, se entiende que se refiere a la constante de disociación, que se obtiene de la relación entre Kd y Ka (es decir, Kd/Ka) y se expresa como una concentración molar (M). Los valores de K_D para los anticuerpos se pueden determinar usando métodos bien establecidos en la técnica. Un método para determinar la K_D de un anticuerpo, es mediante el uso de resonancia de plasmón superficial, o utilizando un sistema biosensor tal como un sistema Biacore®.
- Las expresiones "anticuerpo monoclonal" o "composición de anticuerpo monoclonal" tal y como se emplean en esta memoria, se refieren a una preparación de moléculas de anticuerpo de composición molecular única. Una composición de anticuerpo monoclonal muestra una sola especificidad de unión y una afinidad hacia un epítipo particular.
- Las genotecas de anticuerpos se pueden obtener a partir de inmunoglobulinas o de fragmentos de las mismas, que están sesgadas hacia ciertas especificidades presentes en animales inmunizados o seres humanos inmunizados de forma natural, o infectados. Alternativamente, las genotecas de anticuerpos se pueden obtener a partir de inmunoglobulinas sin activar, o de fragmentos de las mismas, es decir, inmunoglobulinas que no están sesgadas hacia especificidades que se encuentran en el sistema inmune. Tales genotecas se denominan genotecas "no sesgadas". En realizaciones preferidas, la presente descripción proporciona genotecas no sesgadas de anticuerpos, es decir, las genotecas no se exponen previamente al antígeno de interés. Debido a la ausencia de cualquier sesgo, tales genotecas comprenden anticuerpos que se unen a cualquier antígeno diana potencial de interés.
- Típicamente, las genotecas de anticuerpos inmunes se construyen con agrupaciones de genes de VH y VL que se clonan a partir de linfocitos B fuente mediante técnicas de clonación basadas (o relacionadas) en la PCR. De la misma manera, también es posible generar genotecas de anticuerpos no sesgadas, sin activar. Sin embargo, las genotecas de anticuerpos no sesgadas, sin activar también se pueden generar de una manera sintética en la que toda la genoteca está construida enteramente *in vitro*. La tecnología de ADN recombinante se emplea y se puede usar para imitar los sesgos y las redundancias naturales del repertorio de anticuerpos naturales. Tales genotecas de anticuerpos se denominan genotecas "sintéticas" de anticuerpos. La expresión genoteca "totalmente sintética" se refiere a genotecas de anticuerpos que se construyen completamente, es decir, totalmente, de novo mediante la síntesis de ADN, por ejemplo, mediante síntesis génica total, métodos basados en PCR o tecnologías de ADN relacionadas. En tales genotecas, todo el ADN se construye de novo, es decir, la parte que codifica las CDRs, así como las partes (por ejemplo, las regiones estructurales) que codifican las partes que rodean a las CDRs de los anticuerpos de la genoteca. Los términos "sintético" y "totalmente sintético", por tanto, se refieren al origen de novo del ADN. Por el contrario, en una genoteca de anticuerpos "semi-sintética" solo partes de los anticuerpos de la genoteca se construyen de novo, mientras que otras partes, por ejemplo, ciertas regiones de las CDRs, se obtienen a partir de fuentes naturales, (existen numerosas revisiones de esta materia, véase, por ejemplo, Sidhu et al.; Nat Chem Biol (2006), 2, 682-8). En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos de roedor. En un aspecto preferido, la presente invención proporciona una genoteca totalmente sintética de anticuerpos de roedor.
- El término "genoteca" se refiere a una colección de moléculas distintas que comprende típicamente más de 10³, más

de 10^4 , más de 10^5 , más de 10^6 , más de 10^7 , más de 10^8 , más de 10^9 o incluso más de 10^{10} miembros. Una genoteca en el contexto de la presente descripción es una mezcla de polipéptidos o ácidos nucleicos heterogéneos. La genoteca está compuesta de miembros, cada uno de los cuales tiene un solo polipéptido o secuencia de ácido nucleico. En este sentido, genoteca es sinónimo de repertorio. Las diferencias de secuencias entre los miembros de la genoteca son responsables de la diversidad presente en la genoteca. La genoteca puede estar en forma de una simple mezcla de polipéptidos o ácidos nucleicos, o puede estar en forma de organismos o células, por ejemplo, bacterias, virus, células animales o vegetales, transformadas con una genoteca de ácidos nucleicos. Preferiblemente, cada organismo o célula individual contiene solo uno o un número limitado de miembros de la genoteca. Ventajosamente, los ácidos nucleicos se incorporan en vectores de expresión, con el fin de permitir la expresión de los polipéptidos codificados por los ácidos nucleicos. En un cierto aspecto, una genoteca puede estar en forma de una población de organismos hospedadores, conteniendo cada organismo una o varias copias de un vector de expresión que contiene un solo miembro de la genoteca en forma de ácido nucleico que se puede expresar para producir su miembro polipeptídico correspondiente. Por lo tanto, la población de organismos hospedadores tiene el potencial de codificar un gran repertorio de variantes polipeptídicas genéticamente diversas. El término "colección" se utiliza esencialmente de forma intercambiable con el término "genoteca".

El término "línea germinal" se refiere a las secuencias de nucleótidos de los genes de anticuerpos y segmentos de genes, cuando pasan de los progenitores a la descendencia a través de las células germinales. La secuencia de la línea germinal se distingue de las secuencias de nucleótidos que codifican los anticuerpos en los linfocitos B maduros que han sido alteradas por los eventos de recombinación e hipermutación durante el curso de la maduración de linfocitos B.

La expresión "ácido nucleico" se usa en esta memoria de forma intercambiable con el término "polinucleótido" y se refiere a desoxirribonucleótidos o ribonucleótidos y polímeros de los mismos en cualquier forma de cadena sencilla o doble. La expresión abarca ácidos nucleicos que contienen análogos conocidos de nucleótidos o residuos o enlaces modificados de la estructura principal, que son sintéticos, de origen natural y de origen no natural, que tienen propiedades de unión similares a las del ácido nucleico de referencia, y que se metabolizan de una manera similar a la de los nucleótidos de referencia. Ejemplos de tales análogos incluyen, sin limitación, fosforotioatos, fosforamidatos, fosfonatos de metilo, fosfonatos de metilo quirales, 2-O-metil ribonucleótidos y ácidos péptido-nucleicos (PNAs).

A menos que se indique lo contrario, una secuencia de ácido nucleico particular incluye también implícitamente variantes de la misma, modificadas de manera conservadora, (por ejemplo, sustituciones de codones degenerados) y secuencias complementarias, así como la secuencia indicada explícitamente. Específicamente, como se detalla a continuación, las sustituciones de codones degenerados se pueden conseguir generando secuencias en las que se sustituye la tercera posición de uno o varios codones seleccionados (o todos) con residuos de bases mixtas y/o desoxiinosinas (Batzler et al., *Nucleic Acid. Res.* 19:5081, 1991; Ohtsuka et al., *J. Biol. Chem.* 260:2605-2608, 1985; y Rossolini et al., *Mol. Cell. Probes* 8:91-98, 1994).

La expresión "ligado funcionalmente" se refiere a una relación funcional entre dos o más segmentos de polinucleótidos (por ejemplo, ADN). Típicamente, se refiere a la relación funcional de una secuencia reguladora de la transcripción con una secuencia transcrita. Por ejemplo, una secuencia de promotor o potenciador está ligada funcionalmente a una secuencia codificante, si estimula o modula la transcripción de la secuencia codificante en una célula hospedadora apropiada u otro sistema de expresión. Generalmente, las secuencias promotoras reguladoras de la transcripción que se ligan funcionalmente a una secuencia transcrita, están físicamente adyacentes a la secuencia transcrita, es decir, actúan en cis. Sin embargo, algunas secuencias reguladoras de la transcripción, tales como potenciadores, no necesitan estar físicamente adyacentes o estar situadas en estrecha proximidad a las secuencias codificantes cuya transcripción potencian.

Tal como se utiliza en esta memoria, el término, "optimizado" significa que una secuencia de nucleótidos ha sido alterada para codificar una secuencia de aminoácidos usando codones que son preferidos en la célula u organismo de producción, generalmente una célula eucariota, por ejemplo, una célula de *Pichia*, una célula de ovario de hámster chino (CHO) o una célula humana. La secuencia de nucleótidos optimizada está diseñada genéticamente para conservar completamente o tanto como sea posible, la secuencia de aminoácidos codificada originalmente por la secuencia de nucleótidos de partida, que también se conoce como la secuencia "parental". Las secuencias optimizadas en este documento se han diseñado genéticamente para tener codones que son preferidos en células de mamíferos. Sin embargo, la expresión optimizada de estas secuencias en otras células eucariotas o células procariontas también se prevé en el presente documento. Las secuencias de aminoácidos codificadas por secuencias de nucleótidos optimizadas también se conocen como optimizadas.

El término "aminoácido" se refiere a aminoácidos presentes en la naturaleza y sintéticos, así como a análogos de aminoácidos y a miméticos de aminoácidos que actúan de una manera similar a los aminoácidos de origen natural. Aminoácidos de origen natural son los codificados por el código genético, así como aquellos aminoácidos que se modifican posteriormente, por ejemplo, hidroxiprolina, γ -carboxiglutamato y O-fosfoserina. Análogos de aminoácidos se refiere a compuestos que tienen la misma estructura química básica que un aminoácido de origen natural, es decir, un carbono alfa que está unido a un hidrógeno, un grupo carboxilo, un grupo amino y un grupo R, por ejemplo, homoserina, norleucina, sulfóxido de metionina, metionina metil sulfonio. Tales análogos tienen grupos R modificados (por ejemplo, norleucina) o estructuras principales peptídicas modificadas, pero conservan la misma estructura

química básica que un aminoácido de origen natural. Miméticos de aminoácidos se refiere a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que actúan de una manera similar a un aminoácido de origen natural.

5 Los términos "polipéptido" y "proteína" se utilizan indistintamente en este documento para referirse a un polímero de residuos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en los que uno o varios residuos de aminoácidos son un mimético químico artificial de un aminoácido de origen natural correspondiente, así como a polímeros de aminoácidos de origen natural y polímeros de aminoácidos de origen no natural. A menos que se indique lo contrario, una secuencia de polipéptido particular incluye también implícitamente variantes de la misma, modificadas de manera conservadora.

10 La expresión "variante modificada de forma conservadora" se aplica tanto a secuencias de aminoácidos como de ácidos nucleicos. Con respecto a secuencias particulares de ácidos nucleicos, las variantes modificadas de manera conservadora se refieren a aquellos ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o cuando el ácido nucleico no codifica una secuencia de aminoácidos, a secuencias esencialmente idénticas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína dada. Por ejemplo, los codones GCA, GCC, GCG y GCU todos ellos codifican el aminoácido alanina. Por lo tanto, en cada posición en la que una alanina está especificada por un codón, el codón se puede alterar a cualquiera de los codones correspondientes descritos, sin alterar el polipéptido codificado. Tales variaciones del ácido nucleico son "variaciones silenciosas", que son una especie de variaciones modificadas de manera conservadora. Cada secuencia de ácido nucleico en este documento que codifica un polipéptido, también describe cada posible variación silenciosa del ácido nucleico. Un experto reconocerá que cada codón en un ácido nucleico (excepto AUG, que es comúnmente el único codón para metionina, y TGG, que es comúnmente el único codón para triptófano) se puede modificar para producir una molécula funcionalmente idéntica. Por consiguiente, cada variación silenciosa de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

25 Para las secuencias de polipéptidos, las "variantes modificadas de manera conservadora" incluyen sustituciones, deleciones o adiciones individuales de una secuencia de polipéptido que dan como resultado la sustitución de un aminoácido por un aminoácido químicamente similar. Las tablas de sustituciones conservadoras que proporcionan aminoácidos funcionalmente similares son bien conocidas en la técnica. Tales variantes modificadas de manera conservadora se suman y no excluyen variantes polimórficas, homólogos entre especies y alelos de la descripción. Los ocho grupos siguientes contienen aminoácidos que son entre sí sustituciones conservadoras: 1) Alanina (A), Glicina (G); 2) Ácido aspártico (D), Ácido glutámico (E); 3) Asparagina (N), Glutamina (Q); 4) Arginina (R), Lisina (K); 30 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V); 6) Fenilalanina (F), Tirosina (Y), Triptófano (W); 7) Serina (S), Treonina (T); y 8), Cisteína (C), Metionina (M) (véase, por ejemplo, Creighton, *Proteins* (1984)). En algunas realizaciones, la expresión "modificaciones conservadoras de la secuencia" se utiliza para referirse a modificaciones de aminoácidos que no afectan o alteran significativamente las características de la unión, del anticuerpo que contiene la secuencia de aminoácidos.

Los términos "idéntico" o porcentaje de "identidad" en el contexto de dos o más secuencias de ácidos nucleicos o polipeptídicas, se refieren a dos o más secuencias o subsecuencias que son iguales. Dos secuencias son "sustancialmente idénticas" si dos secuencias tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales (es decir, 60% de identidad, opcionalmente 65%, 70%, 75%, 80%, 85%, 90%, 95% o 99% de identidad sobre una región especificada o, si no se especifica, en toda la secuencia), cuando se comparan y se alinean para tener una correspondencia máxima sobre una ventana de comparación, o una región designada según se mide usando uno de los siguientes algoritmos de comparación de secuencias, o mediante alineación manual e inspección visual. Opcionalmente, la identidad existe sobre una región que tiene una longitud de al menos aproximadamente 50 nucleótidos (o 10 aminoácidos), o más preferiblemente sobre una región que tiene una longitud de 100 a 500 o 1000 o más nucleótidos (o 20, 50, 200 o más aminoácidos).

Para la comparación de secuencias, una secuencia actúa típicamente como secuencia de referencia, con la que se comparan las secuencias del ensayo. Cuando se usa un algoritmo de comparación de secuencias, las secuencias del ensayo y de referencia se introducen en un ordenador, se designan las coordenadas de las subsecuencias, si es necesario, y se designan los parámetros del programa del algoritmo de la secuencia. Se pueden utilizar parámetros del programa predeterminados o se pueden designar parámetros alternativos. El algoritmo de comparación de secuencias calcula entonces el porcentaje de identidad de secuencia para las secuencias del ensayo con respecto a la secuencia de referencia, basándose en los parámetros del programa.

Una "ventana de comparación", tal y como se usa en el presente documento, incluye la referencia a un segmento de una cualquiera entre una variedad de posiciones contiguas seleccionadas a partir del grupo que consiste en 20 a 600, por lo general de aproximadamente 50 a aproximadamente 200, más habitualmente de aproximadamente 100 a aproximadamente 150, en donde una secuencia puede compararse con una secuencia de referencia con el mismo número de posiciones contiguas después de que las dos secuencias se alinean de manera óptima. Los métodos de alineación de secuencias para comparación son bien conocidos en la técnica. Una alineación óptima de secuencias para comparar puede llevarse a cabo, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman (1970) *Adv. Appl. Math.* 2:482c, mediante el algoritmo de alineación por homología de Needleman y Wunsch, *J. Mol. Biol.* 48:443, 1970, mediante la búsqueda con el método de similitud de Pearson y Lipman, *Proc. Natl. Acad. Sci.*

USA. 85:2444, 1988, mediante implementaciones informatizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante alineación manual e inspección visual (véase, por ejemplo, Brent et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (2003)).

5 Dos ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul et al., Nucl. Acids Res. 25:3389-3402, 1977; y Altschul et al., J. Mol. Biol. 215:403-410, 1990, respectivamente. El programa informático para realizar el análisis BLAST está a disposición del público a través del Centro Nacional de Información Biotecnológica. Este algoritmo implica primero identificar parejas de secuencias con alta puntuación (HSPs), identificando palabras cortas de longitud W en la secuencia de consulta, que coinciden o satisfacen alguna puntuación umbral T de valor positivo, cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. T se refiere al umbral de puntuación de la palabra vecina (Altschul et al., supra). Estos aciertos iniciales de palabras vecinas actúan como semillas para iniciar búsquedas para encontrar HSPs más largas que las contienen. Los aciertos de palabras se extienden en ambas direcciones a lo largo de cada secuencia tan lejos como la puntuación de alineación acumulativa se pueda incrementar. Las puntuaciones acumulativas se calculan usando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para una pareja de residuos que coinciden; siempre > 0) y N (puntuación de penalización para residuos que no coinciden; siempre < 0). Para secuencias de aminoácidos, se emplea una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabras en cada dirección se detiene cuando: la puntuación de alineación acumulativa disminuye en una cantidad X desde su valor máximo alcanzado; la puntuación acumulativa llega a cero o menos, debido a la acumulación de una o varias alineaciones de residuos con puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros del algoritmo BLAST W , T y X determinan la sensibilidad y la velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) emplea por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, $M = 5$, $N = -4$ y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP usa por defecto una longitud de palabra de 3, y una expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff, Proc. Natl. Acad. Sci. USA. 89:10915, 1989) alineaciones (B) de 50, una expectativa (E) de 10, $M = 5$, $N = -4$, y una comparación de ambas cadenas.

El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, Proc. Natl. Acad. Sci. USA. 90:5873-5787, 1993). Una medición de la similitud proporcionada por el algoritmo BLAST es la menor probabilidad de suma ($P(N)$), que proporciona una indicación de la probabilidad con la que una coincidencia entre dos secuencias de nucleótidos o aminoácidos se produciría por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia, si la menor probabilidad de suma en una comparación del ácido nucleico del ensayo con el ácido nucleico de referencia, es menor de aproximadamente 0,2, más preferiblemente menor de aproximadamente 0,01 y lo más preferiblemente menor de aproximadamente 0,001.

35 El porcentaje de identidad entre dos secuencias de aminoácidos también puede determinarse usando el algoritmo de E. Meyers y W. Miller (Comput. Appl. Biosci., 4:11-17, 1988) que se ha incorporado en el programa ALIGN (versión 2.0), usando una tabla de residuos por peso PAM120, una penalización por longitud de hueco de 12 y una penalización por hueco de 4. Además, el porcentaje de identidad entre dos secuencias de aminoácidos puede determinarse utilizando el algoritmo de Needleman y Wunsch (J. Mol. Biol. 48: 444-453, 1970) que se ha incorporado en el programa GAP en el paquete del programa informático GCG (disponible en www.gcg.com), utilizando ya sea una matriz Blossom 62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 o 4 y un peso de longitud de 1, 2, 3, 4, 5 o 6.

45 Aparte del porcentaje de identidad de secuencia indicado anteriormente, otra indicación de que dos secuencias de ácido nucleico o polipéptidos son sustancialmente idénticas es que el polipéptido codificado por el primer ácido nucleico reacciona inmunológicamente de forma cruzada con los anticuerpos generados contra el polipéptido codificado por el segundo ácido nucleico, como se describe a continuación. Por lo tanto, un polipéptido es típicamente sustancialmente idéntico a un segundo polipéptido, por ejemplo, cuando los dos péptidos difieren solo en sustituciones conservadoras. Otra indicación de que dos secuencias de ácido nucleico son sustancialmente idénticas, es que las dos moléculas o sus complementos, se hibridan entre sí en condiciones rigurosas, como se describe a continuación. Sin embargo, otra indicación de que dos secuencias de ácido nucleico son sustancialmente idénticas es que se pueden utilizar los mismos cebadores para amplificar la secuencia.

55 La expresión "anticuerpo recombinante", tal y como se usa en esta memoria, incluye todos los anticuerpos que se preparan, expresan, crean o aíslan por medios recombinantes, tales como anticuerpos aislados a partir de un animal (por ejemplo, un ratón) que es transgénico o transcromosómico para genes de inmunoglobulina de roedor o un hibridoma preparado a partir de los mismos, anticuerpos aislados a partir de una célula hospedadora transformada para expresar un anticuerpo humano y de roedor, por ejemplo, a partir de un transfectoma, anticuerpos aislados a partir de una genoteca combinatoria recombinante de anticuerpos de roedor, y anticuerpos preparados, expresados, creados o aislados por cualquier otro medio que implica corte y empalme de una porción o la totalidad de un gen, secuencias de inmunoglobulina de roedor, en otras secuencias de ADN. Tales anticuerpos recombinantes de roedor tienen regiones variables en las que las regiones estructurales y las CDRs se obtienen a partir de secuencias de inmunoglobulina de la línea germinal de roedor. En ciertas realizaciones, sin embargo, tales anticuerpos recombinantes de roedor pueden ser sometidos a mutagénesis *in vitro* (o, cuando se usa un animal transgénico para se-

cuencias de Ig de roedor, mutagénesis somática *in vivo*) y por lo tanto las secuencias de aminoácidos de las regiones VH y VL de los anticuerpos recombinantes son secuencias que, aunque se obtienen y están relacionadas con secuencias de la línea germinal de VH y VL de roedor, pueden no existir de forma natural dentro del repertorio de la línea germinal de anticuerpos de roedor *in vivo*.

5 La expresión "célula hospedadora recombinante" (o simplemente "célula hospedadora") se refiere a una célula en la que se ha introducido un vector de expresión recombinante. Se debe entender que tales expresiones son para referirse no solo a la célula objeto en particular, sino a la progenie de tal célula. Debido a que pueden producirse ciertas modificaciones en generaciones sucesivas debido a una mutación o a influencias ambientales, dicha progenie puede ser que de hecho no sea idéntica a la célula progenitora, pero que se siga incluyendo dentro del alcance de la expresión "célula hospedadora" tal y como se usa en el presente documento.

10 El término "vector" es para referirse a una molécula de polinucleótido capaz de transportar otro polinucleótido al que está ligado. Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN bicatenario circular al que se pueden ligar segmentos de ADN adicionales. Otro tipo de vector es un vector vírico, en el que se pueden ligar segmentos de ADN adicionales en el genoma vírico. Ciertos vectores son capaces de una replicación autónoma en una célula hospedadora en la que se introducen (por ejemplo, vectores bacterianos que tienen un origen de replicación bacteriano y vectores episomales de mamífero). Otros vectores (por ejemplo, vectores no episomales de mamífero) se pueden integrar en el genoma de una célula hospedadora después de la introducción en la célula hospedadora, y de ese modo se replican junto con el genoma del hospedador. Además, ciertos vectores son capaces de dirigir la expresión de genes a los que están ligados funcionalmente. Tales vectores se denominan en esta memoria "vectores de expresión recombinantes" (o simplemente, "vectores de expresión"). En general, los vectores de expresión de utilidad en las técnicas de ADN recombinante, están frecuentemente en forma de plásmidos. En la presente memoria descriptiva, "plásmido" y "vector" pueden usarse indistintamente ya que el plásmido es la forma de vector empleada más comúnmente. Sin embargo, la descripción se entiende que incluye otras formas de este tipo de vectores de expresión, tales como vectores víricos (por ejemplo, retrovirus, adenovirus y virus adenoasociados con replicación defectuosa), que sirven para funciones equivalentes.

Los sitios de restricción que son "únicos" son sitios de restricción que existen o aparecen solo una vez en una molécula de ácido nucleico dada. Típicamente, una molécula de ácido nucleico de este tipo es un vector que codifica los miembros de la genoteca de la presente descripción.

30 El término "mVL" es la abreviatura para el dominio variable de la cadena ligera de inmunoglobulina murina. El término "mVH" es la abreviatura para el dominio variable de la cadena pesada de inmunoglobulina murina. El término "rVL" es la abreviatura para el dominio variable de la cadena ligera de inmunoglobulina de rata. El término "rVH" es la abreviatura para el dominio variable de la cadena pesada de inmunoglobulina de rata. El término "IMGT" es la abreviatura para ImMunoGeneTics, una base de datos en línea, (Giudicelli, et al. (1997), *Nucleic Acids Res* 25, 206-211)).

35 El término "sujeto" incluye animales humanos y no humanos. Los animales no humanos incluyen todos los vertebrados, por ejemplo, mamíferos y no mamíferos, tales como primates no humanos, ovejas, perros, vacas, pollos, anfibios y reptiles. Excepto cuando se indique, los términos "paciente" o "sujeto" se utilizan en esta memoria de forma intercambiable.

40 El término "tratamiento" incluye la administración de composiciones o anticuerpos para prevenir o retrasar la aparición de síntomas, complicaciones o indicios bioquímicos de una enfermedad, aliviando los síntomas o deteniendo o inhibiendo un mayor desarrollo de la enfermedad, la afección o el trastorno. El tratamiento puede ser profiláctico (para prevenir o retrasar la aparición de la enfermedad, o para prevenir la manifestación de síntomas clínicos o subclínicos de la misma) o la supresión terapéutica o el alivio de síntomas después de la manifestación de la enfermedad.

45 La expresión "uso de aminoácidos dependiente de la posición" se refiere a la probabilidad de aparición de una secuencia de aminoácidos particular en una posición dada en un polipéptido. En la presente descripción, el uso de aminoácidos dependiente de la posición se determinó para las secuencias de aminoácidos reordenadas, clasificadas por el gen de la línea germinal individual. Esto permite el diseño preciso e individual de las CDRs en su contexto natural de la línea germinal.

50 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos de roedor que comprende miembros de al menos dos familias de la línea germinal. En otros aspectos dicha genoteca comprende miembros de al menos tres, al menos cuatro o al menos cinco familias de la línea germinal. En aspectos preferidos dicha genoteca sintética de anticuerpos de roedor es una genoteca de roedores murinos. En otros aspectos dicha genoteca sintética de anticuerpos de roedor es una genoteca de roedores de tipo rata.

55 En ciertos aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos que comprende más del 50% del repertorio natural de VH de ratón. En otros aspectos dicha genoteca comprende más del 60%, más del 70%, más del 75%, más del 80%, más del 85%, más del 90% o más de 95% del repertorio natural de VH de ratón.

- En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende miembros de al menos dos familias de la línea germinal de VH. En otros aspectos, dicha genoteca comprende miembros de al menos tres, al menos cuatro o al menos cinco familias de la línea germinal de VH. En ciertos aspectos dichas familias de la línea germinal de VH se seleccionan a partir de VH1, VH2, VH3, VH5 y VH9.
- 5 En ciertos aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos que comprende más del 50% del repertorio natural de VL de ratón. En otros aspectos dicha genoteca comprende más del 60%, más del 70%, más del 75%, más del 80%, más del 85%, más del 90% o más del 95% del repertorio natural de VL de ratón.
- 10 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende miembros de al menos dos familias de la línea germinal de VL. En otros aspectos, dicha genoteca comprende miembros de al menos tres familias de la línea germinal de VL. En ciertos aspectos, dichas familias de la línea germinal de VL son las familias de la línea germinal de VL-kappa seleccionadas a partir de VL1, VL3, VL7, VL8 y VL19.
- 15 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende al menos uno, al menos dos, al menos tres, al menos cuatro o al menos cinco de los siguientes genes de la línea germinal de VH: IGHV1-72*01 (VH1), IGHV2-2*01 (VH2), IGHV3-6*01 (VH3), IGHV5-9*04 (VH5) y IGHV9-3*01 (VH9). En otros aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos que comprende los siguientes genes de la línea germinal de VH: IGHV1-72*01 (VH1), IGHV2-2*01 (VH2) y IGHV5-9*04 (VH5).
- 20 En ciertas realizaciones, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende un gen variante de la línea germinal de VH1 IGHV1-72*01 que contiene una mutación K77R y/o T78V. En ciertos aspectos, dicho gen variante de la línea germinal de VH1, IGHV1-72*01, contiene la mutación K77R. En otros aspectos, dicha variante contiene una mutación T78V. En aspectos preferidos, dicha variante contiene una mutación K77R y T78V.
- 25 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco o al menos seis de los siguientes genes de la línea germinal de VL-kappa: IGKV1-117*01 (Vk1), IGKV3-12*01 (Vk3), IGHV3-4*01 (Vk3), IGHV7-33*01 (Vk7), IGHV8*21*01 (Vk8) y GHV19-93*01 (Vk19). En otros aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos que comprende al menos dos de los siguientes genes de la línea germinal de VL-kappa: IGKV1-117*01 (Vk1), IGKV3-12*01 (Vk3) y IGHV3-4*01 (Vk3).
- 30 En ciertos aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos que comprende los genes de la línea germinal de VH IGHV1-72*01 (VH1), IGHV2-2*01 (VH2) y IGHV5-9*04 (VH5) y al menos dos de los siguientes genes de la línea germinal de VL-kappa: IGKV1-117*01 (Vk1), IGKV3-12*01 (Vk3) y IGHV3-4*01 (Vk3).
- 35 En ciertos aspectos de la presente descripción, los genes de la línea germinal requieren una modificación con el fin de incluir sitios de restricción y para lograr una modularidad de la genoteca. Por lo tanto, en ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende uno o varios de los siguientes genes modificados de la línea germinal: IGHV1-72*01 con una mutación Q1E; IGHV1-72*01 con una mutación Q1E, una mutación K77R y una mutación A78V; IGHV2-2*01 con una mutación Q1E; IGHV3-6*01 con una mutación D1E y una mutación T26S; IGHV5-9*04 con una mutación K3Q; IGHV9-3*01 con una mutación Q1E, una mutación I2V y una mutación K53E, y IGKV1-117*01 con una mutación V2I.
- 40 En ciertas realizaciones, la presente invención proporciona una genoteca sintética de anticuerpos murinos que comprende H-CDR3 que incluye más del 50% de las longitudes de H-CDR3 de origen natural del repertorio de H-CDR3 murina. En otros aspectos, dicha genoteca comprende más del 60%, más del 70%, más del 75%, más del 80%, más del 85%, más del 90% o más del 95% de las longitudes de H-CDR3 de origen natural del repertorio de H-CDR3 murina.
- 45 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende H-CDR3 con una longitud de 7-12 aminoácidos. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende H-CDR3s con una longitud de 7 aminoácidos y/o 8 aminoácidos y/o 9 aminoácidos y/o 10 aminoácidos y/o 11 aminoácidos y/o 12 aminoácidos.
- 50 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende un diseño de H-CDR3 como se muestra en la Figura 5 o la Figura 6. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende un diseño de H-CDR3 como se muestra en la Figura 6.
- 55 En ciertos aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos en la que la región H-CDR3 tiene una diversidad de al menos $1,0 \cdot 10^8$. En otros aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos en la que la región H-CDR3 tiene una diversidad de al menos $1,0 \cdot 10^9$, de

al menos $1,0 \cdot 10^{10}$, de al menos $1,0 \cdot 10^{11}$, de al menos $1,0 \cdot 10^{12}$ o de al menos $1,0 \cdot 10^{13}$.

5 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende L-CDR3s que incluyen más del 80% de las longitudes de L-CDR3 de origen natural del repertorio de L-CDR3 murina. En otros aspectos, dicha genoteca comprende más del 85%, más del 90% o más de 95% de las longitudes de L-CDR3 de origen natural del repertorio de L-CDR3 murina.

En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende una L-CDR3 con una longitud de 8 aminoácidos. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que la L-CDR3 de esencialmente todos los miembros de la genoteca, tiene una longitud de 8 aminoácidos.

10 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende un diseño de L-CDR3 como se muestra en la Figura 7 o en la Figura 8. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende un diseño de L-CDR3 tal y como se muestra en la Figura 8.

15 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que la región L-CDR3 tiene una diversidad de al menos $1,0 \cdot 10^4$. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en donde la región L-CDR3 tiene una diversidad de al menos $1,0 \cdot 10^5$, de al menos $1,0 \cdot 10^6$ o de al menos $1,0 \cdot 10^7$.

20 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que al menos uno de los residuos de aminoácidos 52, 52a, 54, 56 y 58 (numeración de Kabat) de la H-CDR2, está diversificado. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que al menos dos, al menos tres, al menos cuatro o al menos cinco de los residuos de aminoácidos 52, 52a, 54, 56 y 58 (numeración de Kabat) de la H-CDR2 están diversificados.

25 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende un diseño de H-CDR2 tal y como se muestra en la Figura 10 o la Figura 11. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende un diseño de H-CDR2 tal y como se muestra en la Figura 11.

30 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que la región H-CDR2 tiene una diversidad de al menos $1,0 \cdot 10^5$. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en donde la región H-CDR2 tiene una diversidad de al menos $1,0 \cdot 10^6$, de al menos $1,0 \cdot 10^7$, de al menos $1,0 \cdot 10^8$, de al menos $1,0 \cdot 10^9$ o de al menos $1,0 \cdot 10^{10}$.

35 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos con un gen de la línea germinal de VH1, que comprende una diversidad de al menos $1,0 \cdot 10^5$ en la región H-CDR2. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos con una línea germinal de VH1 que comprende una diversidad de al menos $1,0 \cdot 10^6$, de al menos $1,0 \cdot 10^7$, de al menos $1,0 \cdot 10^8$ o de al menos $1,0 \cdot 10^9$ en la región H-CDR2.

40 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos con un gen de la línea germinal de VH2 que comprende una diversidad de al menos $1,0 \cdot 10^3$ en la región H-CDR2. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos con un gen de la línea germinal de VH2 que comprende una diversidad de al menos $1,0 \cdot 10^4$, de al menos $1,0 \cdot 10^5$, de al menos $1,0 \cdot 10^6$ o de al menos $1,0 \cdot 10^7$ en la región H-CDR2.

45 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos con un gen de la línea germinal de VH5 que comprende una diversidad de al menos $1,0 \cdot 10^5$ en la región H-CDR2. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos con una línea germinal de VH5 que comprende una diversidad de al menos $1,0 \cdot 10^6$, de al menos $1,0 \cdot 10^7$, de al menos $1,0 \cdot 10^8$ o de al menos $1,0 \cdot 10^9$ en la región H-CDR2.

50 En ciertos aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL se presentan de manera eficaz. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas o sustancialmente todas las combinaciones de VH-VL, con excepción de todas las combinaciones de VLB7, se presentan bien. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende al menos dos, al menos tres, al menos cuatro o al menos cinco genes de la línea germinal de VH y al menos dos, al menos tres, al menos cuatro o al menos cinco genes de la línea germinal de VL, en donde cada una de las combinaciones de VH-VL comprendidas en dicha genoteca se presenta de forma eficaz. La eficacia de la presentación se puede medir mediante ELISA de tipo sándwich en fagos, tal y como se describe en la presente descripción.

55 En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL se expresan bien en formato Fab en *E. coli*. En otros aspectos, la

presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas o sustancialmente todas las combinaciones de VH-VL, con excepción de todas las combinaciones de VLB7, se expresan bien en formato Fab en *E. coli*. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende al menos dos, al menos tres, al menos cuatro o al menos cinco genes de la línea germinal de VH y al menos dos, al menos tres, al menos cuatro o al menos cinco genes de la línea germinal de VL, en donde cada una de las combinaciones de VH-VL comprendidas en dicha genoteca se expresa bien en formato Fab en *E. coli*. La expresión en formato Fab en *E. coli* se puede cuantificar. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL se expresan con niveles superiores a 5 mg/L en un cultivo bacteriano. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL se expresan con niveles superiores a 10 mg/L en un cultivo bacteriano. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que la mayoría de las combinaciones de VH-VL se expresan con niveles superiores a 15 mg/L en un cultivo bacteriano. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que al menos dos, al menos tres, al menos cuatro o al menos cinco combinaciones de VH-VL se expresan con niveles superiores a 20 mg/L en un cultivo bacteriano.

En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL se expresan bien en un sistema de mamífero en formato IgG. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas o sustancialmente todas las combinaciones de VH-VL, con excepción de todas las combinaciones de VLB7, se expresan bien en un sistema de mamífero en formato IgG. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos que comprende al menos dos, al menos tres, al menos cuatro o al menos cinco genes de la línea germinal de VH y al menos dos, al menos tres, al menos cuatro o al menos cinco genes de la línea germinal de VL, en donde cada una de las combinaciones de VH-VL comprendida en dicha genoteca se expresa bien en un sistema de mamífero en formato IgG. La expresión en un sistema de mamífero en formato IgG se puede cuantificar. En ciertos aspectos, dicho sistema de mamífero es un cultivo en suspensión de células de mamífero. En otros aspectos, dicho sistema de mamífero es un cultivo de células de mamífero adherentes. En ciertos aspectos, dicho formato IgG es un formato IgG1. En otros aspectos, dicho formato IgG es un formato IgG2a. En ciertos aspectos, dicho sistema de mamífero comprende células HKB11. En otros aspectos, dicho sistema de mamífero comprende células PERC.6. En todavía otros aspectos, dicho sistema de mamífero comprende células CHO. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL se expresan con niveles superiores a 5 mg/L en un sistema de mamífero en formato IgG. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL se expresan con niveles superiores a 10 mg/L en un sistema de mamífero en formato IgG. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que la mayoría de las combinaciones de VH-VL se expresan con niveles superiores a 15 mg/L en un sistema de mamífero en formato IgG. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que al menos dos, al menos tres, al menos cuatro o al menos cinco combinaciones de VH-VL se expresan con niveles superiores a 20 mg/L en un sistema de mamífero en formato IgG.

En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todos o sustancialmente todos los miembros de dicha genoteca son estables en isopropanol a una concentración del 30% (v/v). Es una suposición que todos los derivados de la CDR de una región estructural estable se comportarán como los genes maestros sometidos a ensayo.

En ciertos aspectos, la presente invención proporciona una genoteca sintética de anticuerpos murinos en la que todas o sustancialmente todas las combinaciones de VH-VL son térmicamente estables. La estabilidad térmica se puede medir tal y como se describe en la presente solicitud. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL tienen una T_m de más de 62°C. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL tienen una T_m de más de 64°C. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que esencialmente todas las combinaciones de VH-VL tienen una T_m de más de 66°C. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que la mayoría de las combinaciones de VH-VL tienen una T_m de más de 68°C. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que muchas combinaciones de VH-VL tienen una T_m de más de 70°C. En ciertos aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas o sustancialmente todas las combinaciones de VH1 tienen una T_m de más de 70°C. En otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas o sustancialmente todas las combinaciones de VH1_RV tienen una T_m de más de 70°C. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas las combinaciones de VLB1 tienen una T_m de más de 68°C. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas las combinaciones de VLB3 tienen una T_m de más de 68°C. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas las combinaciones de

VLB7 tienen una T_m de más de 68°C. En aún otros aspectos, la presente descripción proporciona una genoteca sintética de anticuerpos murinos en la que todas las combinaciones de VLC3 tienen una T_m de más de 68°C.

5 En ciertos aspectos, la presente descripción proporciona anticuerpos de roedor que poseen propiedades biofísicas favorables. Tales anticuerpos están desprovistos de las deficiencias de los anticuerpos obtenidos a partir de genotecas sintéticas que no han sido sometidas a maduración *in vivo*. Tales propiedades biofísicas favorables y deseadas incluyen una mayor estabilidad, niveles de expresión más altos y una baja tendencia a la agregación.

10 En ciertos aspectos, la presente descripción proporciona un anticuerpo aislado a partir de una genoteca de roedor contemplada por la presente descripción. En ciertos aspectos, dicho anticuerpo puede ser un anticuerpo modificado o una variante de un anticuerpo aislado a partir de una genoteca de roedor contemplada por la presente descripción. En otros aspectos dicho anticuerpo modificado o variante puede ser un anticuerpo humanizado.

15 Cada composición y cada aspecto concomitante, y cada método y cada aspecto concomitante que se han descrito anteriormente, se pueden combinar entre sí de una manera coherente con las enseñanzas contenidas en este documento. De acuerdo con las realizaciones de las invenciones, todos los métodos y las etapas de cada método se pueden aplicar en cualquier orden y repetirse muchas veces de una manera coherente con las enseñanzas contenidas en este documento.

20 La invención se describe adicionalmente mediante los siguientes ejemplos, que no limitan de ningún modo la invención. Por ejemplo, ciertas características del diseño que se describen a continuación, pueden ser exclusivas del diseño de "HuCAL MOUSE" o de "HuCAL RAT" y no se requieren en una genoteca de roedor alternativa contemplada en la presente descripción.

Ejemplos

Ejemplo 1: Generación de una genoteca de anticuerpos murinos (HuCAL MOUSE)

Ejemplo 1.1: Selección de secuencias de la línea germinal de VL y VH para la generación de genes maestros murinos

25 Se analizaron los datos públicos disponibles para el uso de familias de la línea germinal de VH murinas (véase, Kaushik et al., Res. Immunol. (1996) 147, 9-26). La siguiente tabla resume la aparición de las diferentes líneas germinales de VH:

Tabla 1:

VH	% en BD de Kabat (cepas var.)	Nombre (Kaushik et al. (1996))	% en Kaushik et al. (BALB/c)	% en Kaushik et al. (cepas var.)
VH1	33	J558	30	45
VH2	24	Q52	14	10
VH3	24	30-60	8	17
VH5	16	7183	11	11
VH9	10	VGAM	9	10
SUMA	87		72	93

30 Nuestro objetivo era generar una genoteca que comprendiera el mayor número posible de familias de la línea germinal. Una genoteca que contiene un miembro de cada una de las familias de la línea germinal mencionadas como genes maestros, incluiría más del 85% del repertorio de VH de ratón, y precisamente este era el objetivo de nuestra tarea.

35 Las familias de la línea germinal comprenden frecuentemente más de un gen de la línea germinal. Para las familias mencionadas, el número de genes de la línea germinal oscila desde 7 (VH2) a 134 (VH1) - el número exacto depende de la cepa de ratón. Como la mayoría de los estudios en animales se centran en cepas de ratones BALB/c y C57BL-6, solamente los genes de la línea germinal favorecidos por estas cepas se consideraron como base para los genes maestros de la genoteca murina. Véase la Tabla 2. Se encontró y se utilizó exclusivamente una intersección razonable de genes de la línea germinal de VH de ambas cepas de ratón.

Tabla 2:

Nº de genes de la línea germinal en la familia		
VH	(cepas var.)	(BALB/c & C57BL-6)

	N° de genes de la línea germinal en la familia	
	VH (cepas var.)	(BALB/c & C57BL-6)
VH1	134	35
VH2	7	2
VH3	5	4
VH5	20	9
VH9	8	7

5 Entre las cadenas ligeras murinas, las cadenas ligeras kappa constituyen la mayoría de todas las cadenas ligeras, lo que representa aproximadamente el 90% de las inmunoglobulinas séricas aisladas a partir de ratones (Frank y Gutman, 1988, Mol Immunol 25, 953-960). Por tanto, solo VL-kappa fue considerada en la genoteca murina. Dado que los ratones BALB/c y C57BL-6 no comparten el mismo uso de la línea germinal de VL, el desarrollo de genes maestros VL para BALB/c y C57BL-6 se realizó por separado. La Tabla 3 muestra las familias de la línea germinal que se producen con más frecuencia.

Tabla 3:

BALB/c		C57BL-6	
Familia de VL-kappa	Uso en %	Familia de VL-kappa	Uso en %
1	40	3	49
3	33	8	19
7	12	19	17
SUMA	85	SUMA	85

10 El objetivo era preparar una genoteca que contuviera un miembro de cada una de las familias de la línea germinal de la lista. Tal genoteca incluirá >85% de las secuencias de VL de ratones BALB/c y C57BL-6. Aunque la familia de la línea germinal kappa-3 es utilizada por ambas cepas, se requieren dos genes maestros distintos porque se prefieren genes diferentes de la línea germinal, con un total de 5 diferencias de nucleótidos. Por lo tanto, en total se requirieron 6 genes maestros de VL. Se precisan determinadas alteraciones de estos genes de la línea germinal para formar genes maestros que conserven la modularidad de la genoteca.

15 Ejemplo 1.2: Selección de genes maestros

Se realizaron dos metodologías para la selección de los genes de la línea germinal utilizados con más frecuencia, que son los que se van a utilizar como genes maestros:

- 20 a) Identificación del gen de la línea germinal más destacado dentro de las respectivas familias de la línea germinal
- b) Identificación del gen de la línea germinal que es más similar al consenso de todos los genes de la línea germinal dentro de una familia.

Idealmente, ambas metodologías darían como resultado el mismo gen de la línea germinal.

25 Ejemplo 1.2.1: Identificación del gen de la línea germinal más destacado dentro de las respectivas familias de la línea germinal

El gen de la línea germinal empleado de forma más abundante se identificó mediante una comparación de los genes de la línea germinal con secuencias reordenadas. Aproximadamente 9400 secuencias de VH murina reordenadas y 2500 secuencias de Vk se extrajeron de una base de datos de archivo plano que se descargó desde IgBLAST en el servidor de NCBI (url: <http://www.ncbi.nlm.nih.gov/igblast/>). Se excluyeron los fragmentos de secuencias y las secuencias de mala calidad. Las secuencias restantes se clasificaron en familias de la línea germinal mediante la aplicación de filtros para un patrón específico y único de aminoácidos de la línea germinal. Esta versión sin pulir de la clasificación de las familias, se corrigió mediante la puntuación de los anticuerpos reordenados frente a genes de la línea germinal. Aquí, el uso de aminoácidos dependiente de la posición de cada secuencia reordenada, se comparó con el uso de aminoácidos de cada gen de la línea germinal de varias cepas. Al hacerlo, cada secuencia reordenada fue asignada a un gen de la línea germinal más próximo. Si se identificaron secuencias reordenadas mal asignadas, se clasificaron en la familia de la línea germinal correcta, lo que llevó a la distribución mostrada en la Tabla 4 (VH) y

5 (VL).

Tabla 4:

VH:

Familia de la línea germinal	Nº de secuencias reordenadas	% de distribución
VH1	5980	70,9
VH2	597	7,1
VH3	246	2,9
VH5	762	9,0
VH9	117	1,4
otros	730	8,7
SUMA	8432	100

Tabla 5:

VL:

Familia de la línea germinal	Nº de secuencias reordenadas	% de distribución
Vk1	543	21,6
Vk3	232	9,2
Vk7	14	0,6
Vk8	187	7,4
Vk19	44	1,8
otros	1494	59,4
SUMA	2514	100

5

El uso de aminoácidos dependiente de la posición, de la secuencia reordenada se comparó con el uso de aminoácidos de los genes de la línea germinal específica de BALB/c y C57BL-6, para identificar los genes de la línea germinal usados más frecuentemente. Cada secuencia reordenada fue asignada a un gen de la línea germinal de BALB/c/C57BL6 más cercana. Si se identificaron dos o más genes de la línea germinal como el gen de la línea germinal más próximo para una secuencia reordenada, todos los genes de la línea germinal fueron considerados como genes de la línea germinal más próximos. Para el cálculo del gen de la línea germinal utilizado preferentemente dentro de una familia, se sumó el uso de cada gen de la línea germinal. Los resultados para los genes de la línea germinal preferidos dentro de una familia eran sorprendentemente inequívocos y claros. La Tabla 6 muestra un ejemplo de los resultados para el uso de los cuatro genes de la línea germinal en la familia de VH3:

10

15

Tabla 6:

Gen de la línea germinal	Uso absoluto	Uso relativo [%]
D13203 IGHV3-3*03 BALB/c	0	0,0
AC073589 IGHV3-3*01 C57BL/6	5	1,9
AJ223544 IGHV3-6*03 BALB/c	14	5,4
AC073590 IGHV3-6*01 C57BL/6	242	92,7
SUMA	261	100
nº de secuencias reordenadas	246	

Para el ejemplo mostrado, la suma del uso absoluto de genes de la línea germinal (261) difiere del número de secuencias reordenadas (246), ya que 15 de las secuencias reordenadas habían sido asignadas a dos o más genes de la línea germinal.

20

La Tabla 7 muestra los genes de la línea germinal que son los más destacados en las diferentes familias de VH. La Tabla 8 muestra los genes de la línea germinal más destacados para las diferentes familias de VL. Se identificaron diferentes genes de la línea germinal para Vk3 de ratones BALB/c y C56BL-6 como los más destacados.

Tabla 7:

Familia de la línea germinal	Nombre del gen destacado de la línea germinal	Nº de secuencias reordenadas	% de genes destacados de la línea germinal en secuencias reordenadas
VH1	AC163348 IGHV1-72*01	5980	49,7
VH2	AC090887 IGHV2-2*01	597	87,0
VH3	AC073590 IGHV3-6*01	246	92,7
VH5	AF120472 IGHV5-9*04	762	49,8
VH9	AC073563_ IGHV9-3*01	117	63,1

Tabla 8:

Familia de la línea germinal	Cepa	Nombre del gen destacado de la línea germinal	Nº de secuencias reordenadas	% de genes destacados de la línea germinal en secuencias reordenadas
Vk1	BALB/c	IGKV1-117*01	543	90%
Vk3	BALB/c	IGKV3-12*01	232	24%
Vk3	C57BL-6	IGHV3-4*01	232	75%
Vk7*)	BALB/c	IGHV7-33*01	246	100%
Vk8	C57BL-6	IGHV8*21*01	187	37%
Vk19*)	C57BL-6	IGHV19-93*01	44	100%
SUMA			1484	

*) La familia de la línea germinal tiene solo un gen maestro

5 Ejemplo 1.2.2: Identificación del gen de la línea germinal que es más parecido al consenso de todos los genes de la línea germinal dentro de una familia

10 Se prepararon secuencias de consenso de los genes de la línea germinal de la intersección de la línea germinal BALB/c-C57BL-6. La comparación de las secuencias de consenso con los genes de la línea germinal se realizó de forma análoga a la descrita anteriormente en esta memoria: el uso de aminoácidos dependiente de la posición de las secuencias de consenso se comparó con el uso de aminoácidos de los genes de la línea germinal específicos de BALB/c y C57BL-6. Al hacerlo, las secuencias de consenso se asignaron a un gen de la línea germinal más próximo de BALB/c y C57BL-6. El resultado se muestra en la Tabla 9.

Tabla 9:

Familia de la línea germinal	Nombre del gen destacado de la línea germinal	Familia de la línea germinal	cepa	Nombre del gen destacado de la línea germinal
VH1	AC163348 IGHV1-72*01	Vk1	BALB/c	IGKV1-117*01
VH2	AC090887 IGHV2-2*01	Vk3	BALB/c	IGKV3-12*01
VH3	AC073590 IGHV3-6*01	Vk3	C57BL-6	IGHV3-4*01
VH5	AF120472 IGHV5-9*04	Vk7	BALB/c	IGHV7-33*01
VH9	AC073563_ IGHV9-3*01	Vk8	C57BL-6	IGHV8*28*01
		Vk19	C57BL-6	IGHV19-93*01

15 La validez de esta metodología fue confirmada por análisis BLAST de las secuencias de consenso frente a todos los genes de la línea germinal. A excepción de Vk8, la identificación del gen de la línea germinal más destacado dentro de las respectivas familias de la línea germinal, y el gen de la línea germinal con la mayor similitud con el consenso o la búsqueda BLAST, surgió del mismo conjunto de genes de la línea germinal. Para Vk8, se seleccionó IGKV8-28*01 como gen maestro, ya que muestra menos desviaciones con la secuencia de consenso que IGKV8-21*01. El uso de IGKV8-28*01 para secuencias reordenadas en la familia de la línea germinal de Vk8 es 11%.

20

Para mVH1, se incluyó una segunda variante con las mutaciones K77R y T78V, porque la formación de un puente salino estabilizador entre R77 y D100 podría añadir una estabilidad adicional elevada a esta variante de VH1 (Kaufmann et al., J. Mol. Biol., (2002) 318, 135-47). Véase la Figura 1.

5 Los genes maestros seleccionados finalmente (véase más abajo) requerían ligeras modificaciones con el fin de lograr una modularidad de la genoteca. Las modificaciones realizadas para generar una restricción se resumen en las Tablas 10 (VH) y 11 (VL).

Tabla 10:

GM	Nombre del gen más próximo de la línea germinal	Desviaciones con el gen más próximo de la línea germinal
mVH1	AC163348 IGHV1-72*01	Q1E
mVH1_RV	AC163348 IGHV1-72*01	Q1E,K77R,A78V (puente salino estabilizante)
mVH2	AC090887 IGHV2-2*01	Q1E
mVH3	AC073590 IGHV3-6*01	D1E, T26S (sitio BspEI)
mVH5	AF120472 IGHV5-9*04	K3Q (sitio Mfel)
mVH9	AC073563_ IGHV9-3*01	Q1E, I2V, K53E (sitio XhoI)

Tabla 11:

GM	Cepa	Nombre del gen más próximo de la línea germinal	Desviaciones con el gen más próximo de la línea germinal
Vk1 = B1	BALB/c	IGKV1-117*01	V2I (sitio EcoRV)
Vk3 = B3	BALB/c	IGKV3-12*01	-
Vk7 = B7	BALB/c	IGHV7-33*01	-
Vk3 = C3	C57BL-6	IGHV3-4*01	-
Vk8 = C8	C57BL-6	IGHV8*28*01	-
Vk19 = C19	C57BL-6	IGHV19-93*01	-

10 Además, para crear genes maestros de longitud completa a partir de los genes de la línea germinal, se tuvo que añadir a las secuencias la región estructural 4 del anticuerpo. IMGT (<http://irnqt.cines.fr/>) proporciona un acceso a las regiones J de murino, que forman la región estructural 4 del anticuerpo en el desarrollo del anticuerpo. La Figura 2 muestra las regiones estructurales para las cadenas pesadas, la Figura 3 para las cadenas ligeras.

15 La región de J JH2 fue seleccionada para el gen maestro de VH, ya que el extremo 3' asociado de HCDR3 (FDY) se prefiere en HCDR3 y JH2 proporciona todos los sitios de ER requeridos (StyI y BlnI). Las regiones de J JK1 (= JK2 en la secuencia de FW4) fueron seleccionadas para la complementación del gen maestro de VL, que incluía el 85% del uso de JK de ratón. El último aminoácido de la región estructural 4 se cambió de Ala a Thr para proporcionar un sitio para BsiWI.

20 Ejemplo 1.3: Diseño de CDRs murinas

Los sitios de contacto más estrecho entre el anticuerpo y el antígeno son las regiones determinantes de complementariedad (CDR) del anticuerpo. H-CDR3 y L-CDR3 tienen un papel importante en la unión al antígeno, por tanto, la variabilidad se introdujo principalmente en estas dos CDRs. Una variabilidad limitada adicional se introdujo en posiciones de contacto con el antígeno de la H-CDR2. Se emplearon secuencias de la línea germinal para todas las demás CDRs. Para mejorar las afinidades de los ligantes murinos seleccionados para obtener afinidades comparables a las del candidato líder humano, se sintetizaron casetes de maduración adicionales para H-CDR2.

Los residuos de cisteína se evitaron generalmente en el diseño de las CDRs, porque los residuos de cisteína pueden estar involucrados en la formación de enlaces disulfuro que unen covalentemente un ligante con la diana de la inmunoadsorción o en la formación de homodímeros de anticuerpo. Tales ligantes es poco probable que sean dianas selectivas.

Las secuencias de anticuerpos reordenadas se compilaron y se analizaron para el diseño de CDRs murinas (véase el Ejemplo 1.1). Las Tablas 4 (VH) y 5 (VL) muestran los conjuntos de datos utilizados.

Ejemplo 1.3.1: Diseño de H-CDR3

El análisis estadístico de la H-CDR3 se hizo para todas las secuencias reordenadas con independencia de la familia de la línea germinal, ya que esta CDR se codifica de forma independiente de la línea germinal por el segmento D y J.

5 Aunque tirosina, glicina y serina predominan tanto en secuencias de H-CDR3 humanas como de ratón, la utilización de aminoácidos difiere entre las secuencias de la misma longitud entre las cadenas pesadas variables de ratón y humanas. Estas diferencias en las frecuencias totales de aminoácidos y en la distribución local de residuos de aminoácidos dentro de los bucles de H-CDR3, contribuyen a enormes diferencias en la diversidad y las estructuras previstas de los repertorios de H-CDR3 de ratón y humana. (Zemlin et al., 2003, J Mol Biol 334, 733-749). Basándose en el análisis de secuencias de ratón reordenadas, se identificaron algunas posiciones, que se desviaban extremadamente de la composición del diseño de H-CDR3 humana. Especialmente las posiciones de Kabat 95, 96 y 102 mostraron otro uso de aminoácidos que la H-CDR3 humana (véase la Tabla 12). Por lo tanto, se requieren H-CDR3s murinas diseñadas especialmente.

Tabla 12:

Ratón			Ser humano		
95	96	102	95	96	102
17% Y	22% Y	77% Y	20% D	10% G	45% Y
16% G	17% G	14% V	20% G	10% R	15% I
12% S	10% D		8% V	10% S	15% V
10% D	7% R		8% E	10% L	10% P
10% R	6% L		8% A	10%P	
5% E	6% S		8% S		
	5% N				

Ilustración reducida a aminoácidos con una aparición de $\geq 5\%$.

15 En comparación con los anticuerpos humanos, el intervalo de longitud de H-CDR3 encontrado en los anticuerpos murinos está restringido. Aunque la H-CDR3 humana incluye longitudes de 2 a 28 aminoácidos, la H-CDR3 murina utiliza solo un conjunto de 3 a 16 aminoácidos, tal y como se observó en el análisis de los datos de Ig-BLAST. Para los resultados véase la Tabla 13. Resultados similares fueron descritos por Zemlin et al. (Zemlin et al., 2003, J Mol Biol 334, 733-749). La tecnología TRIM[®] (véase el documento WO 93/21203) se utilizó para diseñar H-CDR3s de ratón que incluían una longitud de 7 - 12 aminoácidos. Este intervalo abarca aproximadamente el 80% de todas las H-CDR3 murinas.

Tabla 13:

Longitud de HCDR3 (numeración de HuCAL)	Nº de secuencias	%
1_aa	0	0
2_aa	1	0
3_aa	72	1
4_aa	113	1
5_aa	326	4
6_aa	341	4
7_aa	446	6
8_aa	645	8
9_aa	1705	21
10_aa	1311	17
11_aa	1054	13
12_aa	1189	15
13_aa	355	4
14_aa	187	2
15_aa	120	2

Longitud de HCDR3 (numeración de HuCAL)	Nº de secuencias	%
16_aa	46	1
17_aa	19	0
18_aa	9	0
19_aa	2	0
20_aa	0	0
21_aa	0	0
Nº total de secuencias	7941	100

5 La distribución de los aminoácidos en relación con la dependencia de la posición se determinó para cada longitud individual, lo que dio como resultado la distribución natural de aminoácidos. Las distribuciones de aminoácidos individuales se alinearon mediante la agrupación para el segmento D, dando lugar a una alineación centrada derecha. Los resultados se muestran en la Figura 4. Para resumir todas las longitudes en un diseño, se calculó el promedio ponderado de los aminoácidos para cada posición. Véase la Figura 5.

Esta distribución calculada de los aminoácidos se modificó ligeramente para finalizar el diseño de la H-CDR3 en la genoteca:

- * Unificación de la distribución de aminoácidos para limitar el conjunto de diferentes mezclas TRIM
- 10 * Reconocimiento de los aminoácidos con una aparición de $\geq 10\%$
- * Complementación de otros aminoácidos (excepto cisteína) mediante una mezcla de 19 aminoácidos, que se añadió a cada posición
- * Evitar completamente la cisteína
- * Examen y reducción de ciertos sitios de restricción para conseguir una modularidad
- 15 * Menor uso de triptófano, que se va a utilizar solo en la mezcla de 19aa
- * Evitar aminoácidos similares en una posición (por ejemplo, Glu o Asp)
- * Evitar sitios de glicosilación (NxS, NxT)

Después de considerar e implementar las modificaciones requeridas, se utilizó el diseño de H-CDR3 mostrado en la Figura 6.

20 La diversidad de la H-CDR3 con una longitud de 12 aminoácidos es el producto de las diversidades en las posiciones individuales y es de aproximadamente $2,5 \cdot 10^{13}$. La diversidad total es la suma de las diversidades de las distintas longitudes, 7-12 aminoácidos, y es de aproximadamente $2,6 \cdot 10^{13}$.

Ejemplo 1.3.2: Diseño de L-CDR3

25 El análisis estadístico de la L-CDR3 se realizó para todas las secuencias reordenadas con independencia de la familia de la línea germinal, ya que esta CDR está codificada parcialmente de forma independiente de la línea germinal por el segmento J. Basándose en el análisis de secuencias reordenadas de VL, se encontró, que una longitud de L-CDR3 de 8 aminoácidos tiene lugar en aproximadamente el 85% de las secuencias (véase la Tabla 14). Por lo tanto la longitud de la L-CDR3 en la genoteca se fijó en una longitud de 8 aminoácidos.

Tabla 14:

Longitud	Nº de secuencias	%
7aa	173	9,6
8aa	1571	87,6
9aa	50	2,8
SUMA	1794	100,0

30 Se calculó la distribución de los aminoácidos dependientes de la posición de las 1571 secuencias reordenadas de V-kappa con una longitud de 8 aminoácidos. Los resultados se muestran en la Figura 7.

El motivo de la variabilidad en la posición de la región estructural 100, es probablemente el uso de diferentes segmentos JK. Esta distribución de aminoácidos calculada se modificó ligeramente para finalizar el diseño de la L-CDR3 en la genoteca:

- * Unificación de la distribución de aminoácidos para limitar el conjunto de diferentes mezclas TRIM
- 5 * Reconocimiento de aminoácidos con una aparición de $\geq 10\%$
- * Complementación de otros aminoácidos (excepto cisteína) mediante una mezcla de 19 aminoácidos, que se añadió a cada posición
- * Evitar completamente la cisteína
- * Examen y reducción de ciertos sitios de restricción para conseguir una modularidad
- 10 * Menor uso de triptófano, que se va a utilizar solo en la mezcla de 19aa (especialmente las posiciones 91 y 96)
- * Evitar aminoácidos similares en una posición (por ejemplo, Glu o Asp)
- * Evitar sitios de glicosilación (NxS, NxT, presentes en $\sim 5\%$ en una distribución nat.)
- * Fijar la posición de 90 al 100% para Q y la posición de 95 al 100% para P
- 15 Después de considerar e implementar las modificaciones requeridas, se utilizó el diseño de L-CDR3 mostrado en la Figura 8.

La diversidad de la L-CDR3 con una longitud de 8 aminoácidos es el producto de las diversidades en las posiciones individuales y es de aproximadamente $4,7 \cdot 10^7$ que es igual a la diversidad total.

Ejemplo 1.3.3: Diseño de H-CDR2

- 20 Se ha descrito que las posiciones 52, 52a, 54, 56, 58 (numeración de Kabat) de H-CDR2 son las más responsables del contacto con el antígeno en la región H-CDR2. Véase (<http://www.bioc.uzh.ch/antibody/Structures/AgContact/index.html>). Por lo tanto, una variación de estas posiciones debe permitir la selección de anticuerpos con el sitio de unión al antígeno adaptado que comprende H-CDR2.

- 25 Se investigó la distribución de los aminoácidos para distintas familias de secuencias reordenadas de la línea germinal de anticuerpos, representados por los tres genes maestros seleccionados de mVH (mVH1_RV, mVH2 y mVH5). Se recopiló una lista de los residuos que se encontraron con frecuencia en las posiciones particulares de H-CDR2. Con el fin de conservar la prevalencia de la secuencia de la línea germinal, los residuos alternativos presentes en los anticuerpos reordenados pero no en las secuencias de la línea germinal, se infrarrepresentaron intencionadamente en el diseño. La variación de las posiciones se introdujo en los genes maestros seleccionados mediante oscilación.
- 30 Véase la Figura 9.

Los casetes de maduración de H-CDR2 se diseñaron para mejorar las afinidades de ligantes murinos seleccionados.

- 35 La distribución de aminoácidos dependiente de la posición de la H-CDR2 se investigó para distintas familias de secuencias de la línea germinal de anticuerpos reordenados representados por los tres genes maestros seleccionados de mVH (mVH1_RV, mVH2 y mVH5). Basándose en el análisis de las secuencias de ratón reordenadas, se dedujo la distribución de aminoácidos mostrada en la Figura 10 mediante el cálculo de la distribución de aminoácidos dependiente de la posición.

Esta distribución de aminoácidos calculada se modificó ligeramente para finalizar el diseño de la H-CDR2 en la genoteca:

- * Unificación de la distribución de aminoácidos para limitar el conjunto de diferentes mezclas TRIM
- 40 * Reconocimiento de aminoácidos con una aparición de $\geq 10\%$
- * Complementación de otros aminoácidos (excepto cisteína) mediante una mezcla de 19 aminoácidos, que se añadió a algunas posiciones
- * Evitar completamente la cisteína
- * Evitar sitios de glicosilación (NxS, NxT)
- 45 * Examen y reducción de ciertos sitios de restricción para conseguir una modularidad
- * Aumento de la variabilidad del casete VH2-HCDR2 mediante la adición de una mezcla de 19 aminoácidos

en la posición 53 y 54

* Fijación de todas las posiciones desde la posición 62 en adelante en un aminoácido para proporcionar una región constante para el cebador hacia atrás de la PCR

5 Después de considerar e implementar las modificaciones requeridas, se utilizó el diseño de H-CDR2 mostrado en la Figura 11.

La diversidad de los casetes de H-CDR2 es el producto de las diversidades en las posiciones individuales y es de aproximadamente $1,1 \cdot 10^9$ para mVH1-H-CDR2, $9,9 \cdot 10^6$ para mVH2-H-CDR2 y $3,6 \cdot 10^9$ para mVH5-H-CDR2.

Ejemplo 1.4: Clonación, expresión y purificación de la genoteca murina

Cepas bacterianas y vectores utilizados para la clonación

10 La clonación molecular se llevó a cabo utilizando las cepas de *E. coli* XL1-Blue (Stratagene) o Top10F+ (Invitrogen). Los genes maestros fueron sintetizados por GeneArt AG (Regensburg, Alemania). Los productos de la síntesis génica se clonaron en un vector pGA, y luego se volvieron a clonar en el vector tricistrónico CysDisplay pMorph30, como se describe a continuación. El vector pMorph30 presentado en fagos se había obtenido a partir del plásmido pMorph23 (Rothe et al., 2008, J Mol Biol 376, 1182-1200), básicamente de la siguiente manera: el marcador hexa-
15 His fijado al extremo C-terminal de VH fue sustituido por la secuencia de aminoácidos HKHKHK, y un sitio chi en gill fue eliminado. Las endonucleasas de restricción eran de New England Biolabs o Fermentas, y la ADN ligasa de T4 era de Invitrogen. Las condiciones de la reacción se llevaron a cabo de acuerdo con las recomendaciones del fabricante.

Números de orden de los genes maestros originales

20 Números de orden de la sub-base de datos de *Mus musculus* de IMGT/NCBI de las secuencias de la línea germinal que se utilizaron para la construcción de los genes maestros:

mVH1: IGHV1-72*01 / AC163348, mVH2: IGHV2-2*01 / AC090887, mVH3: IGHV3-6*01 / AC073590, mVH5: IGHV5-9*04 / AF120472, mVH9: IGHV9-3*01 / AC073563, mVLB1: IGKV1-117*01 / D00081, mVLB3: IGKV3-12*01 / K02159, mVLB7: IGKV7-33*01 / AF044198, mVLC3: IGKV3-4*01 / Y15968, mVLC8: IGKV8-28*01 / AJ235947,
25 mVLC19: IGKV19-93*01 / AJ235935, Jk2: IGKJ2*01 / V00777, Ck:IGKC*01 / V00807, JH2: IGHJ2*01 / V00770, CH1 + bisagra: IGHG1*02 / L35252

Consideraciones para la síntesis de genes maestros

Los genes correspondientes a las secuencias representativas de la línea germinal seleccionadas, se optimizaron a nivel de ADN mediante la aplicación de los criterios siguientes: (i) el uso de codones raros tiene que ser óptimo para
30 *E. coli*, mientras que se evitaban codones raros murinos/humanos siempre que fue posible; (ii) las características indeseables de ARN/ADN como cajas TATA internas, sitios chi, sitios de entrada ribosómica, segmentos de secuencias ricas en AT o GC, secuencias que impiden el transporte de ARNm en el citoplasma (elementos INS o CRS), secuencias repetidas, estructuras secundarias de ARN y sitios donantes y aceptores de corte y empalme crípticos, se evitaban; (iii) se evitó la introducción de sitios de endonucleasas de restricción (ER) que estaban en conflicto con
35 la modularidad de la clonación con HuCAL[®]; y (iv) se introdujeron sitios de endonucleasas de restricción necesarios para la clonación modular. La glicosilación en los dominios variables no es deseable principalmente porque se ha observado que la fijación de hidratos de carbono afecta frecuentemente a la unión del antígeno debido a un impedimento estérico, y también es responsable de la heterogeneidad de proteínas expresadas en hospedadores eucariotas. Por lo tanto, comprobamos que los genes maestros de VL/VH diseñados no codificaban proteínas con sitios
40 potenciales para la glicosilación ligada a N.

Dominios constantes en formato Fab

El dominio constante de la cadena ligera kappa está codificado por el gen Ck (IGKC*01 / V00807, véase más arriba). El residuo de alanina N-terminal fue reemplazado por treonina lo que permitió la introducción de un sitio BsiWI en
45 esta posición a nivel de ADN. Un residuo de cisteína localizado en el extremo C-terminal de Ck de tipo silvestre, fue reemplazado por alanina.

El dominio constante de la cadena pesada en formato Fab está codificado por el fragmento génico CH1. Hemos elegido el gen IGHG1*02, porque representa el consenso de secuencias CH1 del tipo IgG1 murino, depositado en la base de datos IMGT. El único desemparejamiento era un residuo Glu en un fragmento de la secuencia de la línea
50 germinal TWPSE y, por tanto, el residuo Glu se cambió a un Gln. La adición de un residuo de alanina delante de la secuencia N-terminal KTTPPS, permitió la introducción de un sitio BlnI. Se añadieron los primeros cuatro residuos de la bisagra (secuencia VPRD) seguidos por los residuos Glu-Phe (necesarios para la introducción de un sitio EcoRI) en el extremo C-terminal del dominio CH1.

Ensamblamiento génico en el plásmido pMorph31 tricistrónico mostrado en fagos

Los genes diseñados que codificaban dominios constantes o variables murinos fueron generados mediante ensamblamiento con PCR de una serie de oligonucleótidos, y se clonaron en un vector de clonación pGA en GeneArt AG. Era cómodo y rápido reemplazar los genes de VL, VH, C κ y CH1 humanos por equivalentes murinos en dos moléculas de vectores intermedios diferentes. Dos derivados del plásmido pMorph30 (véase más arriba) denominados "plásmido L" y "plásmido H", se construyeron con fragmentos grandes de ADN simulado con el fin de facilitar la clonación de la siguiente manera: en el "plásmido L", se insertó ADN simulado (aproximadamente 3 kb) en lugar de la región que codificaba la cadena ligera, mientras que en el "plásmido H", el ADN simulado (aproximadamente 2 kb) se insertó en lugar de la región que codificaba la cadena pesada. El ADN simulado en el "plásmido L" fue reemplazado después por mC κ y mVL en dos etapas: el gen mC κ (véase más arriba) se insertó utilizando BsiWI/SphI, y después los genes maestros de mVL kappa (véase más arriba) se insertaron utilizando EcoRV/BsiWI. El ADN simulado en el "plásmido H" fue reemplazado por mCH1 y mVH también en dos etapas: el gen mCH1 (véase más arriba) se insertó utilizando Styl (o Eco130I)/EcoRI, y luego se insertaron los genes maestros de mVH (véase más arriba) mediante el uso de MfeI/Styl. Los fragmentos de ADN que codificaban los genes maestros mVH1, mVH1RV, mVH2, mVH5 o mVH9 (todos ellos ligados a la región mCH1) se combinaron (clonación cruzada) mediante el uso de SphI/HindIII con los genes maestros mVLB1, mVLB3, mVLB7, mVLC3, mVLC8 y mVLC19 en el plásmido pMorph30/mC κ dando como resultado 36 combinaciones de mVL-mVH. Después de la clonación cruzada, el plásmido de presentación se denominó pMorph31, indicando el hecho que el plásmido además de gpIII también codifica dominios variables y constantes murinos y, por lo tanto, es adecuado para la presentación de Fabs murinos.

Reclonación de combinaciones de genes maestros de VL-VH en el vector bicistrónico pMorphx9_Fab_FH para la expresión en E. coli

Con el fin de analizar la expresión soluble de las combinaciones de la región estructural VL-VH en formato Fab, reclonamos las 36 combinaciones de genes maestros de mVL-mVH (como insertos que codifican Fab) utilizando XbaI y EcoRI en el vector pMorphx9_Fab_FH (Rauchenberger et al., 2003, J. Biol. Chem. 278, 38194-38205), en donde FH indica que los marcadores Flag y hexa-His se fijaron en el extremo C-terminal de los Fabs.

Conversión de combinaciones de genes maestros de VL-VH en formato IgG2a

Con el fin de analizar la expresión de las combinaciones de genes maestros en formato kappa IgG2a murino de longitud completa en células de mamífero, subclonamos las 36 combinaciones de genes maestros de mVL-mVH en el vector pMorph4_h/m_IgG que es un derivado del vector pMorph2_h_IgG (Steidl et al. 2008. Mol Immunol 46(1):135-44). En el vector pMorph4_h/m_IgG, las cadenas ligeras y pesadas se codificaban en un plásmido; sin embargo, cada cadena se traducía junto con su secuencia líder N-terminal a partir de un transcrito distinto (la transcripción estaba dirigida por dos sitios de promotor CMV localizados en el plásmido).

Conversión de combinaciones de genes maestros de VL-VH en formato IgG1 murino

Con el fin de analizar la expresión de las combinaciones de genes maestros en formato kappa IgG1 murino de longitud completa en células de mamífero, subclonamos las 36 combinaciones de genes maestros de mVL-mVH en el vector pMorph4_h/m_IgG, que es un derivado del vector pMorph2_h_IgG (Steidl et al. 2008. Mol Immunol 46(1):135-44). En el vector pMorph4_h/m_IgG, las cadenas ligeras y pesadas se codificaban en un plásmido; sin embargo, cada cadena se traducía junto con su secuencia líder N-terminal a partir de un transcrito distinto (la transcripción estaba dirigida por dos sitios de promotor CMV localizados en el plásmido).

Expresión y purificación de Fab

La expresión en células TG1 de *E. coli* a partir del plásmido pMx9_Fab_FH se llevó a cabo en 500 ml de medio 2xYT complementado con cloranfenicol (34 μ g/ml) y glucosa (0,1%). Después de la inducción con IPTG (0,75 mM), las células se cultivaron a 30°C durante 20 h. Los sedimentos celulares se lisaron durante la incubación con tampón de lisis (Tris/HCl 25 mM, pH 7,4, NaCl 0,5 M, imidazol 10 mM, 0,2% de lisozima, MgCl $_2$ 2 mM, 20 U/ml de benzonasa y cóctel inhibidor de proteasa exento de EDTA de Roche) durante 30 min a TA. La purificación de Fab a partir de lisados celulares aclarados (centrifugación y filtración, 0,2 μ m) se llevó a cabo mediante una purificación en 2 etapas, utilizando Äkta Express para el procesamiento automatizado. La primera etapa era la cromatografía IMAC con iones Ni $^{2+}$, en donde en caso de una evaluación de la estabilidad, se llevó a cabo una etapa de lavado adicional con 30% de isopropanol en PBS. Después de la elución con una concentración elevada de imidazol, el pico eluido se volvió a cargar sobre una columna de filtración en gel, y las proteínas se separaron mediante exclusión por tamaño utilizando PBS pH 7,4. Finalmente, se agruparon las fracciones de elución que correspondían al tamaño esperado.

Expresión y purificación de IgG

Células eucariotas HKB11 (Cho et al.; J Biomed Science (2002) 9, 631-8) se transfectaron de forma transitoria con el plásmido pMorph4_h/m_IgG que codificaba las cadenas pesadas y ligeras. El material sobrenadante del cultivo celular se recogió 3 a 7 días después de la transfección. Después de ajustar el pH del sobrenadante a 8,0, de añadir NaCl (concentración final 2 M) y filtrar de forma estéril (0,45 μ m), la solución se sometió a una cromatografía de afinidad con proteína A estándar (MabSelect SURE, GE Healthcare). Si no se indica lo contrario, el intercambio de tampón se realizó con 1 x PBS de Dulbecco (pH 7,2, Desnitrogen) y las muestras se filtraron de forma estéril (0,2 μ m). La pureza de la IgG se analizó en condiciones de desnaturalización mediante SDS-PAGE o mediante el uso de

Agilent BioAnalyzer y en estado natural mediante HP-SEC.

Preparación de fagos

La amplificación y la purificación de los fagos se llevó a cabo esencialmente tal y como se ha descrito anteriormente (Krebs et al., 2001, J. Immunol Methods 254, 67-84) con la siguiente modificación. Los fagémidos (combinaciones de la región estructural VL-VH codificadas en el plásmido de presentación) se propagaron en células Top10F+ de *E. coli* en medio 2xYT/Cam/Tet/glucosa (1%). Después de la infección con fago auxiliar (Hyperphage de Progen, nº de Cat.: PRHYPE, multiplicidad de la infección: 40), la centrifugación y la resuspensión del sedimento celular en 2xYT/Cam/Kan/Tet/IPTG (0,25 mM), los fagos se produjeron mediante incubación durante 20 h a 22°C.

ELISA de fagos

Las placas de microtitulación negras de Maxisorp (Nunc) se recubrieron con anticuerpo anti-pIII (MoBiTec, nº de Cat.: PSKAN3) para la determinación del nivel de pIII relativo (\approx concentración de fagos), o con anticuerpo F(ab')₂ anti-ratón (Jackson ImmunoResearch Lab.; nº de Cat.: 112-006-072) para la determinación del nivel relativo de Fab en las muestras de fago. Aunque el epítipo exacto del anticuerpo F(ab')₂ anti-ratón no se conoce (se supone que se encuentra en el dominio constante), pensamos que la unión del anticuerpo a Fab murino no es probable que esté muy afectada por una variación de secuencia en el dominio Fv y, por lo tanto, el anticuerpo debía ser adecuado para la evaluación general de la presentación de Fab murino en la punta del fago filamentoso. TBS pH 7,4 que contenía Tween 20 (0,05%) y chemi-bloquer (de Chemicon; diluido 1:2) se utilizó para el bloqueo de los sitios de unión no específicos en la placa de microtitulación y en los fagos. Después de incubar con los fagos, los pocillos se lavaron con TBS-T, y se incubaron secuencialmente con anticuerpo anti-bacteriófago biotinilado (Sigma nº de Cat. B2661) y avidina-HRP (BD Pharmingen; nº de Cat.: 554058). Se aplicó el sustrato "Quanta Blu" para HRP, y se registraron las señales de fluorescencia en un aparato Tecan-Spectrafluor (excitación: 320 nm, emisión: 430 nm). Los niveles relativos de pIII y de Fab en la muestra de fago se determinaron mediante el uso de un fago de referencia (que presenta un ligante seleccionado previamente convertido en Fab con dominios constantes murinos) y el modelo de 2-estados estándar aplicado para la regresión no lineal. La tasa de presentación relativa se definió como la relación entre los niveles de Fab y de pIII.

Ejemplo 1.5: Caracterización de la genoteca de anticuerpos murinos por medio del anticuerpo 4D5 ejemplar

Una buena presentación en fagos y una expresión elevada en las células se puede lograr solo con moléculas Fab estables y plegadas de forma adecuada. La estabilidad termodinámica de una molécula Fab depende de la estabilidad intrínseca de los dominios individuales (cadenas pesada y ligera), así como de la interacción no covalente entre los dos dominios (Ewert et al., 2003, J Mol Biol 325, 531-553). Una interacción estrecha entre Fd (cadena pesada) y Ck (cadena ligera) es de particular importancia porque en el sistema CysDisplay[®] los dos dominios no pueden estar unidos covalentemente a través de un enlace disulfuro, ya que el residuo de cisteína en los extremos C-terminales de los dominios constantes se ha eliminado. Por lo tanto, en lugar de someter a ensayo individualmente las regiones estructurales de VL y VH, analizamos combinaciones de VL-VH que se podían plegar de manera eficaz y expresar en *E. coli* (formato Fab) y en células de mamífero (formato IgG).

Como caso tipo utilizamos la secuencia HCDR3 del anticuerpo 4D5 (WGGDGFYAMDY). Esta secuencia ya se había utilizado previamente para determinar genes maestros humanos antes de la construcción de la genoteca HuCAL original (Knappik et al., 2000, J Mol Biol 296, 57-86)

Análisis de la eficacia de la presentación

Una presentación eficaz de los miembros de la genoteca sobre partículas de fago es un requisito previo para el éxito de las selecciones de una presentación en fagos. Utilizamos ELISA de tipo sándwich en fagos para cuantificar la presentación de las 36 combinaciones de VL-VH murinas en formato Fab. Nuestros datos indican que la mayoría de las combinaciones de mVL-mVH sometidas a ensayo, con excepción de todas las combinaciones de mVLB7, se presentaron en los fagos. No había ninguna diferencia significativa en la presentación entre las moléculas Fab que contenían mVH1 o mVH1_RV. Aunque estadísticamente es insignificante, hay una tendencia que indica que las combinaciones de VH5 se presentan ligeramente mejor que todos los demás Fabs murinos. Los resultados se representan en la Figura 12.

Expresión de Fab soluble

A continuación, investigamos si las características favorables del plegamiento y la expresión de Fab murino sometido a ensayo en un hospedador bacteriano, también se pueden observar en formato soluble. En este formato, tanto las cadenas pesadas como las ligeras se dirigen al periplasma de *E. coli*, en donde alcanzan su plegamiento final y se asocian entre sí para formar un fragmento Fab estable.

Como se muestra en la Figura 13, todas las combinaciones de VL-VH murinas sometidas a ensayo, con excepción de todas las combinaciones de mVLB7, se expresaron bien en *E. coli* y el rendimiento de Fab purificado superaba los 5 mg/L de cultivo bacteriano. Las combinaciones de VH5 (en particular, VLC3, VLC8-, VLC19-VH5) parecían estar un poco mejor expresadas que todos los demás Fabs murinos. No había una diferencia significativa en la ex-

presión entre las moléculas Fab que contenían mVH1 o mVH1_RV. La expresión de Fab en formas solubles y de presentación se correlaciona muy bien (véase la Figura 14). Por otra parte, la caracterización del material aislado reveló que el 98-100% del Fab purificado estaba en forma monomérica. Esto indica la utilidad para el desarrollo de anticuerpos murinos sustitutos.

5 *Expresión de IgG soluble*

La gran mayoría de los productos biológicos terapéuticos basados en anticuerpos, actualmente en el mercado, están en formato IgG por una serie de razones: (i) la semivida de las moléculas de IgG en el cuerpo humano es muy alta (alrededor de 3 semanas) debido a la interacción de la IgG con el receptor neonatal (FcRn); (ii) las moléculas de IgG son altamente solubles, termodinámicamente estables y relativamente resistentes a las proteasas en la sangre; y (iii) las IgGs poseen actividad ADCC (citotoxicidad mediada por células dependiente de anticuerpos) y/o CDC (citotoxicidad dependiente del complemento), que se requieren para la eliminación de células tumorales.

Puesto que la expresión de una combinación de VL-VH particular en formato Fab no se correlaciona necesariamente con la expresión de la misma combinación de VL-VH en formato IgG, optamos por caracterizar la expresión de todas las combinaciones compartidas de genes maestros sintetizados de VL y VH también en formato IgG.

15 Dos formatos IgG distintos de ratón, IgG1 e IgG2a, fueron elegidos para la prueba de expresión. El isotipo IgG2a murino se encuentra en aproximadamente el 27% de todas las IgGs monoclonales aisladas a partir de ratón Balb/c (Rousseaux y Bazin, 1979, Vet Immunol. Immunopathol. 1, 61-78), y sometimos a ensayo todas las combinaciones de genes maestros de VL-VH con este formato. Además, unas pocas combinaciones de VL-VH seleccionadas también se sometieron a ensayo en formato IgG1 murino, el cual se encuentra en aproximadamente el 40% de todas las IgGs monoclonales aisladas a partir de ratón Balb/c (Rousseaux y Bazin, 1979, Vet Immunol Immunopathol. 1, 61-78).

20 Todos los genes maestros VL y VH murinos se volvieron a clonar en vectores de expresión de mamífero, y los plásmidos resultantes se usaron para la transfección transitoria de una línea celular de mamífero en suspensión, tal y como se ha descrito en esta memoria anteriormente. Los datos se muestran en las Figuras 15 y 16, lo que indica que el nivel de expresión de casi todas las combinaciones de mVL-mVH sometidas a ensayo, tanto en formato tanto IgG1 como IgG2a murino, exceden en gran medida un valor de 10 mg/L. De manera similar a las observaciones hechas para el formato Fab, no había diferencias en la expresión entre las moléculas de IgG que contenían mVH1 o mVH1_RV. Un nivel muy bajo de expresión de las combinaciones de mVLB7 es paralelo a la expresión procarionota muy baja de las combinaciones de mVLB7 observada en el formato Fab. Resumiendo, contra todo pronóstico, hemos sido capaces de expresar de forma soluble casi todas las combinaciones de mVL-mVH sometidas a ensayo.

25 Todas las combinaciones de genes maestros de VL y VH murinos se purificaron en un proceso de purificación de 2 etapas, usando Äkta express. Los fragmentos Fab expresados están unidos a Ni²⁺ quelado y se lavaron con 15 volúmenes de columna de tampón fosfato de Na 20 mM/NaCl 500 mM, pH 7,4, imidazol 20 mM o, en una purificación distinta, utilizando 20 volúmenes de columna de isopropanol al 30%/fosfato de Na 20 mM pH 7,4. La segunda etapa de la purificación (SEC) era idéntica en ambos procedimientos (con o sin tratamiento con isopropanol) y se ha descrito anteriormente en el presente documento. Después de la purificación, se compararon ambas preparaciones de proteína en relación con el rendimiento, la concentración, la pureza, la proporción de VH y VL y la estabilidad térmica. Se introdujo una etapa de lavado con isopropanol al 30%, a la vez que el Fab todavía estaba unido a la columna de IMAC, con el fin de recibir fragmentos Fab que tenían mayor pureza. Sin embargo, para algunos fragmentos Fab, este tratamiento parecía ser demasiado fuerte porque la interfaz - la interacción hidrofóbica entre la cadena pesada y ligera - no era lo suficientemente fuerte como para soportar este procedimiento. Las cadenas pesada y ligera de tales fragmentos Fab se eluyeron en diferentes etapas y, por lo tanto, la relación entre las dos cadenas se desvió de 1:1.

Estabilidad térmica

35 El nivel de expresión de las proteínas en un hospedador celular depende de muchos parámetros. Algunos de ellos están relacionados con el plegamiento intrínseco de las proteínas; sin embargo, muchos de ellos dependen del estado metabólico del hospedador, tal como la actividad de la maquinaria para la traducción y la translocación, o los niveles de proteasas secretadas y de ubiquitina. Estas últimas propiedades extrínsecas de las proteínas son difíciles de controlar y varían frecuentemente de forma sustancial en experimentos realizados en días diferentes. Debido a las variaciones observadas en las pruebas de expresión descritas anteriormente, caracterizamos las propiedades intrínsecas de las proteínas mediante el uso de muestras proteicas purificadas durante los ensayos de expresión. Prácticamente, la medición de la estabilidad térmica se puede realizar con un pequeño volumen en un termociclador estándar para PCR en tiempo real y, por lo tanto, es adecuada para la comparación relativa de la estabilidad intrínseca entre un grupo numeroso de proteínas sometidas a ensayo. Por otro lado, la desnaturalización térmica de proteínas, al contrario que la desnaturalización química mediante agentes caotrópicos, es típicamente irreversible, y no se puede utilizar directamente para la determinación de parámetros termodinámicos absolutos, tales como ΔG de la transición proteica entre sus dos estados conformacionales.

Los resultados se muestran en la Figura 18. Las combinaciones de VL-VH murinas más termoestables eran las

combinaciones VLB3-VH1_RV y VLB3-VH1_RV. Importante, incluso la estabilidad térmica de las combinaciones de mVLB7 sometidas a ensayo era razonablemente alta. Esto fue inesperado, teniendo en cuenta los niveles de expresión comparablemente bajos, detectados para todas las combinaciones de mVLB7 en los formatos Fab e IgG.

Selección de regiones estructurales de Fv para la construcción de una genoteca

5 Durante la determinación de las 36 regiones estructurales diferentes (combinaciones de 6 genes maestros de VL con 6 de VH) descrita anteriormente, observamos que básicamente todas las regiones estructurales Fv se expresaban en *E. coli* en formato Fab y en formato IgG en células de mamífero, con rendimientos apreciables que excedían generalmente los 10 mg/L. Las únicas excepciones eran todas las combinaciones de mVLB7 y varias de las combinaciones de mVLC19 sometidas a ensayo (véase más arriba). En nuestra selección de las regiones estructurales Fv más favorables, pusimos de relieve la estabilidad termodinámica de las moléculas Fv. En este sentido, mVLB3-mVH1_RV y mVLC3-mVH1_RV con un punto medio de desnaturalización térmica (Tm) a aproximadamente 76°C, eran una clara elección (véase la Figura 18). Además, las secuencias de la línea germinal mVH1, mVLB3 y mVLC3 se encuentran frecuentemente en la base de datos IMGT (ver Tabla 15). Dos otras regiones estructurales Fv seleccionadas comprenden mVLB3 y mVLC3 en combinación con mVH5. mVH5 es muy utilizada en las cadenas pesadas murinas (Tabla 15), las combinaciones de mVH5 estaban bien expresadas en formato Fab en *E. coli* (véanse las Figuras 12 y 13), y las secuencias de mVLB3-mVH5 y mVLC3-mVH5 se encontraron en la base de datos Kabat de anticuerpos murinos reordenados (Tabla 16).

20 Otra región estructural Fv favorable que comprendía distintas VH era mVLB1-mVH2, debido a la abundancia relativamente alta de esta combinación de regiones estructurales Fv en la base de datos Kabat (Tabla 16), y debido al uso elevado de mVLB1 en cadenas ligeras murinas (Tabla 15). Por lo tanto, casetes diversificados de la genoteca Lcdr3 y Hcdr3 se subclonaron en las regiones estructurales Fv seleccionadas, dando como resultado cinco subgenotecas murinas (véase más abajo).

La Tabla 15 indica el uso de familias de la línea germinal de VK y VH para las cepas de ratón BALB/c y C57BL-6.

Tabla 15:

BALB/c		C57BL-6	
VK_B1	40%	VK_C3	49%
VK_B3	33%	VK_C8	19%
otros	27%	VK_C19	17%
		otros	5%
BALB/c & C57BL-6			
VH1	27%		
VH2	10%		
VH5	25%		
VH8	12%		
Otros	26%		

25

La Tabla 16 muestra el resultado del análisis estadístico de la representación de las secuencias de VL-VH combinadas en la base de datos Kabat.

Tabla 16:

Pareja VL-VH	Número de aciertos en Kabat
Vk B3_VH5	3
Vk B1_VH2.	4
VkB3_VH9	0
VkB3_VH3	1
Vk B3_VH1 RVT	0
Vk C3_VH9	0
Vk C3_VH5	1

Pareja VL-VH	Número de aciertos en Kabat
Vk C8_VH1	0
Vk C19_VH1 RVT	0
Vk C3_VH1 RVT	0

Análisis estadístico de la representación de secuencias de VL-VH combinadas en la base de datos de Kabat

Para determinar si las combinaciones de VH-VL que daban como resultado anticuerpos que se expresaban bien en la genoteca de la presente descripción, también existen en anticuerpos seleccionados de forma natural, se analizó la representación de esos anticuerpos en la base de datos Kabat. Las secuencias combinadas de la región estructural VH-VL de los anticuerpos se habían anulado en todas las regiones estructurales de las secuencias de anticuerpos completos de Kabat que comprendían una secuencia para VH y VL.

En la bases de datos de Kabat, están presentes 1501 entradas de anticuerpos de ratón completos, con información de la secuencia completa (cadena pesada y ligera). Para incluir todas las posibles combinaciones de cadenas pesadas y ligeras de la cepa de ratón BALB/c, 101 (cadenas pesadas) x 23 (cadenas ligeras) = se necesitarían 2323 anticuerpos completos en el conjunto de la muestra. Para C57BL6 el número sería aún mayor: 195 (cadenas pesadas) x 74 (cadenas ligeras) = 14430 anticuerpos completos. Teniendo en cuenta el pequeño conjunto de la muestra y el hecho de que en la base de datos Kabat estas dos cepas de ratón no son las únicas que están representadas, es más bien poco probable que una de las combinaciones de cadenas VH-VL de la genoteca de ratón de la presente descripción exista en esta base de datos. Por tanto, cualquier número de coincidencias mayor que uno, se tomó como una indicación de que esa combinación podía ser en realidad una forma natural preferida. Las combinaciones de cadenas mejor representadas fueron elegidas para el diseño final de la genoteca.

Ejemplo 1.6: Diversificación de la genoteca murina

Introducción de ADN de relleno (del inglés "stuffer") para la inserción de casetes diversificados de CDR3 en el plásmido de presentación pMorph31

Las regiones de ADN que codifican L-CDR3 o H-CDR3 son cortas (<40 pb). En nuestro sistema modular, ambas CDR3s originales codificadas en el gen maestro y el casete de CDR3 diversificado, están flanqueadas por una pareja de sitios de restricción únicos, que permiten la introducción de un casete de CDR3 diversificado en lugar de la CDR3 original. Una digestión ineficaz del plásmido en estos sitios de restricción puede dar lugar a un ruido de fondo sustancial que comprende el vector con CDR3 original porque no es técnicamente posible separar eficazmente, mediante electroforesis en gel, los plásmidos linealizados (-5500 pb) en los que está presente o ausente la región que codifica CDR. Con el fin de evitar el ruido de fondo del vector, sin tener en cuenta la eficacia de la digestión, decidimos introducir un relleno de ADN (fragmentos de ADN simulado de -500 pb de tamaño) en lugar de la región que codificaba CDR3. Los rellenos de ADN se introducen en el plásmido pMorph31 en los sitios de ER que flanqueaban CDR3, dando como resultado los plásmidos pMorph31_mVLj/MBP-mVHi/AP. En particular, el relleno AP obtenido a partir del gen de la fosfatasa alcalina, se insertó en lugar de H-CDR3 mediante el uso de BssHII/Styl, y el relleno MBP obtenido a partir del gen de la proteína que se une a maltosa, se insertó en lugar de L-CDR3 mediante el uso de Bpil/BsiWI. Hay que mencionar que debido al modo de digestión de Bpil (ER de tipo II), tuvimos que preparar dos variantes del relleno MBP (una para mVLB1 y otra para mVLB3/mVLC3) que diferían en la proximidad del sitio Bpil dependiente de la secuencia del gen maestro.

Construcción de un casete de mini-genoteca utilizando cebadores con posiciones oscilantes

La variabilidad prevista (cinco variantes de secuencias de aminoácidos por gen maestro) se introdujo en los genes maestros seleccionados en 10 ciclos de PCR con ADN polimerasa Platinum Pfx (Invitrogen), y un conjunto de cebadores directos e inversos. Los cebadores directos que contenían posiciones oscilantes con una proporción definida de nucleótidos, fueron utilizados de la siguiente manera:

GGTCGCGGCCTCGAGTGGATTGGCCGCATT**G**(70%)/**A**(15%)/**T**(15%)ATCCGAACAG(80%)/**A**(20%)CGGCGGCACCAAATATAACGAA para mVH1_RV,

GGCAAAGGTCTCGAGTGGCTGGGCGTGATTTG**G**A(80%)/**G**(20%)GCG**G**(70%)/**A**(30%)CGGC**A**(70%)/**G**(30%)GCACCGATTATAACGCCGCC para mVH2, y

GAAAAACGCCTCGAGTGGGTGGCGACCATTAG(70%)/**A**(30%)CGGTGG**C**A(70%)/**G**(30%)GTG**G**CA**A**(60%)/**G**(40%)CACCTATTATCCAGATAGC para mVH5.

Las regiones H-CDR2 se completaron mediante regiones flanqueantes constantes, que representaban una parte de la región estructural 2 y de la región estructural 3 del anticuerpo. Por lo tanto, era posible amplificar con PCR y clonar los casetes en una etapa posterior. Dado que las regiones flanqueantes de los genes maestros de VH son diferentes, había que sintetizar oligos distintos para los diferentes genes maestros:

El análisis de la secuencia de aproximadamente 30 clones recogidos al azar por cada región estructural, indicaba que las variantes deseadas (cinco por región estructural) estaban representadas en la HCDR2 con una frecuencia similar a nuestro diseño. Solo aproximadamente el 10% de los clones recogidos al azar tenía deleciones de nucleótidos no deseadas, principalmente en las posiciones que oscilaron en los cebadores de PCR. Con el fin de asegurar poca abundancia de clones con desplazamiento de marco en la genoteca, decidimos diluir la agrupación de plásmidos que codificaban variaciones en HCDR2, con plásmido que codificaba la región estructural VH de tipo silvestre. Por lo tanto, en la genoteca final las variantes menores en HCDR2 son aproximadamente 10 veces menos abundantes que en nuestro diseño inicial de HCDR2.

Después de este control de calidad, los productos de la PCR se digirieron con XhoI/BssHII, y se clonaron en los plásmidos respectivos pMorph31_mVLj/MBP-mVHi/AP (véanse los párrafos sobre la introducción de los rellenos de ADN más arriba). La agrupación de plásmidos que contenía la variedad introducida en los genes maestros, se mezcló 1:10 con el plásmido que codificaba el gen maestro original y se utilizó para la introducción del casete en la genoteca de LCDR3.

Síntesis de oligonucleótidos que contienen trinucleótidos y preparación del casete de la genoteca de CDR3

Los oligonucleótidos que contenían trinucleótidos se sintetizaron como se ha descrito previamente (Knappik et al., 2000). Un oligonucleótido que contenía trinucleótidos diversificados ("TRIM") se mezcló con un oligonucleótido "solapante", correspondiente a la hebra complementaria que se solapaba con el oligonucleótido que contenía trinucleótidos, en posiciones no diversificadas. Ambos oligonucleótidos (a concentraciones de 0,1-0,2 µM) se ensamblaron y se amplificaron en 10 ciclos de PCR usando cebadores finales (a una concentración de 1 µM) y ADN polimerasa Platinum Pfx (Invitrogen). Los cebadores finales directos e inversos se reasociaron con los extremos 5' y 3' del casete de CDR, respectivamente, y facilitaron la introducción de los sitios de restricción necesarios para la clonación del casete. Las secuencias de todos los oligonucleótidos utilizados para la preparación de casetes de CDR3 se enumeran en las Tablas 17 (L-CDR3) y 28 (H-CDR3). Los casetes de CDR3 amplificados se purificaron en geles de agarosa y se aislaron tal y como se describe en otro lugar (Rothe et al., 2008, J Mol Biol 376, 1182-1200). Cabe destacar, que se tuvieron que sintetizar cuatro casetes de LCDR3 (uno específico de mVkB3 y mVcC3, otros de mVkB1, VkB7 y VcC8) para cada región estructural VL seleccionada que difería en la proximidad del sitio Bpil (ER de tipo II), lo que era necesario para la subclonación de LCDR3.

Tabla 17:

Cebador directo específico de LCDR3 ratón y rata		
	Nombre	Secuencia
mVkB3	Lib63	GTGGAAGAAGAAGACGCCGCCACCTATTATTGC
mVkB1	Lib64	GTGGAAGCCGAAGACCTGGGCGTGATTATTGC
mVkB7	Lib65	GTGCAGGTGGAAGACCTGACCCATTATTACTGC
mVcC8	Lib66	GTGCAGGCCGAAGACCTGGCCGTGTATTACTGC
Oligos solapantes		
ratón	Lib71	CGGCCACGTACGTTTCAGTTCCAGTTTGGTGCCCGCACCAAAGG
Cebador Vk inv		
ratón	Lib69	CGGCCACGTACGTTTCAGTTC

Tabla 18:

Cebador directo específico de HCDR3 ratón y rata		
	nombre	secuencia
ratón/rata	BHU402	CGCGTGGAAGACTGCGCGC
Oligos solapantes		
ratón/rata	BHU401	CATAACGAAGACGCCTTGCCCCA
Cebador HCDR3 inv		
ratón/rata	BHU403	CATAACGAAGACGCCTTG

Subclonación de casetes de la genoteca CDR3 en genes maestros en formato fagémido

Los casetes amplificados de la genoteca LCDR3 se insertaron en lugar del relleno de ADN MBP en los plásmidos pMorph31_mVLj-mVH//AP con HCDR2 diversificada mediante el uso de Bpil/BsiWI. Utilizamos normalmente ~2 µg de vector para ligar el exceso de 2 molar del casete LCDR3. La mezcla vector-inserto se preincubó durante 10 minutos a 56°C con una adición previa de ADN ligasa de T4 y tampón, y después de la incubación durante la noche a 16°C, la mezcla de ligación se incubó 10 minutos a 65°C con el fin de inactivar la ligasa. A continuación, el ADN ligado se purificó mediante precipitación con 2-butanol/glucógeno y se electroporó en células Top10F' competentes (Invitrogen). Después de la transformación y el crecimiento durante 1 hora a 37°C en medio SOB, se determinó el número de células transformadas por valoración en placas de agar con LB/cam/glucosa (1%). La amplificación de los clones transformados se realizó en medio líquido (LB/cam/glucosa (1%)) durante la noche a 22°C y luego a 30°C hasta que se alcanzó una DO_{600nm} de 2,0. Una parte alícuota del cultivo líquido se utilizó para la reserva congelada con glicerol, y el resto se utilizó para un maxiprep de ADN usando el kit de Qiagen.

El casete L-CDR3 muy variable fue completado mediante regiones flanqueantes constantes, lo que representaba una parte de la región estructural 3 del anticuerpo y la región estructural 4 terminal del anticuerpo VL. Por lo tanto, era posible amplificar con PCR y clonar el casete en una etapa posterior. Ya que las regiones flanqueantes de los genes maestros VL eran diferentes, se tuvieron que sintetizar oligos distintos para los diferentes genes maestros:

mVkB3 = mVkB3

5' GAAGACGCCGCCACCTATTATTGC-T8-CAG-T9-T10-T11-T12-CCG-T13-ACCTTTGGC
GGTGGCACCA

mVkB1

5' GAAGACCTGGGCGTGTATTATTGC-T8-CAG-T9-T10-T11-T12-CCG-T13-ACCTTTGGC
GGTGGCACCA

mVkB7

5' GAAGACCTGACCCATTACTACTGC-T8-CAG-T9-T10-T11-T12-CCG-T13-ACCTTTGGC
GGTGGCACCA

mVkB8

5' GAAGACCTGGCCGTGTACTACTGC-T8-CAG-T9-T10-T11-T12-CCG-T13-ACCTTTGGC
GGTGGCACCA

en donde T8 - T13 son mezclas TRIM.

Para la construcción de H-CDR3, se generó un casete de CDR distinto para cada longitud de H-CDR3. La composición de los casetes distintos se muestra en la Figura 19.

El casete H-CDR3 muy variable fue completado mediante regiones flanqueantes constantes, lo que representaba una parte de la región estructural 3 del anticuerpo y la región estructural 4 terminal del anticuerpo VH. Por lo tanto, era posible amplificar con PCR y clonar los casetes en una etapa posterior. Ya que las regiones flanqueantes de los genes maestros VH eran diferentes, un oligo se pudo utilizar para todos los genes maestros:

5'GTGGAAGACTGCGCGCGG-T1-T1-(T2)-(T3)-(T3)-(T3)-T4-T5-T6-W-
T7-TGGGGCCAAGGCGTCT,

en donde T1 - T7 son mezclas TRIM y W son nucleótidos oscilantes como se han descrito en esta memoria anteriormente.

Los casetes de la genoteca de HCDR3 amplificados y diversificados con diferentes longitudes (7, 8, 9, 10, 11 y 12 residuos) se mezclaron entre sí en relaciones molares que reflejaban la distribución según la longitud de HCDR3 planificada, y se subclonaron en los plásmidos pMorph31_mVLj-mVH//AP con HCDR2 y LCDR3 diversificadas mediante el uso de BssHII/Styl. Típicamente, 15-20 µg de vector se ligaron con un exceso 2 molar de la mezcla del casete HCDR3. Las condiciones de ligación, transformación y propagación de los clones, empleadas para la subclonación de la mezcla de casetes HCDR3, no variaban de las condiciones utilizadas para la subclonación de los casetes LCDR3. El análisis de la secuencia de VL y VH de los -70 clones y el análisis de restricción de 24 clones recogidos al azar y del maxiprep del ADN agrupado de cada subgenoteca, se realizaron con el fin de evaluar la calidad de la genoteca.

La diversidad teórica de $2,4 \cdot 10^8$ clones individuales por subgenoteca se alcanzó rápidamente después de la transformación de la agrupación de plásmidos en *E. coli* (Tabla 19).

Tabla 19:

	Región estructural	LCDR3/HCDR2		LCDR3/HCDR2/HCDR3		
		diversidad teórica	tamaño de la genoteca	diversidad teórica	tamaño de la genoteca	Porcentaje de clones correctos [%]
Ratón	V κ B3-VH1 RV	2.4E+08	5.1E+08	6.1E+21	4.3E+09	83%
	V κ C3-VH1 RV	2.4E+08	5.5E+08	6.1E+21	3.2E+09	91%
	V κ B3-VH5	2.4E+08	2.3E+08	6.1E+21	1.1E+09	89%
	V κ C3-VH5	2.4E+08	2.6E+08	6.1E+21	2.6E+09	85%
	V κ B1-VH2	2.4E+08	2.6E+08	6.1E+21	2.2E+09	89%
	Las 5 genotecas de ratón	1.2E+09		3.1E+22	1.3E+10	87%
Rata	V κ 1-VH5	2.4E+08	3.8E+08	6.1E+21	1.5E+09	88%

5 Obtuvimos aproximadamente $2,7 \cdot 10^9$ clones individuales por subgenoteca, o $1,3 \cdot 10^{10}$ clones individuales para toda la genoteca HuCAL MOUSE que comprendía las cinco regiones estructurales Fv murinas diversificadas (Figuras 19-24). La diversidad lograda de la genoteca de anticuerpos HuCAL Mouse es 2 órdenes de magnitud mayor que la población de linfocitos B de un ratón inmunológicamente maduro (Rajewsky et al., 1987, Science 238, 1088-1094), y la calidad de la genoteca era muy alta - aproximadamente el 87% de las secuencias analizadas estaban en marco y no contenían mutaciones de aminoácidos indeseables (véase la Tabla 19). El análisis de las secuencias de clones
 10 recogidos al azar indicó que la distribución de aminoácidos en LCDR3 y HCDR3 no variaba sustancialmente de nuestro diseño (Figuras 20-25), y también la distribución por longitud real de HCDR3 estaba de acuerdo con el diseño (Figura 26).

Subclonación del casete de la genoteca LCDR3 en el vector de clonación auxiliar para maduración por afinidad

15 Los casetes de la genoteca LCDR3 amplificados se insertaron en lugar de un relleno de ADN en el vector de clonación pa18 mediante el uso de Bpil/BsiWI. Este vector no posee ningún promotor y se replica en bacterias con un número de copias elevado. Las condiciones de ligación, transformación y propagación de los clones empleados para la subclonación de los casetes LCDR3 en el plásmido pMorph31_mVLj-mVH/AP o en el vector de clonación pGA18, eran básicamente las mismas.

20 La diversidad de los casetes de maduración de LCDR3 se muestra en la Tabla 20

Tabla 20:

	Región estructural	Casete de maduración de LCDR3		
		Diversidad teórica	Tamaño de la genoteca	Porcentaje de clones correctos [%]
Ratón	V κ B1	4.7E+07	2.4E+09	89
	V κ B3	4.7E+07	2.0E+09	96
	V κ C3	4.7E+07	2.0E+09	96
Rata	V κ 1	4.7E+07	4.3E+09	95

Amplificación de los casetes de maduración de HCDR2

25 El inserto que comprendía los casetes de maduración de HCDR2 para mVH, se preparó en dos etapas. En la primera etapa, un oligonucleótido diversificado que contenía trinucleótidos ("TRIM") se mezcló con el oligonucleótido "solapante" que se correspondía a la hebra complementaria que se solapaba con el oligonucleótido que contenía trinucleótidos en posiciones no diversificadas. Ambos oligonucleótidos (a concentraciones de 0,1-0,2 μ M) se ensamblaron y se amplificaron en 10 ciclos de PCR utilizando cebadores "mat" directos e inversos (a una concentración de 1 μ M) y la ADN polimerasa Pfx Platinum (Invitrogen), dando como resultado el producto Ai de la PCR. De forma paralela, la región estructural 3 de mVHi se amplificó a partir del molde pMorph31_mVLj/MBP-mVHi/AP (véase más arriba) en 10 ciclos de PCR utilizando cebadores directos e inversos "FW3" (a una concentración de 1 μ M), dando como
 30

5 resultado el producto *Bi* de la PCR. En una siguiente etapa, los productos correspondientes de la PCR de *Ai* y *Bi* (cada uno a una concentración de -20 nM) se ensamblaron y se amplificaron en 10 ciclos de PCR, utilizando los cebadores directos e inversos "mat" y "FW3", correspondientes, dando como resultado productos de PCR que codificaban HCDR2/FW3*i* diversificada y flanqueados por los sitios de ER XhoI y BssHII y 5' y 3', respectivamente. Las secuencias de todos los oligonucleótidos usados para la preparación del casete HCDR2 de maduración, se enumeran en la Tabla 21.

Tabla 21:

mLib74	mu VH1-RVT_HCDR2mat for	GGTCGCGGCCTCGAGTGG
mLib75	mu VH1-RVT HCDR2mat_overlap	GTTTATCCACGGTCAGGGTCACCCGGCTTTTGAA
mLib76	mu VH1-RVT_HCDR2mat rev	GTTTATCCACGGTCAGGGTCACC
mLib77	mu VH1 RVT FW3 for	GGTGACCCTGACCGTGATAAAC
mLib78	mu VH2_HCDR2mat for	GGCAAAGGTCTCGAGTGG
mLib79	mu VH2 HCDR2mat_overlap	GCTGTTATCTTTGCTGATGCTCAGGCGGCTAATAAA
mLib80	mu VH2_HCDR2mat rev	GCTGTTATCTTTGCTGATGC
mLib81	mu VH2 FW3 for	GCATCAGCAAAGATAACAGC
mLib82	mu VH5_HCDR2mat for	GAAAAACGCCTCGAGTGG
mLib83	mu VH5 HCDR2mat_overlap	CGCGGCTAATGGTAAAGCGGCCTTTCACGCT
mLib84	mu VH5_HCDR2mat rev	CGCGGCTAATGGTAAAGC
mLib85	mu VH5 FW3 for	GCTTTACCATTAGCCGCG
mLib86	mu VH1/2/5 AP FW3 rev	CGACGCCAGCGCGCAATAATA
ratLib88	rat VH1_HCDR2mat for	GGCCAGGGCCTCGAGTGG
ratLib89	at VH1 HCDR2mat_overlap	CCACGGTAAAGGTCGCTTTGCCTTTGAATTT
ratLib90	rat VH1_HCDR2mat rev	CCACGGTAAAGGTCGCTTTG
ratLib91	rat VH1 FW3 for	CAAAGCGACCTTTACCGTGG
ratLib92	rat VH1 FW3-AP rev	CGACGCCAGCGCGCAGTAATACACC
ratLib93	rat VH5_HCDR2mat for	ACCAAAGGTCTCGAGTGG
ratLib94	rat VH5 HCDR2mat_overlap	CACGGCTAATGGTAAAACGGCCTTTCACGCT
ratLib95	rat VH5_HCDR2mat rev	CACGGCTAATGGTAAAACGG
ratLib96	rat VH5 FW3 for	CCGTTTTACCATTAGCCGTG
ratLib97	rat VH5 FW3-AP rev	CGACGCCAGCGCGCAATAATAGGTC

10 Los casetes de H-CDR2 se completaron con regiones flanqueantes constantes, que representaban una parte de la región estructural 2 y la región estructural 3 del anticuerpo. Por lo tanto era posible amplificar con PCR y clonar los casetes en una etapa posterior. Dado que las regiones flanqueantes de los genes maestros de VH son diferentes, había que sintetizar oligos diferentes para los diferentes genes maestros.

mVH1_RV_HCDR2mat

5' GGTCGCGGCCTCGAGTGGATTGGC-MMT1-ATT-MMT2-CCG-MMT1-MMT3-MMT4-MMT1-MMT5-MMT1-TAT-AAC-W1-AAATTCAAAGCCGGGTGACC

15 mVH2_HCDR2mat

5' GGCAAAGGTCTCGAGTGGCTGGGC-W2-ATT-MMT6-MMT7-MMT8-GGC-MMT9-AC
C-MMT1-TAT-AAC-MMT10-GCCTTTATTAGCCGCCTGAGCATC

mVH5_HCDR2mat

5' GAAAACGCCTCGAGTGGGTGGCG-MMT1-ATT-MMT11-MMT12-MMT13-W3-MMT
12-MMT14-MMT15-TATTATW4GATAGCGTGAAAGGCCGCTTTACC,

5 en donde MMT1 - MMT15 son mezclas TRIM y W son nucleótidos oscilantes como se han descrito anteriormente en este documento.

La diversidad del casete de maduración HCDR2 se muestra en la Tabla 22:

Tabla 22:

		Casete de maduración de HCDR2		
	Región estructural	Diversidad teórica	Tamaño de la genoteca	Porcentaje de clones correctos [%]
Ratón	VH1 RV	1.1E+09	4.0E+09	83%
	VH2	9.9E+06	2.3E+09	86%
	VH5	3.6E+09	6.0E+09	74%
Rata	VH5	1.2E+08	4.0E+09	88%

Subclonación de los casetes de maduración de HCDR2 en un vector de clonación auxiliar

10 En primer lugar, subclonamos los tres genes maestros murinos de VH seleccionados, mVH1 RV, mVH2 o mVH5 (fragmentos que codifican los residuos 1-94, es decir, el ORF que parte del extremo N-terminal de FW1 hasta el extremo muy C-terminal de FW3) a partir del vector de presentación de fagos pMorph31 (véase más arriba) en un vector de clonación auxiliar pGA18 mediante el uso de MfeI/BssHII. En la etapa siguiente, la región del gen maestro entre los sitios de ER Xho I y BssH II fue reemplazada por un relleno de ADN (aproximadamente 600 pb). Insertamos el ADN de relleno con el fin de facilitar la subclonación del casete de maduración diversificado de HCDR2 amplificada en los vectores de clonación pGA18_mVH*i* mediante el uso de XhoI/BssHII en una etapa posterior. A continuación, subclonamos la HCDR2 amplificada diversificada (véase el capítulo anterior) en el vector de clonación pGA18_mVH*i* en lugar del relleno de ADN mediante el uso de XhoI/BssHII. Típicamente, se ligaron de 15-20 µg de vector, con un exceso 2 molar del casete HCDR2. Las condiciones de ligación, transformación y propagación de los clones empleados para la subclonación del casete HCDR2 no variaron de las condiciones utilizadas para la subclonación de los casetes LCDR3.

Sumario

25 La genoteca de anticuerpos de ratón generada en el estudio presentado comprende regiones LCDR3 y HCDR3 diversificadas. Los casetes de la genoteca de CDR3 se generaron a partir de trinucleótidos mezclados (TRIM) que codificaban solamente los residuos deseables en los casetes de CDR3, impidiendo la introducción de codones de cisteína y de detención. Puesto que se ha reconocido que la interacción antígeno-anticuerpo implica también residuos de la región estructural y que, por lo tanto, podría ser beneficiosa para la detección de ligantes a partir de una conformación espacial más amplia que la ofrecida por una variedad de secuencias distintas solamente en CDRs, seleccionamos cinco regiones estructurales murinas diferentes de Fv. Por ello, cinco genotecas distintas con LCDR3 y HCDR3 diversificadas se generaron en paralelo. La secuenciación de los clones no seleccionados indicó que la incidencia global de mutaciones, deleciones e inserciones no deseadas de ADN, introducidas mediante la amplificación por PCR de los casetes LCDR3 y HCDR3, era baja (~10%). Con el fin de cumplir con los requisitos esenciales de modularidad de la clonación de HuCAL[®], se introdujeron las siguientes mutaciones en las secuencias de la línea germinal: V2I en mVLB1, T25S en mVH3, K46E en mVH9 y K66R/A67V en mVH1_RV. La secuencia de aminoácidos correspondiente a las posiciones 1 - 4 en todas las mVHs se cambió a EVQL.

40 Una revisión general de todos los genes maestros sintetizados que codifican un dominio variable, se presenta en las Figuras 27-31. Kabat indica la posición del residuo de aminoácido en secuencias VL o VH. mVL indica dominio variable murino en la cadena ligera (exclusivamente de tipo kappa), y mVH indica dominio variable murino en la cadena pesada. Los genes de la línea germinal murina de V_k codifican mVL partiendo de la posición 1 (en el extremo N-terminal de la región estructural 1) hasta 95 (en el extremo C-terminal de L-CDR3), y los genes de la línea germinal murina de J_k codifican el residuo muy C-terminal en L-CDR3 (posición 96 constante) y toda la región estructural 4 en mVL (posiciones 97 - 108). Hemos elegido Jk2 murino, ya que Tyr es el residuo más frecuente en la posición 96 de los anticuerpos reordenados que se encuentran en la base de datos IMGT.

Los genes de la línea germinal murina de VH codifican mVH partiendo de la posición 1 (en el extremo N-terminal de la región estructural 1) hasta 94 (en el extremo muy C-terminal de la región estructural 3). La secuencia de HCDR3 está codificada por los genes de la línea germinal D. Demostramos que las propiedades bioquímicas y biofísicas del dominio VH dependen también de su secuencia HCDR3, y por lo tanto decidimos introducir una secuencia HCDR3 particular en todos los genes maestros de mVH con el fin de comparar los genes maestros. La secuencia HCDR3 (WGGDGFYAMDY) se había obtenido a partir del anticuerpo 4D5, y se había utilizado previamente para la evaluación de genes maestros humanos antes de la construcción de la genoteca HuCAL original (Knappik et al., 2000, J Mol Biol 296, 57-86). Los genes de la línea germinal murina de JH codifican una parte C-terminal de HCDR3 y la región estructural 4 completa (posiciones 103-113). Hemos elegido JH2 murina (secuencia WGQGTTTLTVSS) para el diseño de la región estructural 4 porque JH2 se encontraba con frecuencia en anticuerpos reordenados, depositados en la base de datos IMGT, y nos ha permitido introducir dos sitios de endonucleasas de restricción (Styl y BlnI) que son indispensables para la modularidad de la clonación en HuCAL.

Ejemplo 1.7: Prueba de inmunoadsorción con la genoteca murina

M-CSF murino (factor de estimulación de colonias de macrófagos) se utilizó como antígeno para la prueba de inmunoadsorción. En este caso, manejamos cada subgenoteca separadamente y, además, utilizamos una genoteca completa agrupada que consistía en todas las subgenotecas.

Inmunoadsorción

mM-CSF se diluyó en PBS para recubrir con la concentración estándar de 50 µg/ml. Para cada subgenoteca, se recubrieron 2 pocillos de una placa de microtitulación con antígeno, utilizando 300 µl de solución de antígeno por pocillo. La placa se almacenó O/N (del inglés, "overnight") a 4°C. A continuación, la solución de antígeno se eliminó de los pocillos recubiertos, invirtiendo rápidamente la placa sobre una bandeja de plástico. Los pocillos recubiertos se lavaron dos veces con 400 µl de PBS y se bloquearon durante 2 horas a TA en un agitador de placas de microtitulación.

Mientras tanto, las mezclas de bloqueo de fagos se incubaron en tubos de reacción de 2 ml durante 2 h a TA, agitando suavemente. Después del procedimiento de bloqueo, los pocillos se lavaron 2x con 400 µl de PBS y los 300 µl de la mezcla de fagos bloqueada previamente, se transfirió a cada pocillo bloqueado. Se incubó durante 2 horas a TA en un agitador de placas de microtitulación. Después de eso, la solución de fago procedente de los pocillos recubiertos con antígeno, se eliminó invirtiendo rápidamente la placa sobre una bandeja de plástico y las placas se lavaron con las siguientes condiciones de lavado (Tabla 23):

Tabla 23:

1ª ronda	2ª ronda	3ª ronda
3x PBST rápido	1x PBST rápido	10x PBST rápido
2x PBST durante 5 min	4x PBST durante 5 min	5x PBST durante 5 min
3x PBS rápido	1x PBS rápido	10x PBS rápido
2x PBS durante 5 min	4x PBS durante 5 min	5x PBS durante 5 min

Todas las etapas de lavado se realizaron a TA. Después de las etapas de lavado, todas las trazas de la solución de lavado se eliminaron golpeando cuidadosamente la placa de microtitulación sobre una nueva pila de toallas de papel.

Para la elución del fago unido específicamente, añadimos 300 µl de DTT 20 mM en Tris/HCl 10 mM, pH 8,0, a cada pocillo de selección y se incubaron a TA durante 10 min sin agitación. Se recogió el eluido de DTT de cada selección. TG1 de *E. coli* con una DO_{600nm} de 0,6-0,8 se añadió a los fagos eluidos de cada selección y se incubaron en una incubadora sin agitación. Después de la infección, las bacterias se sembraron en placas, de manera uniforme sobre dos grandes placas de agar LB/Cm/Glu para cada selección y se incubaron O/N a 37°C.

Para las siguientes rondas de inmunoadsorción, las suspensiones bacterianas de cada agrupación se recogieron y se usaron para propagar los fagos para una ronda de inmunoadsorción adicional tal y como se ha descrito anteriormente.

Después de cada ronda de inmunoadsorción frente a mM-CSF, se determinó el título de fagos. El intervalo esperado era desde 1×10^{10} - 1×10^{12} fago/mL para la entrada y 10^4 - 10^6 fago/mL para la salida. La Tabla 24 muestra la entrada y la salida después de cada ronda de selección y todos los valores están en el intervalo esperado.

Tabla 24:

Agrupación	1ª ronda		2ª ronda		3ª ronda	
	ENTRADA	SALIDA	ENTRADA	SALIDA	ENTRADA	SALIDA
VH1 VkB3	1x10 ¹²	2,6x 10 ⁵	3,4 x 10 ¹⁰	8x 10 ⁴	2 x 10 ¹⁰	2,6 x 10 ⁶
VH2 VkB1	1x10 ¹²	8 x 10 ⁴	4,6 x 10 ⁹	4,1 x 10 ⁴	3,6 x 10 ⁹	3,2 x 10 ⁶
VH5 VkB3	1x10 ¹²	3,7 x 10 ⁴	2,6x 10 ¹⁰	5,2 x 10 ⁴	4x 10 ⁹	1,8 x 10 ⁵
VH1 VcC3	1x10 ¹²	1,6 x 10 ⁵	2 x 10 ¹⁰	1,8 x 10 ⁴	5,8 x 10 ⁹	4 x 10 ⁵
VH5 VcC3	1x10 ¹²	8 x 10 ⁴	2,8 x 10 ¹⁰	1,2 x 10 ³	1,8 x 10 ¹⁰	1,2 x 10 ⁶
Genoteca completa	1x10 ¹²	2,4 x 10 ⁵	2,4 x 10 ¹⁰	2,88 x 10 ³	1,4 x 10 ¹⁰	2 x 10 ⁶

Subclonación de agrupaciones seleccionadas

5 El ADN se preparó y se digirió con *EcoRI*, *XbaI* y *BmtI*. Los ácidos nucleicos se separaron en un gel para ADN con agarosa preparativa al 1% y se extrajo la banda relacionada con el ADN que codificaba Fab (~ 1470 pb), se ligó en el vector pMorphx9_FH y se sometió a electroporación en células competentes TG1 F-. Las colonias individuales se seleccionaron y se incubaron en placas de microtitulación de fondo redondo de 96 pocillos, en 100 µl de medio 2xYT/Cm/1% de Glu, O/N a 30°C. A la mañana siguiente, se añadieron 100 µl de glicerol estéril al 30% en 2xYT a cada pocillo de la placa maestra, O/N y se almacenaron a -80°C.

Microexpresión de la proteína Fab

10 Para la preparación de las placas de expresión, se replicaron 5 µl por pocillo desde la placa maestra al pocillo correspondiente de la placa de microtitulación de 96 pocillos que contenía 100 µl de medio 2xYT/Cm/0,1% de Glu por pocillo. Después, las placas de microtitulación se agitaron a 30°C y 400 rpm hasta que los cultivos estuvieron ligeramente turbios (~ 2-4 horas) con una DO600nm de ~ 0,5, después de la inducción mediante la adición de 20 µl de 2xYT/Cm que contenía IPTG 3 mM. Cada placa se selló con cinta permeable a los gases y se cubrió con una tapa y se incubó O/N a 30°C con agitación a 400 rpm.

Escrutinio con ELISA primaria

20 Se recubrieron 384 pocillos de placas de microtitulación Maxisorp con 20 µl de solución de antígeno (mM-CSF) por pocillo a una concentración de 1 µg/ml, diluido en PBS y se incubó a 4°C O/N. Los controles se aplicaron en la forma descrita a continuación. La solución de antígeno se retiró de la placa ELISA recubierta, invirtiéndola rápidamente sobre el fregadero y lavando los pocillos una vez con TBST. La placa de ELISA se bloqueó con 100 µl de MTBST al 5% durante 2 horas a TA, agitando suavemente. A cada pocillo de la placa de expresión, se añadieron 40 µl de tampón BEL que contenía 2,5 mg/ml de lisozima y se agitó durante 1 hora a 22°C a 400 rpm. Después de la lisis de las bacterias, se añadieron 40 µl de MTBS al 12,5% por pocillo y se incubó durante 30 min adicionales a 400 rpm y 22°C para el bloqueo. La placa del escrutinio con ELISA bloqueada se lavó una vez con TBST y se golpeó sobre una pila de toallas de papel. Después, 20 µl de los extractos BEL bloqueados se transfirieron desde la placa de expresión al pocillo correspondiente de la placa de ELISA bloqueada. Asimismo, 20 µl del extracto de control positivo para la expresión se transfirieron a un pocillo de la placa de ELISA. La placa de ELISA se incubó durante 1,5 h a TA en un agitador de placas de microtitulación, agitando suavemente. Las placas de ELISA se lavaron 5x rápidamente con TBST. Después, se añadieron 20 µl de conjugado de anticuerpo de cabra anti-F(ab')₂ de ratón-AP, diluido 1:5000 en TBST y se incubó 1 h a TA. Después de 1 h de incubación, las placas de ELISA se lavaron 5 veces rápidamente con TBST y 20 µl de sustrato AttoPhos™ se añadieron a cada pocillo (diluidos 1:5 en TBST). Se registraron las lecturas de la fluorescencia con una excitación de 440 ± 25 nm y una emisión de 550 ± 35 nm.

Controles:

35 pocillo E12 => control positivo para la expresión y la extracción
 pocillo F12 => control de ruido de fondo del conjugado de anticuerpo de cabra anti-F(ab')₂ de ratón-AP
 pocillo G12 => control de ruido de fondo del sustrato AttoPhos™
 pocillo H12 => ruido de fondo de la placa de ELISA

40 Para la selección de los mejores ligantes, se realizó un examen con ELISA después de tres rondas de inmunoadsorción de fago sólido. Se recogieron 184 de los clones transformados para cada subgenoteca y 2x184 para la genoteca completa y se trasladaron a una placa de microtitulación de 96 pocillos para las placas maestras. Después se indujo una expresión periplasmática y se realizó un escrutinio con ELISA con mM-CSF como antígeno para el recubrimiento. La Tabla 25 muestra como ejemplo de la fluorescencia, las lecturas de ELISA de una placa del escrutinio de la genoteca completa. Solo se seleccionaron los aciertos con una señal de intensidad de 5-10 veces (en verde) y

más de 10 veces (en azul) el ruido de fondo. Los aciertos con una señal de intensidad 5,2 veces el ruido de fondo están en amarillo.

Tabla 25:

	1	2	3	4	5	6	7	8	9	10	11	12
A	1379	6959	579	954	1573	506	368	353	262	346	631	2045
B	970	720	723	583	362	411	2898	294	232	373	263	303
C	586	4922	919	6965	1151	341	291	433	322	747	248	284
D	2970	3118	520	527	427	388	323	220	317	455	241	570
E	1026	1419	696	662	475	307	319	8265	301	276	285	12819
F	709	3998	1126	563	544	1962	561	298	741	261	310	251
G	818	1472	1303	2730	562	328	321	291	896	244	550	150
H	9040	853	1583	788	941	375	555	314	397	291	2065	298

5 *Secuenciación de los Fabs positivos para ELISA*

Se secuenció la cadena pesada y ligera de los aciertos primarios. Las Tablas 26 (VH1 Vkc3), 27 (VH5Vkc3), 28 (genoteca completa), 29 (VH1 Vkb3), 30 (VH2Vkb1) y 31 (VH5Vkb3) muestran las regiones CDRs de algunos ligantes únicos. Los ligantes aislados a partir de las mismas subgenotecas comprenden las mismas regiones CDR1 y CDR2.

10 Tabla 26 (ligantes seleccionados a partir de la subgenoteca VH1 Vkc3):

Cadena pesada

Clon	HCDR1	HCDR2	HCDR3
mM-CSF1.3.7C4	YTFTSYWMH	RIDPNSGGTKYNEKFKS	APYYAFAY
mM-CSF1.3.7E5	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SDYRGAFAY
mM-CSF1.3.7G2	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYYYAFDV
mM-CSF1.3.7G7	YTFTSYWMH	RIDPNSGGTKYNEKFKS	DAGVDDFDV
mM-CSF1.3.8A3	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SRAYAGYVFDY
mM-CSF1.3.8B9	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYSYEAFDY
mM-CSF1.3.7G5	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYTSRFDY

Cadena ligera

Clon	LCDR1	LCDR2	LCDR3
mM-CSF1.3.7C4	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QQYNSYPL
mM-CSF1.3.7E5	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QQYWSRPY
mM-CSF1.3.7G2	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	SQYWSYPP
mM-CSF1.3.7G7	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QQSSHYP
mM-CSF1.3.8A3	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QQSNAYPR
mM-CSF1.3.8B9	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	SQKIQRPY
mM-CSF1.3.7G5	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QQDYEPY

Tabla 27 (ligantes seleccionados a partir de la subgenoteca VH5Vkc3):

Cadena pesada

Clon	HCDR1	HCDR2	HCDR3
mM-CSF1.3.9C8	FTFSSYTMSWVRQTPEKRL E	TISGGGGNTYYPDSVKG	RWYYAFDY
mM-CSF1.3.10G5	FTFSSYTMSWVRQTPEKRL E	TISGGGGNTYYPDSVKG	RYYHYFSDFDY
mM-CSF1.3.10H5	FTFSSYTMSWVRQTPEKRL E	TISGGGGNTYYPDSVKG	WIYYAFDY

Cadena pesada

Clon	HCDR1	HCDR2	HCDR3
------	-------	-------	-------

Cadena ligera

Clon	LCDR1	LCDR2	LCDR3
mM-CSF1.3.9C8	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QLGQEKPY
mM-CSF1.3.10G5	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QQGNHFPA
mM-CSF1.3.10H5	KASQSVDYDGDSYMN	YAASNLESGIPARFSGS	QQGTSKPY

Tabla 28 (ligantes seleccionados a partir de la genoteca completa):

Cadena pesada

Clon	Región estructural	HCDR1	HCDR2	HCDR3
mM-CSF1.3.12D6	VH1	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SDGSTYAFAY
mM-CSF1.3.13A8	VH1	YTFTSYWMH	RIYPNSGGTKYNEKFKS	SSSYTYAFDY
mM-CSF1.3.13D4	VH1	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SDSYTYAFAY
mM-CSF1.3.13G11	VH1	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYYYAFDY
mM-CSF1.3.14C2	VH1	YTFTSYWMH	RIDPNSGGTKYNEKFKS	TYYYAFDY
mM-CSF1.3.14D1	VH1	YTFTSYWMH	RIDPNSGGTKYNEKFKS	TYKYAFDY
mM-CSF1.3.13E10	VH1	YTFTSYWMH	RIDPNSGGTKYNEKFKS	STAYGYRYFDY
mM-CSF1.3.14G10	VH1	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYKYAFDY

Cadena ligera

Clon	Región estructural	LCDR1	LCDR2	LCDR3
mM-CSF1.3.12D6	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYSEYPF
mM-CSF1.3.13A8	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYSSYPF
mM-CSF1.3.13D4	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	FQYYERPH
mM-CSF1.3.13E10	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	EQIYSFPL
mM-CSF1.3.13G11	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYYSYPF
mM-CSF1.3.14C2	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYHKYPF
mM-CSF1.3.14D1	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYYEKPY
mM-CSF1.3.14G10	VkB3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYNHVFP

Tabla 29 (ligantes seleccionados a partir de la subgenoteca VH1 VkB3):

Cadena pesada

Clon	HCDR1	HCDR2	HCDR3
mM-CSF1.3.15B8	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYRSYMDY
mM-CSF1.3.15G8	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYRSYFDY
mM-CSF1.3.15A10	YTFTSYWMH	RIDPNSGGTKYNEKFKS	SYRSYFDV

Cadena ligera

Clon	LCDR1	LCDR2	LCDR3
mM-CSF1.3.15B8	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQHWEYPH
mM-CSF1.3.15G8	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYWQYPF
mM-CSF1.3.15A10	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	SQYWSYPF

Tabla 30 (ligantes seleccionados a partir de la subgenoteca VH2VkB1):

Cadena pesada			
Clon	HCDR1	HCDR2	HCDR3
mM-CSF1.3.17F6	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	SYYGSKYPFDY
mM-CSF1.3.18D8	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	SYYGIFDY
mM-CSF1.3.18H4	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	WIYYAFDY
Cadena ligera			
Clon	LCDR1	LCDR2	LCDR3
mM-CSF1.3.17F6	RSSQSIVHSNGNTYLE	YKVSNRFSGVPDRFSGS	AQRSHYPL
mM-CSF1.3.18D8	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQEEHVPL
mM-CSF1.3.18H4	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQYSSYPF

Tabla 31 (ligantes seleccionados a partir de la subgenoteca VH5VkB3):

Cadena pesada			
Clon	HCDR1	HCDR2	HCDR3
mM-CSF1.3.19C9	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	MYYSYGMTFDY
mM-CSF1.3.20C3	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	MYYASGMAFDY
mM-CSF1.3.20D7	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	YRYGSGYYFDY
mM-CSF1.3.20F1	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	YTYSMGYYFDY
mM-CSF1.3.20F9	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	RYYDNLFAV
mM-CSF1.3.20A6	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	SGYGYYFAY
mM-CSF1.3.20A12	FTFSSYTMSWVRQTPEKRLE	TISGGGGNTYYPDSVKG	TSSSFWYAFDY
Cadena ligera			
Clon	LCDR1	LCDR2	LCDR3
mM-CSF1.3.19C9	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQDDHYPY
mM-CSF1.3.20C3	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQDNEYPY
mM-CSF1.3.20D7	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQGSHYPR
mM-CSF1.3.20F1	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQGSHYPR
mM-CSF1.3.20F9	RASKSVSTSGYSYMH	YLASNLESGVPARFSGS	QQEEEWPD
mM-CSF1.3.20A6	KASOSVDYDGD SYMN	YAASNLESGIPARFSGS	SQRSHYPQ
mM-CSF1.3.20A12	KASQSVDYDGD SYMN	YAASNLESGIPARFSGS	QQLYSYPK

- 5 Una diversidad de combinaciones diferentes de las regiones estructurales se aisló a partir de la inmunoadsorción con la genoteca completa, comprendiendo los ligantes más abundantes una combinación VH1 VkB3. La Figura 32 muestra una vista general de las combinaciones identificadas.

Expresión del fragmento Fab de anticuerpo

- 10 En primer lugar, un precultivo (10 ml) de ligantes transformados recientemente en TG1 F- de *E. coli* en el vector de expresión pMORPHx9_FH, se inoculó en medio 2xYT/Cm/1% de Glu y se incubó durante 3-5 horas a 30°C. Después de esto, 500 ml de medio 2xTY/cm/0,1% de Glu con todo el volumen del precultivo, se inoculó y se incubó a 30°C en un agitador a 200 rpm hasta que se alcanzó una DO600nm de 0,5. Se añadió IPTG (1 M) hasta tener una concentración final de 0,75 mM para la inducción de la expresión (375 µl/500 ml de cultivo). Los Fabs se expresaron a 30°C O/N con agitación a 200 rpm. Al día siguiente, las bacterias se centrifugaron a -4000 g durante 30 min a 4°C
- 15 y los sedimentos se congelaron a 20°C durante al menos O/N.

Purificación empleando IMAC y SEC

Los sedimentos bacterianos se resuspendieron en tampón de lisis que comprendía Tris 25 mM pH 8,0, NaCl 0,5 M, 0,1% de lisozima, MgCl₂ 2 mM y 10 U/ml de benzonasa. Después de la incubación, el sobrenadante se filtró a través de un filtro de 0,2 µm. La purificación IMAC del Fab con marcador His6 usando Aekta Express, se aplicó para la purificación totalmente automatizada a temperatura ambiente de 20°C: las muestras filtradas se cargaron en una columna IMAC (HiTrap 1 ml Chelating HP), se lavaron con tampón fosfato de Na 20 mM pH 7,4, NaCl 500 mM, imidazol 20 mM, y se eluyeron con imidazol 250 mM/tampón fosfato de Na 20 mM /NaCl 500 mM, pH 7,4. Los picos eluidos se aplicaron a una columna de filtración en gel (HiLoad 16/60 Superdex 75) y la proteína purificada se eluyó en una placa de pocillos profundos en PBS. La Figura 33 muestra un resumen de los rendimientos de la expresión obtenidos de fragmentos Fab específicos de mM-CSF.

5

10 *Determinación de las constantes de afinidad mediante ELISA*

Los pocillos de una placa de microtitulación se recubrieron con 1 µg/ml de mM-CSF en PBS, 50 µl/pocillo, O/N a 4°C. Después de lavar, los pocillos se bloquearon con 350-400 µl de MTBST al 5% durante 2 h a TA. Mientras tanto, se prepararon diluciones de cada Fab purificado para la prueba: 1000/100/10/1/0,1/0,01 nM de concentración (por duplicado). Las placas de microtitulación de ELISA bloqueadas se lavaron una vez con TBST. Después de esto, 50 µl de las diluciones preparadas de los Fabs purificados se transfirieron a los pocillos correspondientes de las placas de ELISA.

15

NK1: Recubrimiento con antígeno de selección, + bloqueo, + PBS, - anticuerpo secundario conjugado + sustrato

NK2: Recubrimiento con PBS, + bloqueo, + PBS, + PBS, + anticuerpo secundario conjugado, + sustrato

20

NK3: Recubrimiento con antígeno de la selección, + bloqueo, + PBS, + anticuerpo secundario conjugado, + sustrato

Las placas de ELISA se incubaron con diluciones de proteína durante 2 h a TA en un agitador de placas de microtitulación, se lavaron rápidamente 3x con TBST y 100 µl de conjugado de anticuerpo Fab de cabra anti-ratón-AP se añadieron y se incubaron durante 1 h a TA agitando suavemente. La placa se lavó 3 veces rápidamente con TBST y se añadieron 100 µl de sustrato AttoPhos (1:5) a cada pocillo y se midió la fluorescencia (lector de ELISA). Los datos se analizaron utilizando un ajuste de 4 parámetros. La Figura 34 muestra los resultados para algunos fragmentos Fab específicos de mM-CSF. Más del 30% de los Fabs aislados tiene un valor de CE50 en el intervalo de 1 -10 nM, 17% de los Fabs en el intervalo de 100-1000 nM, 15% en el intervalo de 10-100 nM y 7% en el intervalo picomolar.

25

Conversión de Fab y expresión y purificación de un mlgG2

Uno de los fragmentos Fab del anticuerpo que se unía a M-CSF murino y que mostró actividad de unión en un ELISA, se convirtió al formato IgG2a murino por clonación de las secuencias respectivas de VH y VL en un vector de expresión eucariota apropiado que albergaba las cadenas ligera constante y gamma 2a murina, respectivas. Este vector de expresión se utilizó para la transfección de las células HEK-293. Los anticuerpos IgG2a se purificaron posteriormente a partir del material sobrenadante de un cultivo de células HEK293 mediante cromatografía de afinidad con proteína A en un sistema Äkta.

30

35

Determinación de la afinidad de un anticuerpo IgG2a murino mediante BiaCore

La unión del anticuerpo de ratón anti-mM-CSF tal como el subtipo mlgG2a se midió frente a mM-CSF inmovilizado. Para ello, se utilizó HBS-EP como tampón de migración a 30 µl/min a 20°C y el anticuerpo se diluyó en tampón de migración hasta las siguientes concentraciones: 10, 5, 2,5, 1, 0,5 nM. A continuación, se seleccionó la configuración del ciclo: asociación durante 180 s, disociación durante 900 s, regeneración durante 30 s utilizando glicina 10 mM, pH 2,0 a 30 µl/min y una estabilización final durante 60 s. Todas las concentraciones se midieron por duplicado y la referencia se restó (canal en blanco 1). Los sensogramas se ajustaron utilizando un modelo de unión de "ligando bivalente" y las constantes se muestran en la tabla 32.

40

Tabla 32:

Ligando	ka (1/Ms)	SE (ka)	kd (1/s)	SE (kd)	Rmax (RU)	KD (M)	tc	SE (tc)	Chi ² (RU ²)
mM-CSF	2.83 E+05	1.3 E+03	< E-06	< E-08	local	< E-12	1.41 E+09	1.5 E+07	4.5

45

Ejemplo 1.8: Prueba de inmunoadsorción con la genoteca murina

GM-CSF murino (factor de estimulación de colonias de granulocitos macrófagos) se utilizó como antígeno para otra prueba de inmunoadsorción. La inmunoadsorción, la selección y la caracterización del ligante eran esencialmente idénticas a las del Ejemplo 1.7 anterior en esta memoria, excepto que se utilizó GM-CSF en lugar de M-CSF.

Uno de los fragmentos de anticuerpo Fab que se une a GM-CSF murino y que muestra actividad en un ensayo de proliferación con FDCP-1, se convirtió al formato IgG2a murino mediante clonación de las respectivas secuencias de VH y VL en un vector de expresión eucariota apropiado que albergaba la cadena ligera constante y la cadena gamma 2a murina, respectivas. Este vector de expresión se utilizó para la transfección de células HKB11. Los anticuerpos IgG2a se purificaron posteriormente a partir del material sobrenadante del cultivo de células HKB11 mediante cromatografía de afinidad con Proteína A. El fragmento Fab y el anticuerpo IgG2a murino se sometieron entonces a un ensayo comparativo en un ensayo de proliferación con FDCP-1. Brevemente, la línea celular murina FDCP-1 prolifera en presencia de GM-CSF murino. Diferentes diluciones de IgG2a o Fab anti-GM-CSF murino se añadieron a los pocillos que contenían células FDCP-1 y GM-CSF murino. Después de una incubación durante 72 h (37°C; 5% de CO₂) se midió la viabilidad celular mediante la adición de reactivo XTT (Roche) de acuerdo con las recomendaciones del fabricante.

Los resultados se muestran en la Figura 46. Ambos ligantes eran activos en el ensayo con una eficacia comparable, lo que demuestra que los ligantes funcionales se pueden aislar de la genoteca, lo que también conserva su actividad después de la conversión en formato de inmunoglobulina de longitud completa.

15 Ejemplo 2: Generación de una genoteca de anticuerpos de rata (HuCAL RAT)

Ejemplo 2.1: Diseño de los genes maestros de anticuerpo de rata

Dado que no se han publicado datos sobre el uso de familias de la línea germinal de VH-VL de rata, se sometieron a análisis todas las secuencias públicas de anticuerpos de rata disponibles (498 secuencias en total). Las secuencias se clasificaron en familias de la línea germinal mediante la comparación de las secuencias reordenadas para los genes de la línea germinal. En total, 498 secuencias de VH de rata reordenadas se compararon con 119 genes de la línea germinal de rata, extraídos a partir de IMGT. El uso de aminoácidos dependiente de la posición de cada secuencia reordenada se comparó con el uso de aminoácidos de cada gen de la línea germinal de la cepa *Rattus norvegicus*. Al hacerlo, cada secuencia reordenada fue asignada a un gen de la línea germinal más cercana, conduciendo a la siguiente distribución (Tabla 33):

25 Tabla 33:

VH1	48
VH2	25
VH3	
VH4	
VH5	415
VH6	3
VH7	4
VH8	
VH9	
VH10	1
VH11	2
SUMA	498

Un análisis más detallado para la identificación de los genes maestros de VH dentro de los genes de la línea germinal, se realizó exclusivamente para las familias de la línea germinal de VH que incluían una cantidad razonable de secuencias reordenadas de VH1 y VH5.

30 Se emplearon dos metodologías para la selección de los genes de la línea germinal usados más frecuentes, que se iban a utilizar como genes maestros:

- a) Identificación del gen más destacado de la línea germinal dentro de las respectivas familias de la línea germinal
- b) Identificación del gen de la línea germinal que es más parecido al consenso de todos los genes de la línea germinal dentro de una familia.

Idealmente, ambas metodologías darían como resultado el mismo gen de la línea germinal.

Ejemplo 2.1.1: Identificación del gen más destacado de la línea germinal de VH dentro de las respectivas familias de la línea germinal

40 El aminoácido dependiente de la posición presente en cada secuencia reordenada se comparó con el aminoácido presente en los genes de la línea germinal de rata, con el fin de identificar los genes de la línea germinal de uso más

frecuente. Cada una de las secuencias reordenadas fue asignada a un gen de la línea germinal de rata más próximo. Si se identificaban dos o más genes de la línea germinal como el gen de la línea germinal más próximo para una secuencia reordenada, todos los genes de la línea germinal se consideraron genes de la línea germinal más próximos. Para el cálculo del gen de la línea germinal usado preferiblemente dentro de una familia, se sumó el uso de cada gen de la línea germinal.

5

Los resultados para los genes de la línea germinal preferidos dentro de una familia no eran tan claros como para la selección del gen de la línea germinal de ratón. Los siguientes genes de la línea germinal son los dos más destacados en las diferentes familias de VH (Tabla 34):

Tabla 34:

Familia de la línea germinal	Nombre del gen destacado de la línea germinal	Nº de secuencias reordenadas	% de genes destacados de la línea germinal en secuencias reordenadas
VH1	AABR03048898 IGHV1S15*01	48	10
VH1	>AABR03050300 IGHV1S23*01	48	10
VH5	AABR03048730 IGHV5S45*01	415	16
VH5	AABR03048499 IGHV5S23*01	415	14

10

Ejemplo 2.1.2: Identificación del gen de la línea germinal que es el más similar al consenso de todos los genes de la línea germinal dentro de una familia

Se prepararon las secuencias consenso de la línea germinal de 25 genes individuales de la línea germinal de VH1 y 21 genes individuales de la línea germinal de VH5.

15 Consenso_ de todos los genes de la línea germinal de VH1:

QVQLQQSGAELVKPGASVKISKASGYTFTSYMH..VVKQRPGQGLEWIGYIYPGS..GGTNYNE
KFKGKATLTVDKSSSTAYMQLSSLTSEDSAVYYCAR

Consenso_ de todos los genes de la línea germinal de VH5:

EVQLVESGGGLVQPGRSLKLSCAASGFTFSNYGMA..WVRQAPTKGLEWVASISTDG..SSTYYRD
SVKGRFTISRDNKSTLYLQMDSLRSEDATYYCAR

20

Estas secuencias de consenso se compararon con cada gen individual de la línea germinal de la familia VH1 y la familia VH5, respectivamente, y los genes individuales de la línea germinal se evaluaron en relación con la identidad con la secuencia de consenso (Tabla 35).

Tabla 35:

Familia de la línea germinal	Nombre del gen destacado de la línea germinal	Nº de desviaciones frente al consenso [aa]
VH1	>AABR03050300 IGHV1S23*01	11 (de 98 aa)
VH5	AABR03048730 IGHV5S45*01	2 (de 98 aa)

25

Este resultado fue confirmado por análisis BLAST de las secuencias de consenso frente a todos los genes de la línea germinal.

Ejemplo 2.1.3: Confección de un gen maestro de VL de rata

Las cadenas ligeras kappa constituyen la mayoría de las cadenas ligeras de rata, que se encuentran en el 90% de las cadena ligeras agrupadas aisladas de ratas de laboratorio (Frank y Gutman, 1988, Mol Immunol 25, 953-960). Puesto que no estaban disponibles secuencias de la línea germinal de VL y solo estaba disponible un conjunto limitado de secuencias reordenadas, se realizó una metodología diferente para el desarrollo de un gen maestro VL de rata. Las 44 secuencias disponibles se agruparon de acuerdo con la longitud de la L1CDR, seguido por la formación de secuencias de consenso.

30

Se utilizaron las secuencias de consenso de un primer grupo (longitud de L1CDR = 17 aa, número de secuencias =

11) y de un segundo grupo (longitud de LCDR1 = 11 aa, número de secuencias = 33) como genes maestros de VL de rata en la genoteca. Véase la Figura 35.

Ejemplo 2.1.4: Resumen

5 No se requirieron modificaciones de los genes de la línea germinal (para VH) o de la secuencia de consenso (para VL) por razones de modularidad, puesto que todos los sitios de restricción requeridos ya estaban presentes en los genes seleccionados. La Tabla 36 resume los genes de la línea germinal más próximos a los genes maestros seleccionados:

Tabla 36:

GM	Nombre del gen de la línea germinal más próximo
rVH1	IGHV1S23*01
rVH5	IGHV5S45*01
rVL1	Secuencia de consenso 1
rVL2	Secuencia de consenso 2

10 Para completar los genes de la línea germinal de rVH en genes maestros de longitud completa, se añadió a las secuencias, la región estructural 4 del anticuerpo. IMGT proporciona acceso a las regiones J de rata, que forman la región estructural 4 del anticuerpo en el desarrollo. JH2 de la región J (secuencia WGQGVMVTSS) fue seleccionada para la confección del gen maestro VH, ya que se encuentra con frecuencia en anticuerpos reordenados, depositados en la base de datos IMGT. Además, JH2 proporciona los sitios de ER requeridos (Styl y Blpl). Véase la Figura 36.

15 Para completar la rVL en un gen maestro de longitud completa, se seleccionó Jk2-3 de las regiones J, ya que Tyr se encontraba con más frecuencia en el conjunto de secuencias reordenadas que se utilizaron para el análisis. Se añadieron los últimos aminoácidos de la región estructural 4 (TA) para incluir un sitio BsiWI. Véase la Figura 27.

Ejemplo 2.2: Diseño de CDRs de rata

20 *Diseño de casetes de genotecas de L-CDR3 y H-CDR3*

Diseño del casete de maduración de H-CDR2

25 Una visión general de los genes maestros sintetizados que codificaban un dominio variable se presenta en la Figura 21. La numeración de Kabat indica la posición de un residuo de aminoácido en una secuencia de VL o VH. rVL indica un dominio variable de rata en la cadena ligera (exclusivamente de tipo kappa) y rVH indica un dominio variable de rata en la cadena pesada. Las dos secuencias de consenso Vk de rata rVL1 y VL2 codifican VL partiendo de la posición 1 (en el extremo N-terminal de la región estructural 1) hasta 95 (en el extremo C-terminal de L-CDR3). Los genes de la línea germinal de Jk de rata codifican el residuo muy C-terminal en L-CDR3 (posición 96 constante) y toda la región estructural 4 en rVL (posiciones 97 - 108). Hemos elegido Jk2 de rata (Burstein et al., 1982, Proc Natl Acad Sci USA 79, 5993-5997), ya que Tyr es el residuo más frecuente en la posición 96 de anticuerpos reordenados que se encuentran en la base de datos IMGT.

30 Los genes de la línea germinal de VH de rata codifican rVH partiendo de la posición 1 (en el extremo N-terminal de la región estructural 1) hasta 94 (en el extremo muy C-terminal de la región estructural 3). La secuencia H-CDR3 está codificada por los genes de la línea germinal D. Llegamos a la conclusión de que las propiedades bioquímicas y biofísicas del dominio VH dependen también de su secuencia HCDR3 y, por lo tanto, con el fin de comparar los genes maestros, decidimos introducir una secuencia HCDR3 particular obtenida a partir del anticuerpo 4D5 en los dos genes maestros rVH (véase más abajo). Los genes de la línea germinal de JH de rata codifican una parte C-terminal de HCDR3 y la región estructural 4 completa (posiciones 103-113). Hemos elegido JH2 de rata (secuencia WGQGVMVTSS) para el diseño de la región estructural 4 porque JH2 se encontraba con frecuencia en anticuerpos reordenados depositados en la base de datos IMGT, y nos permitía introducir dos sitios de endonucleasas de restricción (Styl y Blpl) que son indispensables para la modularidad de clonación de HuCAL.

35 Con el fin de cumplir con los requisitos esenciales de la modularidad de clonación de HuCAL[®], la secuencia de aminoácidos correspondiente a las posiciones de 1 - 4 en ambas rVHS, se fijó a EVQL.

Ejemplo 2.3: Clonación, expresión y purificación de la genoteca de rata

Cepas bacterianas y vectores utilizados para la clonación

45 Como se ha descrito anteriormente para la genoteca murina, la clonación molecular se llevó a cabo usando las cepas de *E. coli* XL1-Blue (Stratagene) o Top10F+ (Invitrogen). Los productos de la síntesis de genes (genes maes-

5 tros) se clonaron primero en un vector pGA en GeneArt AG, y luego se volvieron a clonar en el vector pMorph30 tricistrónico CysDisplay tal y como se describe a continuación. El vector pMorph30 presentado en fagos se había obtenido a partir del plásmido pMorph23 (Rothe et al., 2008, J Mol Biol 376, 1182-1200), básicamente de la siguiente manera: el marcador hexa-His fijado al extremo C-terminal de VH fue sustituido por la secuencia de aminoácidos HKHKHK, y un sitio chi en gIII fue eliminado. Las endonucleasas de restricción eran de New England Biolabs o Fermentas, y la ADN ligasa de T4 era de Invitrogen. Las condiciones de la reacción se llevaron a cabo de acuerdo con las recomendaciones del fabricante.

Construcción de genes

10 Números de orden de las secuencias de la línea germinal de la sub-base de datos de *Rattus norvegicus* de IMGT/NCBI que se utilizaron para la construcción de los genes maestros:

rVH1: IGHV1S23*01/AABR03050300, rVH5: IGHV5S45*01/AABR03048730, Jk2: IGKJ2-3*01/ J00746, JH2: IGHJ2*01 / X56791, Ck:IGKC*01 / V01241, CH1 + bisagra: IGHG1*01/ AABR03048905.

Consideraciones para la síntesis de genes maestros

15 Como se ha descrito anteriormente, los genes correspondientes a las secuencias seleccionadas, representativas de la línea germinal (los números de orden se proporcionan a continuación), se optimizaron a nivel de ADN mediante la aplicación de los criterios siguientes: (i) el uso de codones raros tiene que ser óptimo para *E. coli*, mientras que se evitaban codones raros murinos/humanos siempre que fue posible; (ii) las características indeseables de ARN/ADN como cajas TATA internas, sitios chi, sitios de entrada ribosómica, segmentos de secuencias ricas en AT o GC, secuencias que impiden el transporte de ARNm en el citoplasma (elementos INS o CRS), secuencias repetidas, 20 estructuras secundarias de ARN y sitios donantes y aceptores de corte y empalme crípticos se evitaban; (iii) se evitó la introducción de sitios de endonucleasas de restricción (ER) que estaban en conflicto con la modularidad de la clonación de HuCAL[®]; y (iv) se introdujeron sitios de endonucleasas de restricción necesarios para la clonación modular. La glicosilación en los dominios variables no es deseable principalmente porque se ha observado que la fijación de hidratos de carbono afecta frecuentemente a la unión del antígeno debido a un impedimento estérico, y también es responsable de una heterogeneidad de las proteínas expresadas en hospedadores eucariotas. Por lo tanto, comprobamos que los genes maestros de VL/VH diseñados no codificaban proteínas con sitios potenciales para la glicosilación ligada a N.

Dominios constantes en formato Fab

30 El dominio constante de la cadena ligera kappa está codificado por el gen Ck. El residuo de alanina N-terminal fue reemplazado por Thr lo que permitió la introducción del sitio BsiWI en esta posición a nivel de ADN. Un residuo de cisteína localizado en el extremo C-terminal de Ck y que normalmente participa en la formación del enlace disulfuro que une las cadenas ligera y pesada en moléculas de IgG o Fab naturales, fue reemplazado por alanina. Puesto que el residuo Cys en el extremo C-terminal "libre" se elimina, las cadenas ligeras y pesadas ya no estaban unidas covalentemente entre sí en nuestro formato Fab modificado. Esta modificación es la característica principal de CysDisplay[®] que permite la elución selectiva de fagos unidos (Rothe et al., 2008, J Mol Biol 376, 1182-1200).

35 El dominio constante de la cadena pesada en formato Fab está codificado por el fragmento génico CH1. Hemos elegido el gen de la línea germinal IGHG1*01, que se parece más a la secuencia de CH1 en un anticuerpo anti-CD52 de rata con éxito terapéutico, CAMPATH-1 (Cheetham et al., 1998, J Mol Biol 284, 85-99). La adición de un residuo Ala en frente de la secuencia N-terminal ETTAPS, nos permitió la introducción del sitio BspI. Se añadieron los primeros tres residuos de la bisagra (secuencia PRD), seguidos por los residuos Glu-Phe necesarios para la introducción de un sitio EcoRI en el extremo C-terminal del dominio CH1.

Ensamblamiento génico en el plásmido tricistrónico presentado en fagos

45 Los genes diseñados que codificaban dominios constantes o variables de rata, fueron generados mediante ensamblamiento con PCR de una serie de oligonucleótidos sintetizados, y se clonaron en un vector de clonación pGA intermedio en GeneArt AG. Era cómodo y rápido reemplazar los genes de VL, VH, Ck y CH1 humanos por equivalentes de rata en dos moléculas de vectores intermedios diferentes. Los derivados del plásmido pMorph30 (Rothe et al., 2008, J Mol Biol 376, 1182-1200) denominados "plásmido L" y "plásmido H", se construyeron con un fragmento relativamente grande de ADN simulado con el fin de facilitar la clonación de la siguiente manera: en el "plásmido L", se insertó ADN simulado (~3 kb) en lugar de la región que codificaba la cadena ligera, mientras que en el "plásmido H", el ADN simulado (aproximadamente 2 kb) estaba presente en lugar de la región que codificaba la cadena pesada. El ADN simulado en el "plásmido L" fue reemplazado en dos etapas: el gen rCk (véase más arriba) se insertó en el vector utilizando BsiWI/SphI, y después los genes maestros de rVL kappa (véase más arriba) se insertaron en el vector utilizando EcoRV/BsiWI. El ADN simulado en el "plásmido H" fue reemplazado también en dos etapas: el gen rCH1 (véase más arriba) se insertó en el vector utilizando StyI (o Eco130I)/EcoRI, y luego se insertaron los genes maestros de rVH (véase más arriba) en el vector mediante el uso de MfeI/StyI. Dos fragmentos de ADN que codificaban los genes maestros rVH1 o rVH5 (ambos ligados a la región mCH1) se escindieron del "plásmido H", y se insertaron mediante el uso de SphI/HindIII en un panel de plásmidos "L", que son plásmidos que codifican una variedad de genes maestros de rVL dando como resultado 4 combinaciones de rVL-rVH. El "plásmido L" tricistrónico que

codificaba también gpIII, rVL/rCk y rVH/rCH1 se denominó pMorph32, indicando el hecho que el plásmido codificaba dominios variables y constantes de rata.

Reclonación de combinaciones de genes maestros de VL-VH de rata en el vector bicistrónico pMorphx9_Fab_FH para la expresión en E. coli

5 Como se ha descrito anteriormente, con el fin de analizar la expresión soluble de las combinaciones de la región estructural VL-VH en formato Fab, clonamos de nuevo, utilizando XbaI y EcoRI, las cuatro combinaciones de genes maestros de VL-VH (como insertos que codifican Fab) en el vector pMorphx9_Fab_FH (Rauchenberger et al., 2003, J. Biol. Chem. 278, 38194-38205), en donde FH indica que los marcadores Flag y hexa-His se fijaron en el extremo C-terminal del Fab.

10 *Conversión de combinaciones de genes maestros de rVL-rVH en formato IgG2b y 2c de rata*

Con el fin de analizar la expresión de las combinaciones de genes maestros en los formatos kappa IgG2b y 2c de longitud completa en células de mamífero, subclonamos las 4 combinaciones de genes maestros de rVL-rVH en el vector pMorph2_h/r_IgG que es un derivado del vector pMorph2_h_IgG (Steidl et al. 2008. Mol Immunol 46(1):135-44). La transcripción estaba dirigida en este plásmido por el promotor CMV.

15 *Expresión y purificación de Fab*

Como se ha descrito en el caso de ratones, la expresión en células TG1 de *E. coli* a partir del plásmido pMx9_Fab_FH se llevó a cabo en 500 ml de medio 2xYT complementado con cloranfenicol (34 mg/ml) y glucosa (0,1%). Después de la inducción con IPTG (0,75 mM), las células se cultivaron a 30°C durante 20 h. Los sedimentos celulares se lisaron durante la incubación con tampón de lisis (fosfato de sodio 200 mM, pH 7,4, NaCl 0,5 M, imidazol 10 mM, 0,2% de lisozima, MgCl₂ 2 mM, 20 U/ml de benzonasa y cóctel inhibidor de proteasa exento de EDTA de Roche), durante 30 min a TA. La purificación de Fab a partir de lisados celulares aclarados, se llevó a cabo mediante IMAC con iones Ni²⁺. Después de la elución, el tampón que contenía imidazol se reemplazó por PBS sobre las columnas de desalación PD10 (Amersham Pharmacia Biotech).

Expresión y purificación de IgG

25 Como se ha descrito anteriormente, las células eucariotas HKB11 se transfectaron de forma transitoria con el plásmido pMorph4_h/m_IgG que codificaba las cadenas pesadas y ligeras. El material sobrenadante del cultivo celular se recogió 3 a 7 días después de la transfección. Después de ajustar el pH del sobrenadante a 8,0, de añadir NaCl (concentración final 2 M) y filtrar de forma estéril (0,45 µm), la solución se sometió a una cromatografía de afinidad con proteína A estándar (MabSelect SURE, GE Healthcare). Si no se indica lo contrario, el intercambio de tampón se realizó con 1 x PBS de Dulbecco (pH 7,2, Invitrogen) y las muestras se filtraron de forma estéril (0,2 µm). La pureza de la IgG se analizó en condiciones de desnaturalización mediante SDS-PAGE o mediante el uso de Agilent BioAnalyzer y en estado natural mediante HP-SEC.

Preparación de fagos

35 La amplificación y la purificación de los fagos se llevó a cabo esencialmente tal y como se ha descrito anteriormente (Krebs et al., 2001, J. Immunol Methods 254, 67-84) con la siguiente modificación. Los fagémidos (combinaciones de la región estructural VL-VH codificada en el plásmido de presentación) se propagaron en células Top10F+ de *E. coli* en medio 2xYT/Cam/Tet/glucosa (1%). Después de la infección con fago auxiliar (Hyperphage de Progen, nº de Cat.: PRHYPE, multiplicidad de la infección: 40), la centrifugación y la resuspensión del sedimento celular en 2xYT/Cam/Kan/Tet/IPTG (0,25 mM), los fagos se produjeron mediante incubación durante 20 h a 22°C.

40 *ELISA de tipo sándwich con presentación en fagos*

Las placas de microtitulación negras de Maxisorp (Nunc) se recubrieron con anticuerpo anti-pIII (MoBiTec, nº de Cat.: PSKAN3) para la determinación del nivel relativo de pIII (≈ concentración de fagos), o con anticuerpo anti-F(ab')₂ de rata (Jackson ImmunoResearch Lab.; nº de Cat.: 112-006-072) para la determinación del nivel relativo de Fab en las muestras de fago. Aunque el epítipo exacto del anticuerpo anti-F(ab')₂ de rata no se conoce (se supone que se encuentra en el dominio constante), se razonó que la unión del anticuerpo a Fab de rata no es probable que esté muy afectada por una variación de secuencia en el dominio Fv y, por lo tanto, el anticuerpo debe ser adecuado para la evaluación general de la presentación de Fab de rata en la punta del fago filamentosos. TBS pH 7,4 que contenía Tween 20 (0,05%) y chemi-bloquer (de Chemicon; diluido 1:2) se utilizó para el bloqueo de los sitios de unión no específica en la placa y en los fagos. Después de incubar con fagos, los pocillos se lavaron con TBS-T, y se incubaron secuencialmente con anticuerpo anti-bacteriófago biotinilado (Sigma nº de Cat. B2661) y avidina-HRP (BD Pharmingen; nº de Cat.: 554058). Se aplicó el sustrato "Quanta Blu" para HRP, y se registraron las señales de fluorescencia en un aparato Tecan-Spectrafluor (excitación: 320 nm, emisión: 430 nm). Los niveles relativos de pIII y de Fab en la muestra de fago se determinaron mediante el uso de un fago de referencia (que presenta un Fab de HuCAL seleccionado previamente convertido en formato de rata) y el modelo de 2 estados estándar aplicado para la regresión no lineal. La tasa de presentación relativa se definió como la relación entre los niveles de Fab y de pIII.

Ejemplo 2.4: Caracterización de la genoteca de anticuerpos de rata por medio del anticuerpo 4D5 ejemplar

Pensamos que una buena presentación en fagos y una expresión elevada en las células se puede lograr solo para moléculas Fab estables y plegadas de forma adecuada. Se sabe que la estabilidad termodinámica de una molécula Fab depende de la estabilidad intrínseca de los dominios individuales (cadenas pesada y ligera), así como de la interacción no covalente entre los dos dominios (Ewert et al., 2003, J Mol Biol 325, 531-553). Una interacción estrecha entre Fd (cadena pesada) y Ck (cadena ligera) tiene una importancia particular porque en el sistema CysDisplay® los dos dominios no pueden estar unidos covalentemente a través de un enlace disulfuro, ya que el residuo de cisteína en los extremos C-terminales de los dominios constantes se ha eliminado. Por lo tanto, en lugar de someter a ensayo individualmente las regiones estructurales de VL y VH, intentamos identificar las combinaciones de VL-VH que se podían plegar de manera eficaz y expresar en *E. coli* (formato Fab) y en células de mamífero (formato IgG).

Como caso tipo se utilizó la secuencia HCDR3 del anticuerpo 4D5 (WGGDGFYAMDY). Esta secuencia ya se había utilizado previamente para determinar genes maestros humanos antes de la construcción de la genoteca HuCAL original (Knappik et al., 2000, J Mol Biol 296, 57-86).

Análisis de la eficacia de la presentación

Puesto que una presentación eficaz de los miembros de la genoteca sobre partículas de fago es un requisito previo para el éxito de las selecciones para una presentación en fagos, decidimos utilizar una cuantificación de ELISA de tipo sándwich en fagos, de la presentación de las cuatro combinaciones de VL-VH de rata en formato Fab. Nuestros datos indican que las cuatro combinaciones de regiones estructurales de rVL-rVH sometidas a ensayo, se presentaban claramente en los fagos (Figura 38).

Expresión de Fab soluble

También la expresión soluble de las cuatro combinaciones de VL-VH de rata en formato Fab fue sometida a ensayo en *E. coli*. Encontramos que las combinaciones de VL1 con VH1 o VH5 se expresaban sorprendentemente con niveles 3-4 veces más elevados que las combinaciones de VL2 (Figura 39). El nivel de expresión de VL1-VH5 era notablemente alto (aproximadamente 15 mg/L), y nos condujo finalmente a nuestra decisión de elegir esta Fv como región estructural para la construcción de la genoteca de rata.

Expresión de IgG soluble

Debido al hecho de que, en la actualidad, el formato de anticuerpo con mayor éxito aprobado para la terapia humana es IgG, decidimos caracterizar la expresión de las cuatro combinaciones de genes maestros de VL-VH de rata también en formato IgG. Dos formatos IgG de rata distintos, IgG2b y 2c, se eligieron para el ensayo de la expresión. Aunque la cadena pesada gamma de rata $\gamma 1$ y $\gamma 2a$ comparten 94% de homología de secuencia, y se parecen mucho a $\gamma 1$ de ratón, $\gamma 2b$ de rata es equivalente a la pareja $\gamma 2a/\gamma 2b$ de ratón en lo que respecta tanto a la secuencia de nucleótidos como a las funciones efectoras del anticuerpo (fuerte activación del complemento y unión a Fc γ R) mientras $\gamma 2c$ de rata se parece a $\gamma 3$ de ratón (homología de secuencia elevada) y activa débilmente el complemento (Brüggenmann, 1988, Gene 74, 473-482). IgG2b de rata se une débilmente a las proteínas A y G con una fuerza iónica baja (fisiológica). Por el contrario, IgG2c de rata se une fuertemente a las proteínas A y G; sin embargo, es poco soluble en soluciones de fuerza iónica baja (Rousseaux y Bazin, 1979, Vet Immunol Immunopathol.1, 61-8).

En la práctica, los genes maestros de VL y VH se volvieron a clonar en vectores de expresión de mamífero, y los plásmidos resultantes se usaron para la transfección transitoria de una línea celular de mamífero en suspensión tal como se ha descrito en Materiales y Métodos. No observamos grandes diferencias a nivel de expresión de las cuatro combinaciones de rata en formato IgG2b de rata; sin embargo, las combinaciones de VH5 parecían expresarse aproximadamente 5 veces mejor que las combinaciones de VH1 en formato IgG2c de rata, en donde VL2-VH5 se expresaba ligeramente mejor que VL1-VH5. Los resultados para IgG2b se muestran en la Figura 40, los resultados para IgG2c en la Figura 41. Los niveles de expresión de las combinaciones de VL-VH de rata sometidas a ensayo en formatos Fab e IgG no se correlacionaron.

Todas las combinaciones de genes maestros de VL y VH de rata se purificaron en un proceso de purificación de 2 etapas usando Äkta express. Los fragmentos Fab expresados se unieron a Ni²⁺ quelado y se lavaron con 15 volúmenes de columna de tampón de fosfato de Na 20 mM/NaCl 500 mM, pH 7,4, imidazol 20 mM o, en una purificación distinta, utilizando 20 volúmenes de columna de isopropanol al 30%/fosfato de Na 20 mM pH 7,4. La segunda etapa de la purificación (SEC) era idéntica en ambos procedimientos (con o sin tratamiento con isopropanol) y se ha descrito anteriormente en el presente documento. Después de la purificación, se compararon ambas preparaciones de proteína en relación con el rendimiento, la concentración, la pureza, la proporción de VH y VL y la estabilidad térmica. Se introdujo una etapa de lavado con isopropanol al 30%, mientras que el Fab todavía está unido a la columna de IMAC, con el fin de recibir fragmentos Fab que tenían mayor pureza. Sin embargo, para algunos fragmentos Fab, este tratamiento parecía ser demasiado fuerte porque la interfaz - la interacción hidrofóbica entre la cadena pesada y ligera - no era lo suficientemente fuerte como para soportar este procedimiento. Las cadenas pesada y ligera de tales fragmentos Fab se eluyeron en diferentes etapas y, por lo tanto, la relación entre las dos cadenas se desvió de 1:1. La Figura 42 muestra un resumen de los rendimientos de la expresión de todas las combinaciones de la región estructural de Fab de rata sometidas a ensayo después de lavar en IMAC con y sin isopropanol con un código de

color para criterios de selección.

Estabilidad térmica

5 Como se ha descrito para la genoteca murina, la estabilidad de la proteína y la resistencia al estrés ambiental (tal como proteasas, agentes térmicos o caotrópicos) son una propiedad intrínseca de la proteína y son típicamente independientes de las condiciones de expresión. Los valores de T_m (todos dentro del intervalo 67-75°C) mostrados en la Figura 43, indican que la estabilidad térmica de cuatro variantes de VL-VH de rata sometidas a ensayo en los formatos IgG2b y IgG2c, no varió; sin embargo, la estabilidad de las combinaciones de VL2 de rata en formato Fab (T_m aproximadamente 48°C) era inferior a la estabilidad de las combinaciones de VL1 de rata (T_m aproximadamente 65°C).

10 *Selección de regiones estructurales de Fv para la construcción de una genoteca*

Teniendo en cuenta la escasez de secuencias disponibles de anticuerpos de rata reordenados en las bases de datos, lo que dificulta la evaluación estadística sólida de la prevalencia de la secuencia de la línea germinal, optamos por la construcción de una genoteca combinatoria de fagémidos de rata usando una región estructural Fv única resistente. Seleccionamos la combinación VL1-VH5 debido a que: (i) tenía la mayor presentación en formato Fab en la punta de fagos filamentosos (véase la Figura 35); (ii) se expresaba claramente mejor en formato Fab soluble en *E. coli* (aproximadamente 17 mg/L); (iii) su nivel de expresión en los formatos IgG sometidos a ensayo era bueno (> 10 mg/L); y (iv) su estabilidad térmica era alta (T_m de aproximadamente 70°C). Casetes diversificados de genotecas de LCDR3 y HCDR3 se clonaron en esta región estructural Fv seleccionada, tal y como se describe a continuación.

Ejemplo 2.5: Diversificación de la genoteca de rata

20 *Construcción del casete de la mini-genoteca de HCDR2 utilizando cebadores con posiciones oscilantes*

Se ha observado (véase <http://www.bioc.uzh.ch/antibody/Structures/AgContact/index.html>) que los residuos más importantes en HCDR2 implicados en el contacto con el antígeno se encuentran en las posiciones 52, 52a, 54, 56 y 58 (numeración de Kabat). Una variación en estas posiciones permite la selección natural *in vivo* de anticuerpos con el sitio de unión al antígeno adaptado que comprende HCDR2. Investigamos la distribución de los aminoácidos para familias de secuencias de la línea germinal de anticuerpos reordenados, representados por nuestros tres genes maestros seleccionados de rVH (rVH1_RV, rVH2 y rVH5), recopilamos una lista de los residuos que se encontraban con frecuencia en posiciones particulares de HCDR2 e introducimos esas variaciones en los genes maestros seleccionados. Con el fin de conservar la prevalencia de la secuencia de la línea germinal, los residuos alternativos presentes en los anticuerpos reordenados, pero no en las secuencias de la línea germinal, se infrarrepresentaron intencionadamente en el diseño (véase la Figura 44). El análisis de la secuencia de aproximadamente 30 clones escogidos al azar por cada región estructural, indicaba que las variantes deseadas (cinco por región estructural) estaban representadas en la HCDR2 con una frecuencia similar a nuestro diseño; sin embargo, aproximadamente 10% de los clones recogidos al azar era portador de deleciones de nucleótidos no deseadas, principalmente en las posiciones que habían oscilado en los cebadores de la PCR (datos no mostrados). Con el fin de asegurar poca abundancia de clones con el marco desplazado en la genoteca, decidimos diluir las variaciones que codificaban la agrupación de plásmidos en HCDR2 con plásmido que codificaba la región estructural VH de tipo silvestre. Por lo tanto en la genoteca final, las variantes menores en HCDR2 son aproximadamente 10 veces menos abundantes que en nuestro diseño inicial de HCDR2.

40 La variabilidad prevista (cinco variantes de la secuencia de aminoácidos en el gen maestro) fue introducida mediante amplificación del ADN en 10 ciclos de PCR con ADN polimerasa Pfx Platinum (Invitrogen) y un conjunto de cebadores directos e indirectos. La variabilidad en rVH5 se introdujo usando el cebador directo que contenía posiciones oscilantes con una relación de nucleótidos definida:

ACCAAAGGTCTCGAGTGGGTGGCGACCATTAG(70%)/C(30%)CTATGATGGCAGCAG(6
0%)/A(20%)/C(20%)CACCTATTATCGTGATAGC.

45 El producto de la PCR resultante se digirió con XhoI/BssHII, y se clonó en el gen maestro de rVH5 en el plásmido pMorph32. La agrupación de plásmidos que contenía la variedad introducida en el gen maestro, se diluyó 1:10 con el plásmido que codificaba el gen maestro original, y se utilizó para la introducción del casete de la genoteca LCDR3.

Construcción del casete de la genoteca LCDR3

50 El relleno MBP para VL1 de rata es idéntico a mVLB1 y los casetes de la genoteca LCDR3 para mVLB1 y rVL1 son idénticos. El casete de la genoteca LCDR3 se construyó como se ha descrito anteriormente en el caso del casete de la genoteca de ratón. El diseño en sí era idéntico al de LCDR3 murino, mientras que las regiones flanqueantes difieren, para reflejar la región estructural natural de la LCDR3:

Estas regiones flanqueantes, representan una parte de la región estructural 3 del anticuerpo y de la región estructural 4 terminal de VL del anticuerpo. Por lo tanto era posible amplificar con PCR y clonar el casete en una

etapa posterior.

rVk1

5' GAAGACGTGGCGACCTATTTTTGC-T8-CAG-T9-T10-T11-T12-CCG-T13
ACCTTTGGTGGCGGGCACCA

en donde T8 - T13 son mezclas TRIM.

- 5 Las secuencias de todos los oligonucleótidos utilizados para la preparación del casete de LCDR3 se enumeran en la Tabla

Tabla 37:

Cebador directo específico de LCDR3 de rata		
	nombre	secuencia
rVk2	Lib67	CTGCAACCGGAAGACGTGGCGACCTATTTTTGC
Oligos solapantes		
rata	Lib70	CGGCCACGTACGTTTGATTTCCAGTTTGGTGCCACCGCCAAAGG
Cebador Vk inv		
rata	Lib68	CGGCCACGTACGTTTGATTC

Construcción del casete de la genoteca de HCDR3

- 10 El casete de la genoteca de HCDR3 se construyó también exactamente como se ha descrito anteriormente en el caso del casete de la genoteca de ratón. El diseño del casete HCDR3 es idéntico al diseño desarrollado para la HCDR3 murina. Por lo tanto, toda la información acerca de la HCDR3-TRIM y los oligos para la amplificación con PCR, se puede encontrar en el capítulo correspondiente para la genoteca murina.

Ejemplo 2.6: Prueba de inmunoadsorción con la genoteca de rata

- 15 La genoteca de rata que consistía en una combinación de la región estructural rVL1_rVH5 con diversificaciones en L-CDR3, H-CDR3 y H-CDR2, se utilizó para generar anticuerpos específicos para GM-CSF de ratón.

Los fagos de la genoteca de rata se sometieron a tres rondas de selección sobre GM-CSF murino (AbD Serotec, Alemania). El GM-CSF se biotiniló de acuerdo con las instrucciones del fabricante (Amersham, GE Healthcare). El GM-CSF murino biotinilado se denominó en lo sucesivo "Bio-GM-CSF". Los fagos se bloquearon previamente con 2x ChemiBLOCKER que contenía 0,1% de Tween, durante dos horas y, posteriormente, se adsorbieron previamente sobre placas de NeutrAvidin® no recubiertas durante una hora adicional, a temperatura ambiente. Esta etapa se incluye para eliminar los fagos específicos de NeutrAvidin que no es diana. El material sobrenadante de fagos de la genoteca de rata aclarada previamente, obtenida como se ha descrito anteriormente, se incubó con Bio-GM-CSF 200 nM durante 1 hora a temperatura ambiente en un rotor. Los fagos unidos a Bio-GM-CSF fueron capturados sobre pocillos recubiertos con NeutrAvidin (tiras de NeutrAvidin; Pierce) durante 30 min y posteriormente los fagos no unidos se eliminaron por lavado tal y como se indica en la Tabla 38. Los fagos unidos específicamente se eluyeron mediante la adición de 100 µl de DTT 25 mM en Tris HCl 10 mM, pH 8,0 a cada pocillo e incubando durante 10-15 min a TA con agitación. Se recogieron los eluidos de DTT de cada selección. TG1 de *E. coli* con una DO600nm de 0,6-0,8 se añadió a los eluidos de fago de cada selección y se incubó durante 30 min en un baño de agua. Después las bacterias para infección se sembraron de manera uniforme en dos grandes placas de agar LB/Cm/Glu para cada selección y se incubaron durante una noche a 37°C. Para las siguientes rondas de inmunoadsorción, se recogieron y se usaron suspensiones bacterianas de cada agrupación para propagar los fagos durante una ronda adicional de inmunoadsorción, tal y como se ha descrito anteriormente. En total, se realizaron tres rondas posteriores de selección de fagos con un incremento del rigor del lavado (véase la Tabla 38)

- 35 Tabla 38: Condiciones de lavado aplicadas en la inmunoadsorción de fagos

1ª Ronda	2ª Ronda	3ª Ronda
3x PBS 0,05% Tween, rápido	1 x PBS 0,05% Tween, rápido	10x PBS 0,05% Tween, rápido
2x PBS 0,05% Tween, 5 min de agitación	4x PBS 0,05% Tween, 5 min de agitación	5x PBS 0,05% Tween, 5 min de agitación

1ª Ronda	2ª Ronda	3ª Ronda
3x PBS, rápido	1 x PBS, rápido	10x PBS, rápido
2x PBS, 5 min de agitación	4x PBS, 5 min de agitación	5x PBS, 5 min de agitación

5 La subclonación de los fagos seleccionados, microexpresión de fragmentos Fab se realizaron tal y como se ha descrito para la genoteca de ratón, más arriba en el presente documento. El escrutinio primario en ELISA se realizó en solución con el antígeno biotinilado en modo captura. Esencialmente 20 µl de anticuerpo anti-rata (IgG de cabra anti-rata, F(ab')₂, Jackson ImmunoResearch Laboratories, Inc) se inmovilizaron en placas Maxisorb a una concentración de 1,5 µg/ml en PBS, se incubaron O/N a 4°C.

10 El bloqueo de los lisados BEL de rata se realizó en 3% de BSA/TBST, 30-60 min a 22°C. Los fragmentos Fab de rata fueron capturados a partir de lisado crudo bacteriano (BEL) mediante incubación durante 1 hora a TA. Después de lavar 5 veces con TBST, se añadieron 20 µl de mGM-CSF biotinilado (diluido hasta 0,75 µg/ml en 0,5% de BSA TBST) a las placas y se incubaron durante 0,5 h. Después de lavar 5 veces con TBST, el mGM-CSF biotinilado capturado a través de Fab específico de rata, se detectó mediante la adición de 20 µl de avidina-HRP diluida 1:5000 en 0,5% de BSA TBST a TA durante 1 h. Después de lavar 5 veces adicionalmente con TBST, se añadieron 20 µl de sustrato quanta blue para las placas de escrutinio, utilizando MultiDrop384. Después de 5-10 minutos, se midió la fluorescencia con GENios Pro (Tecan). Las señales 5-10 veces más fuertes que el ruido de fondo fueron consideradas como una unión específica de mGM-CSF a los fragmentos Fab.

15 La Tabla 39 resume el resultado del escrutinio de la inmuoadsorción de mGM-CSF. Se escrutaron 262 clones en ELISA sobre bio-mGM-CSF capturado, tal y como se ha descrito anteriormente. Se secuenció VH de 52 clones que mostraban 5 veces una señal más clara que el ruido fondo, para identificar fragmentos Fab únicos. De esos 51 clones secuenciados se identificaron cinco secuencias Fab únicas, que mostraban diversidad de secuencia en la secuencia de H-CDR3. La H-CDR1 y la H-CDR2 no muestran variación de la secuencia y son idénticas a la secuencia de rata de la línea germinal de VH5.

Tabla 39: Resumen de la selección de la genoteca HuCAL de rata frente a GM-CSF de ratón

Subgenoteca	Nº escrutado	Aciertos primarios	Tasa de aciertos	Nº secuenciado	Secuencia única
rVH-Vk1	262	51	19%	51	5

25 Estos cinco fragmentos de Fab específicos se expresaron y purificaron como se ha descrito en las secciones de ratón descritas anteriormente en el presente documento. Las secuencias de CDR de la cadena pesada se muestran en la Tabla 40. Se resumen las características de expresión y biofísicas en la Tabla 41. Los cinco clones se expresan bien en *E. coli* con un rendimiento de la expresión en un intervalo de 5 a 15 mg/L. En una columna de cromatografía de exclusión por tamaño, los cinco fragmentos Fab mostraron un perfil de elución monómero, lo que indica que el fragmento de rata no tiene tendencia hacia un plegamiento erróneo o una agregación de anticuerpos. La estabilidad térmica se evaluó mediante el método descrito anteriormente en la sección "Estabilidad térmica" de la genoteca de ratón en el presente documento. Excepto el clon mGM-CSF_rat2, los cuatro otros fragmentos Fab de rata muestran una temperatura T_m aparente sin plegamiento, en el intervalo de 68,5°C hasta 72°C, lo que refleja la temperatura de desnaturalización de la región estructural rVL1_VH5 de la genoteca.

35 Tabla 40: Secuencia de la región determinante de complementariedad de la cadena pesada de anticuerpos específicos de GM-CSF de ratón

Clon	H-CDR1	H-CDR2	H-CDR3
mGM-CSF_rat_1	FTFSNYGMAWVRQAPTKGLE	TISYDGSSTYYRDSVKG	DSWGYRKRKFDY
mGM-CSF_rat_2	FTFSNYGMAWVRQAPTKGLE	TISYDGSSTYYRDSVKG	GYTGSFD
mGM-CSF_rat_3	FTFSNYGMAWVRQAPTKGLE	TISYDGSSTYYRDSVKG	NDYGGWRYKFDY
mGM-CSF_rat_4	FTFSNYGMAWVRQAPTKGLE	TISYDGSSTYYRDSVKG	SDWGGMRYKMDY
mGM-CSF_rat_5	FTFSNYGMAWVRQAPTKGLE	TISYDGSSTYYRDSVKG	YGWYKGRYKFDY

Tabla 41: Expresión y datos de la desnaturalización térmica de anticuerpos de rata específicos de mGM-CSF

Expresión y estabilidad

Clon Combinación de la región estructural Tasa de expresión mg/l Porción de monómero [%] T_m [°C]

mGM-CSF_rat_1	rVH5-K1	5	100	70
mGM-CSF_rat_2	rVH5-K1	7	99	55
mGM-CSF_rat_3	rVH5-K1	15,5	100	72
mGM-CSF_rat_4	rVH5-K1	12	100	71,5
mGM-CSF_rat_5	rVH5-K1	9,5	100	68,5

La especificidad de cuatro de los fragmentos Fab de rata para GM-CSF se demostró en un ELISA de especificidad (véase la Figura 41). Los cuatro anticuerpos de rata proporcionaron una señal muy específica sobre GM-CSF inmovilizado directamente sobre placas Maxisorp, pero ninguna señal con el antígeno no relacionado Estradiol-BSA.

- 5 Estos ejemplos ponen de relieve el hecho de que ambas, la región estructural, así como las CDRs diversificadas, contribuyen a la estabilidad general de una molécula de anticuerpo. Mediante una selección cuidadosa de regiones estructurales muy estables del anticuerpo y un análisis sofisticado y la síntesis de nuevo de CDR, se pueden construir sintéticamente genotecas de anticuerpos que muestran un comportamiento superior, con respecto a la expresión y las características biofísicas y unas características de unión bien definidas.

10

REIVINDICACIONES

- 5 1. Una genoteca de anticuerpos murinos completamente sintética, que comprende regiones estructurales codificadas por los genes de la línea germinal VH IGHV1-72*01 (VH1), IGHV2-2*01 (VH2) y IGHV5-9*04 (VH5) y regiones estructurales codificadas por al menos dos de los siguientes genes de la línea germinal VL-kappa: IGKV1-117*01 (Vk1), IGKV3-12*01 (Vk3) y IGHV3-4*01 (Vk3).
2. La genoteca según la reivindicación 1, en donde dicha genoteca comprende más del 50% del repertorio natural de VH de ratón.
3. La genoteca según cualquiera de las reivindicaciones precedentes, en donde dicha genoteca comprende más del 50% del repertorio natural de VL de ratón.
- 10 4. La genoteca según cualquiera de las reivindicaciones precedentes, en donde dicha genoteca tiene una diversidad de al menos $1,0 \cdot 10^8$ en la región H-CDR3.
5. La genoteca según cualquiera de las reivindicaciones precedentes, en donde esencialmente todas las combinaciones de VH-VL de dicha genoteca se presentan de forma eficaz.
- 15 6. La genoteca según cualquiera de las reivindicaciones precedentes, en donde todas las combinaciones de VH-VL son térmicamente estables.
7. La genoteca según cualquiera de las reivindicaciones precedentes, en donde las regiones H-CDR3 de esencialmente todos los miembros de la genoteca están flanqueadas por sitios de restricción únicos.
- 20 8. La genoteca según cualquiera de las reivindicaciones precedentes, en donde dicha genoteca comprende H-CDR3s que incluyen más del 50% de las longitudes de H-CDR3 presentes en la naturaleza del repertorio de H-CDR3 murino.
- 25 9. La genoteca según cualquiera de las reivindicaciones precedentes, en donde dicha genoteca comprende regiones estructurales codificadas por uno o varios de los siguientes genes de la línea germinal modificados: IGHV1-72*01 con una mutación Q1E; IGHV1-72*01 con una mutación Q1E, una mutación K77R y una mutación A78V; IGHV2-2*01 con una mutación Q1E; IGHV3-6*01 con una mutación D1E y una mutación T26S; IGHV5-9*04 con una mutación K3Q; IGHV9-3*01 con una mutación Q1E, una mutación I2V y una mutación K53E y IGKV1-117*01 con una mutación V2I.
10. Una colección de moléculas de ácido nucleico que codifica los miembros de la genoteca según cualquiera de las reivindicaciones precedentes.
11. Una genoteca de vectores que codifica las moléculas de ácido nucleico según la reivindicación 10.
- 30 12. Una genoteca recombinante de células hospedadoras que comprende las moléculas de ácido nucleico según la reivindicación 10 o el vector según la reivindicación 11.
13. Un método para aislar un anticuerpo específico de un antígeno, comprendiendo dicho método las etapas de:
 - (a) poner en contacto la genoteca según una cualquiera de las reivindicaciones 1-9 con un antígeno;
 - 35 (b) eliminar aquellos miembros de la genoteca que no se unen al antígeno; y
 - (c) recuperar aquellos miembros de la genoteca unidos al antígeno.

Figura 1

		FMS																						
77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101
K	A	T	L	T	V	D	K	P	S	S	T	A	Y	M	Q	L	S	S	L	T	S	E	D	S
R	V	T	L	T	V	D	K	P	S	S	T	A	Y	M	Q	L	S	S	L	T	S	E	D	S

mVH1
mVH1_RV

puente salino estabilizante

Figura 2

<i>JH de ratón</i>		FW4										<i>sitio Blpl</i>		
		W	G	A	G	T	T	V	T	V	T	V	S	S
JH1	Y	W	Y	F	D	V	W	G	A	G	T	T	V	S
JH2			Y	F	D	Y	W	G	Q	G	T	T	L	S
JH3		A	W	F	A	Y	W	G	Q	G	T	L	V	S
JH4	Y	Y	A	M	D	Y	W	G	Q	G	T	S	V	S

Figura 3

Uso	JK de ratón	W	T	F	G	G	G	G	T	K	L	E	I	K	R	A	no codificado por JK
45%	JK1	W	T	F	G <td>G <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td></td>	G <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td>	G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td>	G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td>	T	K	L	E	I	K	R	A	A
40%	JK2	Y	T	F	G <td>G <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td></td>	G <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td>	G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td>	G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td>	T	K	L	E	I	K	R	A	A
	JK3	I	T	F	S <td>D <td>G <td>G <td>T</td> <td>R</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>P</td> <td>A</td> <td>A</td> </td></td></td>	D <td>G <td>G <td>T</td> <td>R</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>P</td> <td>A</td> <td>A</td> </td></td>	G <td>G <td>T</td> <td>R</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>P</td> <td>A</td> <td>A</td> </td>	G <td>T</td> <td>R</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>P</td> <td>A</td> <td>A</td>	T	R	L	E	I	K	P	A	A
	JK4	F	T	F	G <td>S <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td></td>	S <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td>	G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td>	G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td>	T	K	L	E	I	K	R	A	A
	JK5	L	T	F	G <td>A <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>L</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td></td>	A <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>L</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td></td>	G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>L</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td> </td>	G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>L</td> <td>K</td> <td>R</td> <td>A</td> <td>A</td>	T	K	L	E	L	K	R	A	A
	FW4 de ratón HuCAL	?	T	F	G <td>G <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>T</td> <td>T</td> </td></td></td>	G <td>G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>T</td> <td>T</td> </td></td>	G <td>G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>T</td> <td>T</td> </td>	G <td>T</td> <td>K</td> <td>L</td> <td>E</td> <td>I</td> <td>K</td> <td>R</td> <td>T</td> <td>T</td>	T	K	L	E	I	K	R	T	T

Sitio BsIWI

Figura 4a

	95	96	97	98	99	100	a	b	m	n	101	102	FW4
7 sec													
446 sec	27% G 11% D 11% S 10% Y 7% E 6% L 5% R	25% G 15% Y 12% D 7% T 7% S 7% N 5% R	33% Y 12% G 9% D 9% T 6% S 5% W	22% G 21% Y 14% A 8% W 6% D 5% R	32% Y 19% A 8% G 5% W 5% S					54% F 17% M	62% D 26% A	90% Y	
645 sec													
645 sec	18% G 16% Y 10% S 9% D 9% R 6% E 5% H 5% W	26% G 17% Y 11% D 7% R 7% L 5% T 4% S 4% N	23% Y 16% G 11% N 9% D 6% R 6% T 6% S	25% Y 16% G 10% D 9% W 8% S 5% T	32% Y 19% A 8% G 5% W 5% S					64% F 19% M	76% D 18% A	86% Y 9% V	
1705 sec													
1705 sec	44% Y 22% G 6% S	42% Y 17% G 6% D 5% R	56% Y 7% N 7% G	52% G 16% Y 6% S	20% S 19% Y 13% R 11% G	39% Y 13% H 12% A 10% S				75% F 13% M 6% L	87% D 8% A 1% G	75% Y 16% V 5% F	

Figura 4b

10 ab	36% Y 12% G 10% R 7% D	34% Y 19% G 11% D 5% N	46% Y 14% G 9% L 5% S	40% G 15% Y 7% D 7% S	30% S 18% Y 13% G 9% R	30% S 23% Y 10% G 7% W	38% Y 19% A 8% W 7% H	71% F 20% M	82% D 14% A	86% Y 10% V	
11 ab	33% Y 13% G 8% S 7% D 7% R 5% N	31% Y 15% D 15% G 5% E 5% R	46% Y 11% G 8% R 6% L	32% Y 28% G 5% S 5% D 5% R	21% G 20% Y 14% S 8% R 7% P 6% N 5% T	30% Y 21% S 15% G 5% D	34% Y 16% S 10% W 9% F 6% G	38% Y 30% A 7% G 7% W 5% S	61% F 31% M	85% D 13% A	86% Y 11% V
12 ab	47% Y 8% G 7% D 7% R 7% S	38% N 15% Y 14% G 5% D	68% Y 6% G	48% G 22% Y 5% D 5% L	52% G 14% Y 11% S 7% R	51% S 15% G 8% Y 5% D 5% L	37% T 18% Y 16% S 6% R 5% G	59% Y 10% W 6% S 5% R	67% Y 16% A	92% D 6% A	85% Y 12% V

(numeración según Kabat)

Ilustración reducida a aminoácidos con una aparición >=5%.

Figura 5

FWS		95	96	97	98	99	100	a	b	m	n	101	102	FWA	
		32% Y	31% Y	47% Y	39% G	21% S	24% S	27% Y	27% Y	38% Y	68% F	83% D	82% Y		
		17% G	19% G	11% G	22% Y	20% G	22% Y	23% S	18% S	19% A	20% M	13% A	12% V		
		8% S	10% D	6% L	6% D	19% Y	18% G	10% G	10% G	6% G	4% L	1% G	2% F		
		7% D	5% N	5% D	6% S	9% R	6% D	8% R	8% W	6% H	1% E	1% P	1% H		
		7% R	5% R	5% N	4% L	5% D	5% R	6% D	7% R	6% S	1% G		1% S		
		4% E	4% L	4% R	4% N	4% L	4% L	6% T	5% D	6% W	1% S				
		4% H	4% S	4% S	4% R	4% T	3% N	4% L	4% A	3% L	1% V				
		3% N	3% A	3% T	4% T	3% L	3% P	3% H	4% F	2% D					
		3% W	3% E	2% A	2% A	3% P	3% T	3% P	4% T	2% F	1% Y				
		2% A	3% T	2% F	2% F	2% A	3% V	2% A	3% N	2% P					
		2% K	2% F	2% I	2% P	2% F	2% A	2% N	2% L	2% R					
		2% L	2% P	2% P	1% E	2% V	2% H	2% V	2% P	2% T					
		2% P	2% V	2% V	1% I	1% E	1% C	1% E	2% V	2% V					
		1% F	2% W	2% W	1% K	1% H	1% E	1% F	1% E	1% C					
		1% I	1% H	1% E	1% V	1% I	1% F	1% I	1% H	1% I					
		1% Q	1% I	1% H	1% W	1% K	1% I	1% K	1% I	1% N					
		1% T	1% K	1% Q	1% W	1% Q	1% K	1% W							
		1% V	1% Q	1% Q	1% W	1% W	1% W								

(numeración según Kabat)

Figura 6

FW/S	95	96	97	98	99	100	a	b	m	n	101	102	FW/F
	20%Y	20%Y	40%Y	15%Y	15%Y	15%Y	15%Y	30%Y	30%Y	60%F	80%D	80%Y	
	20%G	20%G	10%G	15%G	15%G	15%G	15%G	10%S	20%A	20%M	20%A	20%V	
	10%D	10%D	5%D	15%S	15%S	15%S	15%S	10%G	2.63%all	1.05%all			
	10%S	10%S	5%L	5%R	5%R	5%R	5%R	2.63%all					
	10%R	10%R	5%R	5%D	5%D	5%D	5%D						
	1.59%all	1.59%all	5%S	5%N	5%N	5%N	5%N						
		1.58%all	2.10%all	2.10%all	2.10%all	2.10%all	2.10%all						
mezcla TRIM	T1	T2	T3	T3	T3	T3	T3	T4	T5	T6	oscilante	T7	
Diversidad	19	19	19	19	19	19	19	19	19	19	2	2	

Figura 7

FW4		FW4											
87	88	89	90	91	92	93	94	95	96	97	98	99	100
67% Y 21% F 2% H	100% C	53% Q 11% F 10% S 8% L 7% H 4% W 2% M 1% K 1% G 1% A 1% V	90% Q 6% H 4% N	24% G 22% S 15% Y 13% W 8% H 5% R 3% D 3% N 3% F 1% T 1% A 1% V 1% L	35% S 18% T 14% N 10% Y 5% H 3% D 3% R 3% L 3% W 2% K 1% G 1% V 1% L	34% S 27% H 15% E 5% T 4% G 3% R 3% N 2% K 2% Y 1% D 1% Q 1% I	27% Y 22% V 10% L 10% F 7% N 5% D 5% T 5% S 4% I 2% W 1% Q 1% P	97% P 1% T 1% L 1% X	28% Y 23% L 17% W 10% P 10% F 7% R 1% H 1% Q 1% C 1% V	89% T	99% F	99% G	62% G 29% A 13% S 1% T

Figura 8

FW3		89	90	91	92	93	94	95	96	FW4			
Y	C									T	F	G	
		50%Q 10%F 10%S 1.57%all	100%Q	20%G 20%S 10%Y 2.63%all	30%S 15%T 15%N 10%Y 1.57%all	30%S 25%H 15%E 1.57%all	25%Y 20%V 10%L 10%F 1.84%all	100%P	30%Y 20%L 10%P 10%F 1.57%all				
Mezcla TRIM		T8	-	T9	T10	T11	T12	-	T13				
Diversidad		19	1	19	19	19	19	1	19				

Figura 9

mVH1_RV									
50	51	52	52a	52b	54	55	56	57	58
R	I	70D	P	N	80S	G	G	T	K
		15N			20N				
		15Y			W2				
		W1							
VH2									
V	I	W	80S		70G	G	70S	T	D
			20G		30D		30G		
			W3		W4		W5		
VH5									
T	I	70S	G	G	70G	G	60N	T	Y
		30N			30S		40S		
		W8			W9		W10		

Figura 10

	50	51	52	52a	53	54	55	56	57	58	59	60	61	62	63	64	65
IN#1 5979 sec.	56% R 10% E 7% Y 6% N 5% D 3% V 3% M 3% W 2% A 1% K 1% Q 1% G 1% L	39% I 8% V 2% F 5% H 3% L 3% V 1% S 1% F	60% D 15% N 13% Y 5% H 3% L 1% S 1% F	98% P 1% T	54% N 17% G 15% S 8% Y 1% D 1% E 1% K 1% R	66% S 14% N 12% D 3% T 2% Y 1% A 1% G	84% G 9% S 4% D 1% A 1% V 4% N 3% E 2% F 1% R 1% T 1% A 1% I 1% V	58% G 10% S 10% Y 8% D 4% N 3% E 2% F 1% R 1% T 1% A 1% I 1% F	84% T 11% I 2% S 1% A 1% P 3% Y 1% E 1% T 1% A 1% I 1% F	51% K 23% N 8% R 5% S 3% H 1% E 1% T 1% A 1% I 1% F	96% Y 2% F 1% L	94% N 4% S 1% D 1% T	75% E 21% Q 3% G	93% K 3% N 2% R 1% M	99% F 1% L	96% K 2% R 1% E	43% S 28% G 13% N 9% D 5% T 1% R 1% A 1% I
IN#2 586 sec.	68% V 10% M 1% T 1% I	98% I 1% M	39% W		34% A 30% S 17% G 12% R 6% T	75% G 23% D 1% S 1% N 1% A	99% G 8% G 4% N 2% T 1% D 1% R 1% I 1% X	82% S 8% G 4% N 2% T 1% D 1% R 1% I 1% X	99% T 41% D 10% T 2% Y 1% E 1% S 1% G 1% I	43% N 4% H 1% X	99% Y	84% N 4% H 1% X	69% S 30% A	97% A 1% T 1% G 1% V	66% L 33% F	51% M 27% K 21% I	100% S
IN#5 667 sec.	43% T 24% Y 3% E 3% S 3% N 2% V 1% K 1% I 1% M	99% I	70% S 24% N 3% T 1% D 1% R 1% G	67% S 9% D 8% N 5% G 3% R 3% Y 2% T 1% P 1% I	66% G 15% D 12% N 2% L 1% R 1% S	82% G 2% A 4% T 3% Y 2% K 2% I 2% N 1% V	50% S 33% G 4% D 4% T 3% Y 2% R 2% N 1% V	39% Y 31% S 15% T 8% N 2% K 2% J 1% R 1% F	76% T 20% I 2% A 1% S 1% L	94% Y 3% F 1% H 1% N	99% Y	74% P 16% A 5% L 2% S	96% D 1% E 1% G	48% T 46% S 3% N 1% G 1% A 1% I	85% V 9% M 4% L	84% K 9% E 5% T 1% R 1% Q	87% G 9% R 2% D 2% S

Figura 11

	50	51	52	52a	53	54	55	56	57	58	59	60	61	62	63	64	65	EV3
mVH1	5.26%all	5.26%all	50%D 20%N 10%Y 10%H	P	5.26%all	50%S 10%N 10%D 1.5%all	80%G 10%S 10%D	5.26%all	70%T 10%all	5.26%all	Y	N	60%E 40%Q	K	F	K	S	
Mezcla TRIM	MMT-1	-	MMT-2	-	MMT-1	MMT-3	MMT-4	MMT-1	MMT-5	MMT-1	-	-	W1	-	-	-	-	
Diversidad	19	1	4	1	19	19	3	19	19	19	1	1	2	1	1	1	1	1
mVH2	80%Y 20%M	70%W 1.58% all			15%A 15% S 15% G 15% R 2.11%all	60%G 10%D 1.58% all	G	50% S 2.6%all	T	5.26%all	Y	N	60%S 40%A	A	F	I	S	
Mezcla TRIM	W2	-	MMT-6	-	MMT-7	MMT-8	-	MMT-9	-	MMT-1	-	-	MMT-10	-	-	-	-	
Diversidad	2	1	19		19	19	1	19	1	19	1	1	2	1	1	1	1	1
mVH5	5.26%all	20%S 4.2%all	20%S 20%G 3.2%all	20%S 4.7%all 3.2%all	10%G 3.2%all	70%G 30%S 3.2%all	20%S 20%G 3.2%all	20%Y 20%N 3.2%all	30%T 10%all 3.2%all	Y	Y	60%P 40%A	D	S	V	K	G	
Mezcla TRIM	MMT-1	-	MMT-11	MMT-12	MMT-13	W3	MMT-12	MMT-14	MMT-15	-	-	W4	-	-	-	-	-	
Diversidad	19	1	19	19	19	2	19	19	19	1	1	2	1	1	1	1	1	1

Figura 12

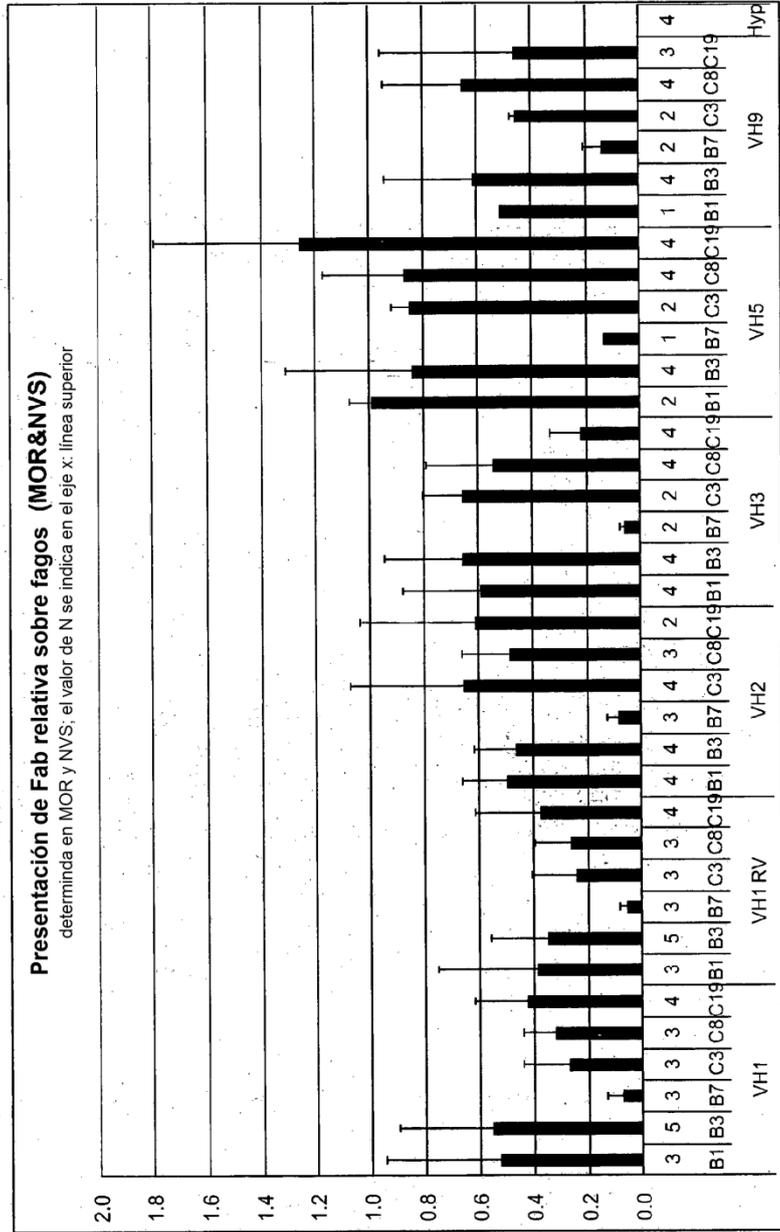


Figura 13
Expresión en Fab murino: NVS & MOR

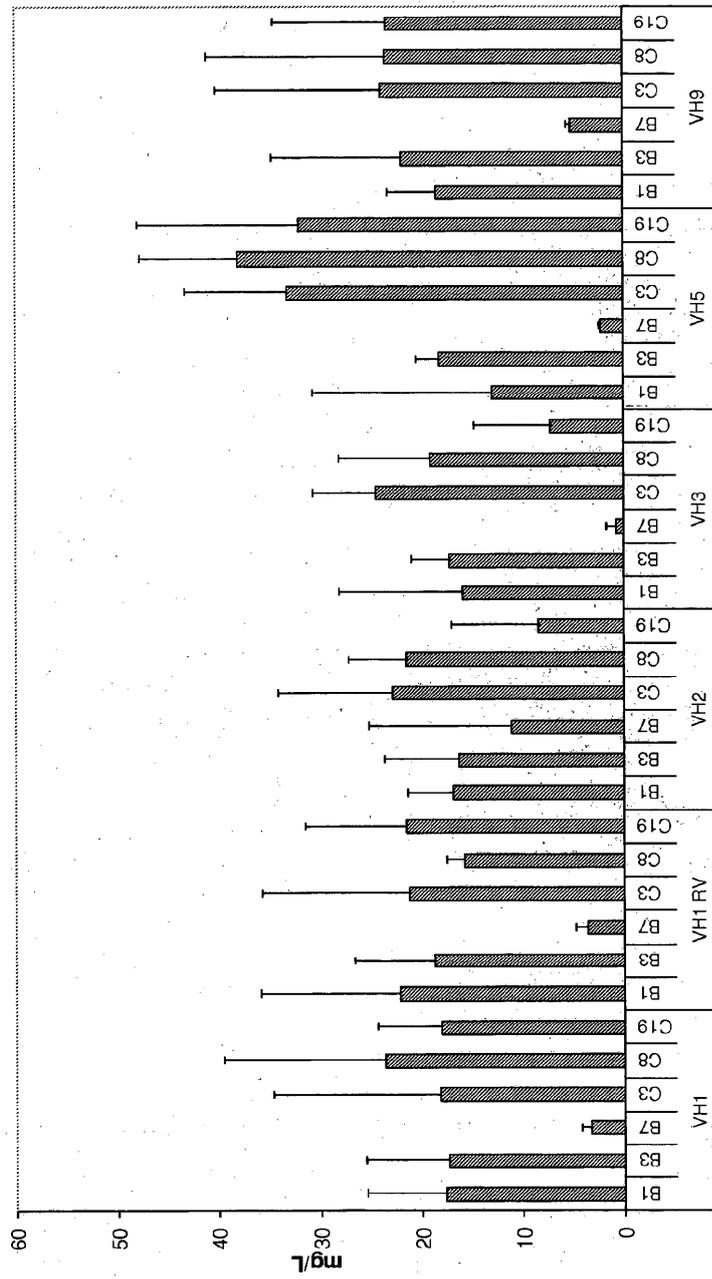


Figura 14

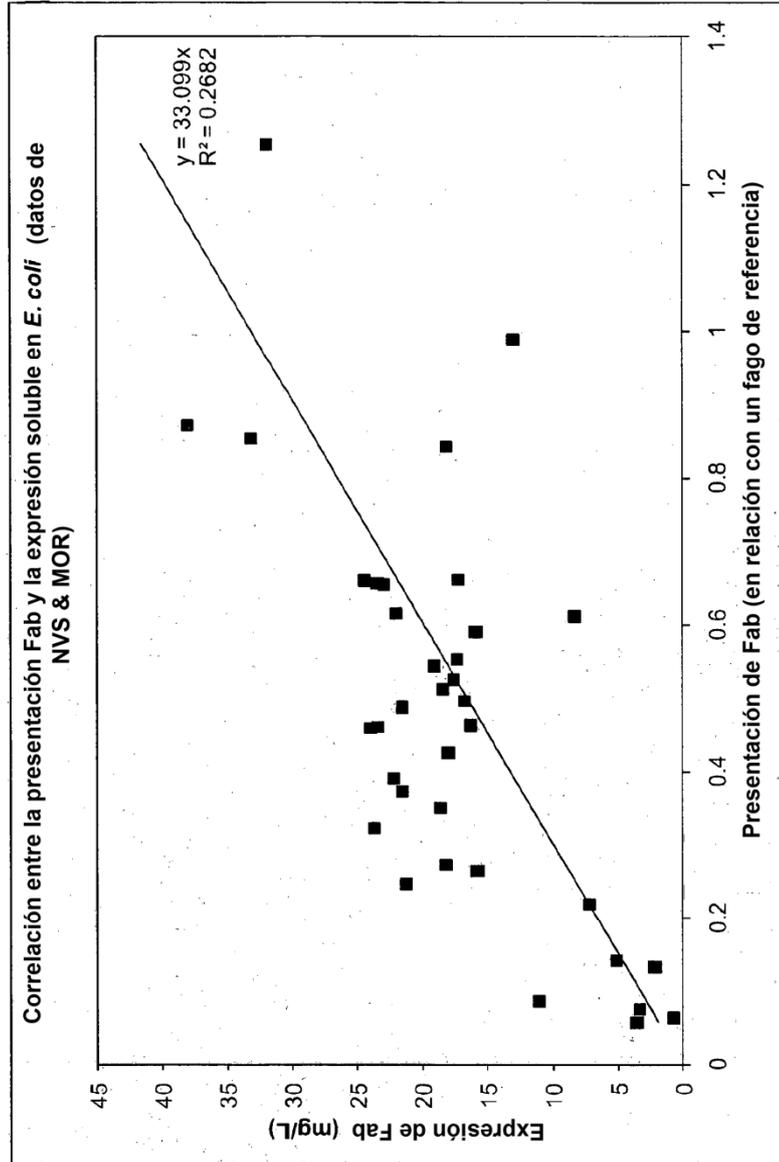


Figura 15
expresión de IgG2s murina: solo MOR

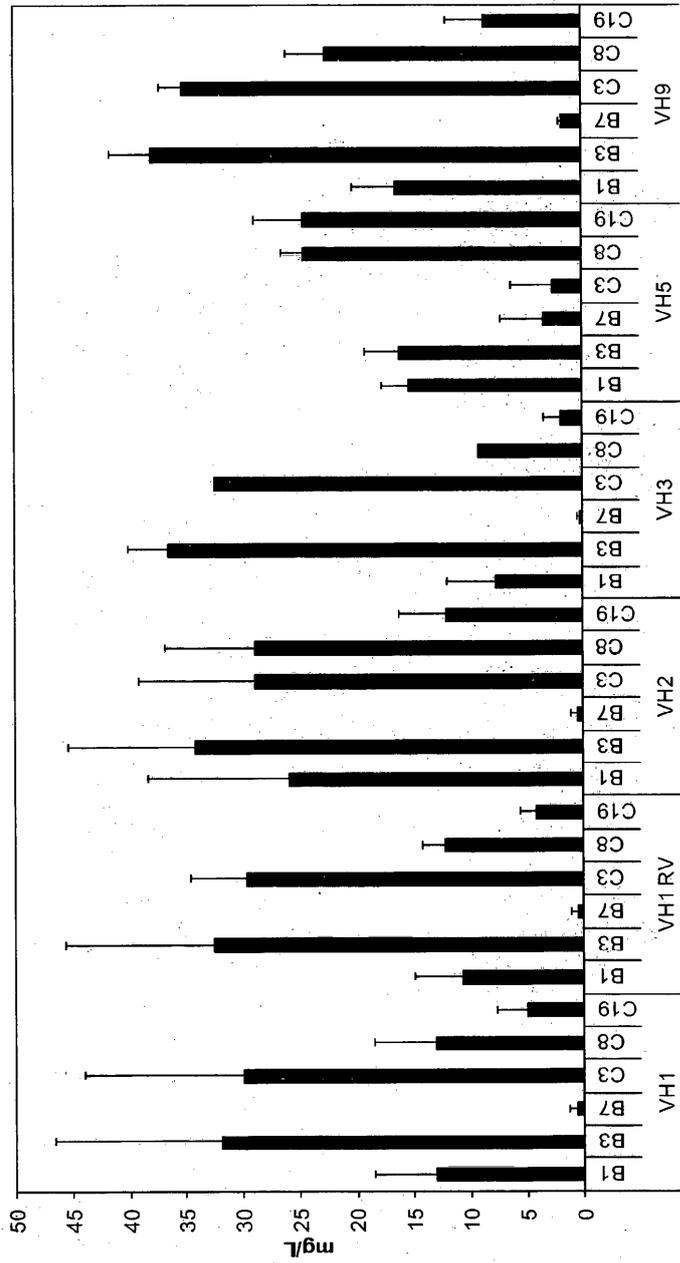


Figura 16

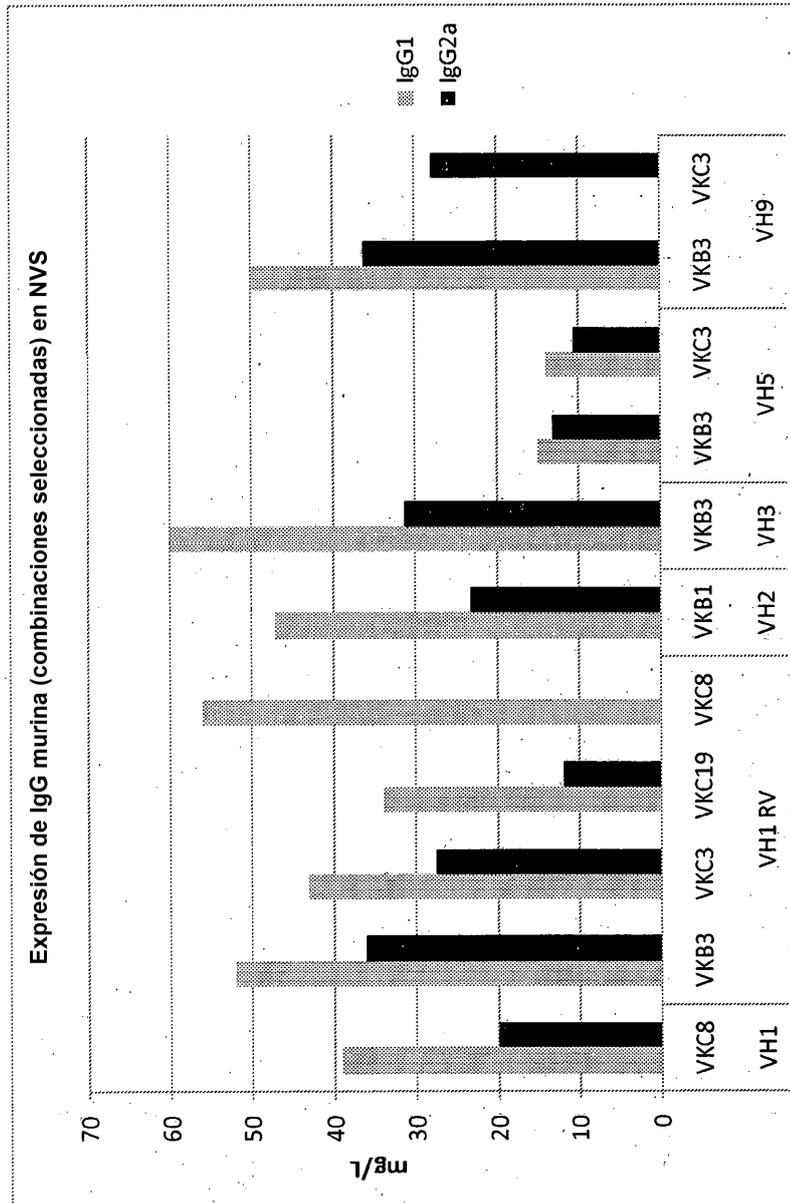


Figura 17a
 Datos para la purificación con isopropanol

Estruct. artificial	conc [mg/ml]	rend. [mg]	rend. [mg/L]	Tm [°C]
VkB1_VH3	0.14	0.9	3.70	71.1
VkB3_VH3	1.00	5.4	21.59	70.4
VkB7_VH3	0.00	0.0	0.00	0
VkC3_VH3	0.50	3.8	15.10	70.2
VkC8_VH3	0.32	2.1	8.25	61.2
VkC19_VH3	0.19	1.0	3.90	62.4
VkB1_VH5	0.64	4.6	18.53	68.4
VkB3_VH5	1.02	7.1	28.53	70.4
VkB7_VH5	0.13	0.7	2.65	68.9
VkC3_VH5	0.91	6.8	27.23	68.7
VkC8_VH5	1.14	8.5	34.19	62.4
VkC19_VH5	0.81	6.1	24.46	66.3
VkB1_VH9	1.11	7.8	31.06	68.7
VkB3_VH9	0.84	7.2	28.60	68.2
VkB7_VH9	0.16	0.8	3.30	69.4
VkC3_VH9	1.51	11.4	46.43	70.4
VkC8_VH9	1.40	9.8	39.27	65.055
VkC19_VH9	1.05	7.9	31.59	65.94
VkB1_VH1 RVT	0.95	7.2	28.64	71.385
VkB3_VH1 RVT	0.77	5.8	23.09	75.69
VkB7_VH1 RVT	0.12	0.5	1.95	72.715
VkC3_VH1 RVT	0.96	7.2	28.62	75.42
VkC8_VH1 RVT	0.25	1.8	7.00	71.055
VkC19_VH1 RVT	0.90	6.7	26.67	71.22
VkB1_VH1	0.68	5.0	20.04	70.97
VkB3_VH1	1.03	7.7	30.62	72.3
VkB7_VH1	0.12	0.7	2.66	72.67
VkC3_VH1	0.52	3.9	15.51	74.9
VkC8_VH1	1.03	7.7	30.69	70.9
VkC19_VH1	0.59	4.5	17.91	70.93
VkB1_VH2	0.81	5.4	21.47	69.43
VkB3_VH2	0.64	4.5	17.87	70.595
VkB7_VH2	0.00	0.0	0.00	0
VkC3_VH2	0.84	5.9	23.64	68.725
VkC8_VH2	1.21	9.1	36.32	67.17
VkC19_VH2	0.00	0.0	0.00	0

<0.5 mg/ml	< 10 mg	< 68 °C
0.5 - 1 mg/ml	10 - 20 mg	68 - 72 °C
1 mg/ml	> 20 mg	> 72 °C

Figura 17b
 Datos para la purificación sin isopropanol

Estruc. artificial	conc [mg/ml]	rend. [mg]	rend. [mg/L]	Tm [°C]
VkB1_VH3	0.27	1.8	7.20	71.6
VkB3_VH3	0.71	5.0	19.91	69.9
VkB7_VH3	0.00	0.0	0.00	0
VkC3_VH3	0.97	7.2	28.81	70.7
VkC8_VH3	0.49	3.2	12.64	61.1
VkC19_VH3	0.07	0.5	1.84	62.4
VkB1_VH5	0.07	0.1	0.53	68.6
VkB3_VH5	0.90	4.9	19.79	70.2
VkB7_VH5	0.12	0.6	2.39	68.7
VkC3_VH5	1.12	10.1	40.25	69.4
VkC8_VH5	1.13	11.2	44.96	62.6
VkC19_VH5	1.14	10.8	43.36	66.1
VkB1_VH9	0.73	5.5	21.90	68.7
VkB3_VH9	1.11	7.7	30.95	68.7
VkB7_VH9	0.16	1.2	4.83	70.2
VkC3_VH9	1.15	8.9	35.44	70.2
VkC8_VH9	1.00	9.0	36.01	65.61
VkC19_VH9	0.87	7.8	31.30	65.86
VkB1_VH1 RVT	0.88	8.0	31.86	71.42
VkB3_VH1 RVT	0.64	6.1	24.20	73.745
VkB7_VH1 RVT	0.13	0.7	2.68	72.155
VkC3_VH1 RVT	1.04	7.9	31.47	75.63
VkC8_VH1 RVT	0.57	4.3	17.01	71.39
VkC19_VH1 RVT	0.95	7.1	28.60	72.595
VkB1_VH1	0.64	5.8	23.19	72.18
VkB3_VH1	0.83	5.8	23.12	74.455
VkB7_VH1	0.12	0.7	2.69	73.3
VkC3_VH1	1.00	7.5	29.90	74.9
VkC8_VH1	1.16	8.7	34.83	70.9
VkC19_VH1	0.75	5.6	22.52	71.01
VkB1_VH2	0.59	5.0	20.03	68.43
VkB3_VH2	0.73	5.4	21.51	70.84
VkB7_VH2	0.55	5.3	21.07	70.51
VkC3_VH2	0.91	7.7	30.81	70.51
VkC8_VH2	0.64	6.4	25.50	65.98
VkC19_VH2	0.08	0.5	2.17	63.66

<0.5 mg/ml	< 10 mg	< 68 °C
0.5 - 1 mg/ml	10 - 20 mg	68 - 72 °C
1 mg/ml	> 20 mg	> 72 °C

Figura 18

Estabilidad térmica en formato Fab:NVS
sin isopropanol

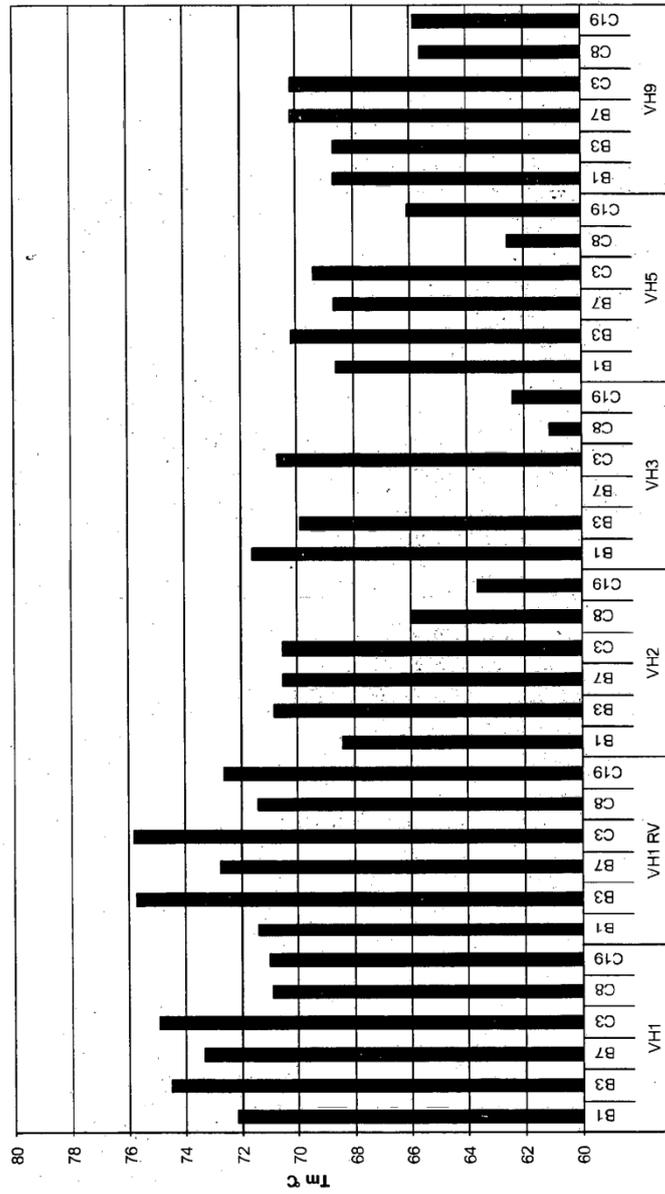


Figura 19

FW3	95	96	97	98	98	99	100	a	b	m	n	101	102	FW4
	T1	T1							T4	T5	T6	oscilante	T7	
	T1	T1	T2						T4	T5	T6	oscilante	T7	
	T1	T1	T2	T3					T4	T5	T6	oscilante	T7	
	T1	T1	T2	T3	T3				T4	T5	T6	oscilante	T7	
	T1	T1	T2	T3	T3	T3			T4	T5	T6	oscilante	T7	
	T1	T1	T2	T3	T3	T3	T3	T3	T4	T5	T6	oscilante	T7	

L_7
L_8
L_9
L_10
L_11
L_12

Figura 20

89 T8	90			91 T9			92 T10			93 T11			94 T12			95			96 T13		
	F	P		F	P		F	P		F	P		F	P		F	P		F	P	
0.9	1.6			2.9	2.6		0.9	1.6		0.8	1.6		1.8	1.8					1.2	1.6	
0.7	1.6			2.1	2.6		1.5	1.6		1.6	1.6		1.1	1.8					0.8	1.6	
1.7	1.6			3.5	2.6		2.5	1.6		18.7	16.6		1.9	1.8					3.0	1.6	
10.6	11.6			2.7	2.6		1.8	1.6		1.6	1.6		13.9	11.8					14.8	11.6	
0.9	1.6			16.4	22.6		0.8	1.6		0.5	1.6		0.8	1.8					1.2	1.6	
2.4	1.6			3.6	2.6		1.5	1.6		31.1	26.6		2.1	1.8					1.9	1.6	
1.2	1.6			2.0	2.6		1.2	1.6		1.0	1.6		1.6	1.8					1.5	1.6	
0.9	1.6			2.9	2.6		1.7	1.6		1.1	1.6		1.3	1.8					1.7	1.6	
1.3	1.6			1.2	2.6		1.9	1.6		1.3	1.6		6.6	11.8		0.1			11.4	21.6	
1.5	1.6			3.7	2.6		1.8	1.6		1.6	1.6		2.5	1.8					2.7	1.6	
0.7	1.6			1.9	2.6		12.4	16.6		1.3	1.6		1.0	1.8					1.4	1.6	
1.5	1.6			2.5	2.6		1.3	1.6		0.9	1.6		0.7	1.8		99.9	100		1.6	11.6	
56.1	51.6			3.2	2.6		1.8	1.6		1.6	1.6		1.1	1.8					2.3	1.6	
0.6	1.6			1.4	2.6		1.1	1.6		1.3	1.6		1.3	1.8					1.0	1.6	
11.0	11.6			21.0	22.6		28.2	31.6		27.6	31.6		1.6	1.8					1.4	1.6	
1.2	1.6			2.9	2.6		17.4	16.6		1.4	1.6		1.2	1.8					1.1	1.6	
2.3	1.6			5.0	2.6		2.9	1.6		2.7	1.6		25.6	21.8					2.4	1.6	
3.0	1.6			4.3	2.6		1.8	1.6		1.6	1.6		2.5	1.8					3.9	1.6	
1.5	1.6			16.8	12.6		17.4	11.6		2.1	1.6		31.4	26.8					44.8	31.6	

A C D E F G H I K L M N P Q R S T V W Y

Figura 21

95-96		97		98-100a		100b		100m		100n		101		102	
T1		T2		T3		T4		T5		T6		W1		T7	
F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P
1.3	1.6	0.9	1.6	1.8	2.1	2.1	2.6	19.6	22.6	0.7	1.1	15.4	20.0		
9.7	11.6	3.4	6.6	5.6	7.1	1.0	2.6	1.5	2.6	0.5	1.1	84.6	80.0		
1.8	1.6	0.6	1.6	1.7	2.1	2.8	2.6	3.9	2.6	0.5	1.1				
1.9	1.6	2.1	1.6	2.9	2.1	3.4	2.6	2.6	2.6	83.5	61.1				
16.7	21.6	7.5	11.6	10.2	17.1	7.7	12.6	1.0	2.6	0.3	1.1				
1.2	1.6	2.4	1.6	1.9	2.1	2.3	2.6	2.8	2.6	0.7	1.1				
1.2	1.6	1.3	1.6	0.9	2.1	1.3	2.6	1.3	2.6	0.5	1.1				
2.5	1.6	2.4	1.6	2.8	2.1	3.3	2.6	5.4	2.6	0.7	1.1				
1.2	1.6	1.9	6.6	2.5	2.1	2.6	2.6	2.0	2.6	0.7	1.1				
1.0	1.6	1.1	1.6	2.0	2.1	1.6	2.6	0.7	2.6	6.2	21.1				
2.0	1.6	1.7	1.6	8.4	7.1	2.1	2.6	2.5	2.6	0.7	1.1				
0.5	1.6	0.6	1.6	0.9	2.1	1.0	2.6	1.5	2.6	0.2	1.1				
0.5	1.6	0.6	1.6	0.8	2.1	1.3	2.6	1.0	2.6		1.1				
5.5	11.6	2.8	6.6	4.0	7.1	1.3	2.6	1.1	2.6	0.5	1.1				
14.2	11.6	8.8	6.6	25.0	17.1	15.5	12.6	4.1	2.6	1.8	1.1				
2.8	1.6	2.4	1.6	3.1	2.1	2.6	2.6	3.4	2.6	0.5	1.1				
3.2	1.6	1.5	1.6	2.2	2.1	3.3	2.6	2.8	2.6	0.8	1.1				
0.7	1.6		1.6	0.8	2.1	0.2	2.6	0.7	2.6	0.2	1.1			17.5	20.0
32.2	21.6	57.7	41.6	22.4	17.1	44.4	32.6	42.2	32.6	1.3	1.1			82.5	80.0

A C D E F G H I K L M N P Q R S T V W Y

Figura 22

	50		53		56		58		51		52		52a		54		55		57		59		60		61		62		63		64	
	Med	F	F	P	F	P	F	P	F	P	F	P	F	P	F	P	MMT-3	F	P	F	P	MMT-5	F	P	F	P	F	P	F	P		
A	3.2	5.3														2.2	1.5															
C	4.0	5.3									39.8	50.0				8.6	11.5	8.6	10.0	2.2	1.0											
D	5.1	5.3														3.2	1.5			3.2	1.0											
E	5.6	5.3														1.1	1.5															
F	3.2	5.3														1.1	1.5	73.1	80.0	1.1	1.0											
G	8.3	5.3									18.3	10.0				2.2	1.5			2.2	1.0											
H	5.4	5.3							100	100																						
I	3.5	5.3														1.1	1.5			1.1	1.0											
K	5.4	5.3														2.2	1.5			1.1	1.0											
L	8.1	5.3														2.2	1.5			3.2	1.0											
M	3.2	5.3									21.5	20.0				7.5	11.5			1.1	1.0											
N	5.4	5.3														1.1	1.5															
P	5.1	5.3																														
Q	4.3	5.3																														
R	5.9	5.3														61.3	51.5	18.3	10.0	2.2	1.0											
S	5.6	5.3														1.1	1.5			69.9	71.0											
T	6.7	5.3														1.1	1.5			2.2	1.0											
V	5.1	5.3														1.1	1.5															
W	6.7	5.3									20.4	20.0				3.2	1.5															
Y																																

Figura 23

	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65
	MMT-6		MMT-7		MMT-8		MMT-9		MMT-1		MMT-10					
	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P
A			1.6	202	17.1	1.1	1.6									
C			4.3	1.6	1.1	2.1	7.4	1.6								
D			1.6	3.2	2.1		1.6									
E			1.1	1.6	2.1	2.1	1.6									
F			1.6	12.8	17.1	66.0	61.6	100.0	100.0							
G			1.6	6.4	2.1	2.1	1.6									
H			100.0	100.0												
I			1.1	1.6	1.1	2.1	2.1	1.6								
K			1.1	1.6	2.1	2.1		1.6								
L			1.1	1.6	2.1	2.1	2.1	1.6								
M	25.5	20.0														
N			1.6	1.1	2.1		1.6									
P			2.1	1.6	2.1	2.1	1.6									
Q			4.3	1.6	2.1	2.1	1.6									
R			1.6	16.0	17.1	1.1	1.6									
S			1.6	17.0	17.1	1.1	1.6									
T			3.2	1.6	1.1	2.1	1.6									
V	74.5	80.0	2.1	1.6	4.3	2.1	4.3	1.6								
W			76.6	71.6	1.1	2.1	2.1	1.6								
Y			2.1	1.6	2.1	2.1	4.3	1.6								

Figura 24

50	31		32		32a		35		53		54		56		57		58		59		60		61		62		63		64		65			
	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P				
MMT-1																																		
A	60	53	24	42	12	32	72	47						24	32	24	32																	
C																																		
D	48	53	36	42	42	32	36	47						48	32	32																		
E	60	53	84	42	24	32	36	47						32	48	32																		
F	84	53	36	42	24	32	84	47						24	32	12	32																	
G	24	53	42	127	232	108	147	651	700					36	32	36	32																	
H	72	53	72	42	42	32	72	47						36	32	48	32																	
I	36	53	36	42	24	32	36	47						24	32	145	132																	
K	36	53	12	42	24	32	60	47						12	32	12	32																	
L	72	53	24	42	24	32	24	47						24	32	72	32																	
M	36	53	72	42	48	32	36	47						36	32	36	32																	
N	60	53	24	42	24	32	36	47						181	232	36	32																	
P	36	53	36	42	36	32	72	47						12	32	108	32																	
Q	72	53	60	42	84	32	72	47						84	32	32																		
R	24	53	24	42	24	32	36	47						24	32	24	32																	
S	48	53	253	242	271	232	32	47	349	300				36	32	12	32																	
T	72	53	60	42	48	32	24	47						12	32	265	332																	
V	60	53	72	42	60	32	96	47						48	32	96	32																	
W	84	53	36	42	42	32	48	47						24	32	32																		
Y	12	53	36	42	48	32	36	47						313	232	24	32	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000		

Figura 25

50	51	52	52a	53	54	55	55	56	57	58	59	60	61	62	63	64	65
MMT-22	MMT-23	MMT-1	MMT-24	MMT-25	MMT-26	MMT-27	MMT-28	MMT-29	MMT-30	MMT-31	MMT-32	MMT-33	MMT-34	MMT-35	MMT-36	MMT-37	MMT-38
F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P	F P F P
		26 48 53		21 21 32 21		223 200											
A																	
C		21 26 48 53		21 21 21		100.0 100.0											
D		53 26 59 53		1.1 21 21 21													
E		1.1 26 59 53		21 21 21													
F		21 26 27 53		702 700 309 421		149 200											100.0 100.0
G		26 59 53		43 21 21 21													
H		32 26 32 53		1.1 21 21 21													
I		1.1 26 37 53		32 21 32 21													
K		21 26 59 53		32 21 21 21													
L		1.1 26 64 53		21 21 32 21													
M		32 26 27 53		32 21 106 171													
N		32 26 96 53		43 21 21													
P		32 26 96 53		74 21 53 21													
Q		1.1 26 43 53		32 21 21 21													
R		394 326 48 53		298 300 234 221 383 321		245 200											
S	394 500	223 226 74 53		1.1 21 191 171													
T	255 250	32 26 32 53		32 21 21 21													
V		1.1 26 37 53		21 21 11 21													
W		53 26 59 53		21 21 32 21													
Y	351 250																

Figura 26

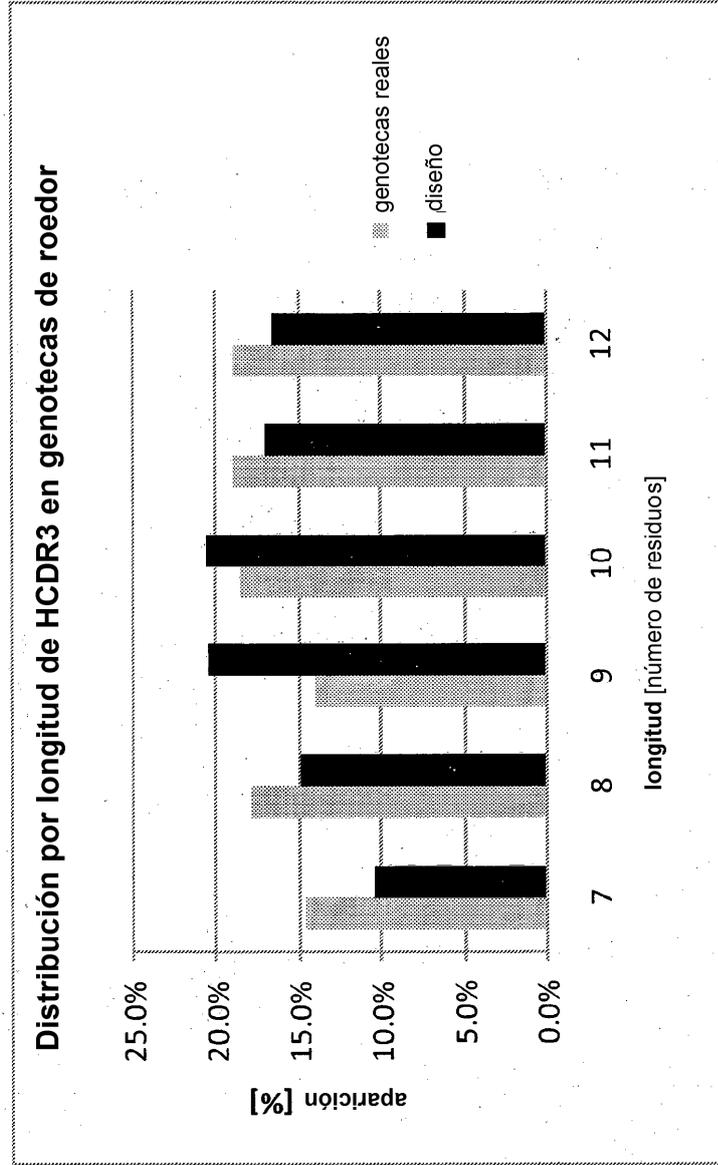


Figura 27a

CDRs de HuCAL	CDRs de SAS	CDRs de Kabat	Nº de Kabat	línea germinal más próxima	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
mV _{LB} 1 (BALB/c)				D0008 IGKV 4-177 ⁰ IM _{US}	D	I	L	M	T	Q	T	P	L	S	L	P	V	S	L	G	D	Q	
mV _{LB} 3 (BALB/c)				K02169 IGKV 3-12 ⁰ IM _{US}	D	I	V	L	T	Q	S	P	A	S	L	A	V	S	L	G	Q	R	
mV _{LB} 7 (BALB/c)				AFO44198 IGKV 7-33 ⁰ IM _{US}	D	I	V	M	T	Q	S	P	T	F	L	A	V	T	A	S	K	K	
mV _{LC} 3 (C75BL6)				Y5968 IGKV 3-4 ⁰ IM _{US}	D	I	V	L	T	Q	S	P	A	S	L	A	V	S	L	G	Q	R	
mV _{LC} 8 (C75BL6)				AJ235947 IGKV 6-26 ⁰ IM _{US}	D	I	V	M	T	Q	S	P	S	S	L	S	V	S	A	G	E	K	
mV _{LC} 19 (C75BL6)				AJ235936 IGKV 6-93 ⁰ IM _{US}	D	I	Q	M	T	Q	S	P	S	S	L	S	A	S	L	G	G	K	
rVL1				grupo de consenso 1	D	I	V	M	T	Q	S	P	S	S	L	A	V	S	A	G	E	K	
rVL2				grupo de consenso 2	D	I	Q	M	T	Q	S	P	S	S	L	S	A	S	L	G	D	R	

ratón/rata

Figura 27b

		LCDR1																				
		LCDR1																				
		LCDR1																				
10	20	21	22	23	24	25	26	27	28	29	30	30a	30b	30c	30d	30e	30f	31	32	33	34	
A	S	I	S	C	R	S	S	Q	S	I	V	H	S	N	G	N	T	Y	L	E		
A	T	I	S	C	R	A	S	K	S	V	S	T	S	G		Y	S	Y	M	H		
V	T	I	S	C	T	A	S	E	S	L	Y	S	S	K	H	K	V	H	Y	L	A	
A	T	I	S	C	K	A	S	Q	S	V	D	Y	D	G		D	S	Y	M	N		
V	T	M	S	C	K	S	S	Q	S	L	L	N	S	G	N	Q	K	N	Y	L	A	
V	T	I	T	C	K	A	S	Q	D	I	N	K							Y	I	A	
V	S	I	S	C	R	S	S	Q	S	L	V	H	S	D	G	K	K	T	Y	L	N	
V	T	I	E	C	K	A	S	Q	N	I	Y	K							Y	L	A	

Figura 27c

CDRs de HUCAL	ratón/rata																	
CDRs de SAS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
CDRs de Kabat	Escribir																	
Nº de Kabat																		
mVLB1	GAT	ATC	CTG	ATG	ACC	CAG	ACC	CCA	CTG	AGC	CTG	CCA	GTT	AGC	CTG	GGC	GAT	CAG
mVLB3	GAT	ATC	GTG	CTG	ACC	CAG	AGC	CCA	GCA	AGC	CTG	GCC	GTG	AGC	CTG	GCC	CAA	CGC
mVLB7	GAT	ATC	GTG	ATG	ACC	CAG	AGC	CCG	ACC	TTT	CTG	GCC	GTG	ACC	GCC	AGC	AAA	AAA
mVLC3	GAT	ATC	GTG	CTG	ACC	CAG	AGC	CCA	GCG	AGC	CTG	GCC	GTG	AGC	CTG	GCC	CAA	CGC
mVLC8	GAT	ATC	GTG	ATG	ACC	CAG	AGC	CCG	AGC	AGC	CTG	AGC	GTT	AGC	GCA	GCC	GAA	AAA
mVLC19	GAT	ATC	CAG	ATG	ACC	CAG	AGC	CCG	AGC	AGC	CTG	AGC	GCC	AGC	CTG	GCC	GCC	AAA
rVL1	GAT	ATC	GTG	ATG	ACC	CAG	TCT	CCG	AGC	AGC	CTG	GCC	GTT	AGC	GCG	GGT	GAA	AAA
rVL2	GAT	ATC	CAG	ATG	ACC	CAG	AGC	CCG	AGC	AGC	CTG	AGC	GCG	AGC	CTG	GGT	GAT	CGT

Uno de los dos aminoácidos (T o V) se puede encontrar en estos residuos (solo VK1 y VK3)
Región de la secuencia recuperada del „Sequence Analysis Software“ (SAS)
Posición de la región diversificada
X
Variabilidad potencial de la longitud
No hay aminoácidos en esta posición

Figura 27d

LCDR1																						
LCDR1																						
LCDR1																						
19	20	21	22	23	24	25	26	27	28	29	30	30a	30b	30c	30d	30e	30f	31	32	33	34	
					Psi																	
GCC	ACC	ATT	AGC	TGC	CGC	AGC	AGC	CAG	AGC	ATT	GTG	CAT	AGC	AAC	GGC	AAC	AGC	TAC	CTG	GAA		
GCC	ACC	ATT	AGC	TGC	CGC	GCC	AGC	AAA	AGC	GTT	AGC	ACC	AGC	GGC				TAT	AGC	TAT	ATG	CAG
GTG	ACC	ATT	AGC	TGC	ACC	GCC	AGC	GAA	AGC	CTG	TAT	AGC	AGC	AAA	CAT	AAA	GTG	CAG	TAT	CTG	GCC	
GCC	ACC	ATT	AGC	TGC	AAA	GCC	AGC	CAG	AGC	GTG	GAT	TAT	GAT	GGC			GAT	AGC	TAT	ATG	AAC	
GTG	ACC	ATT	AGC	TGC	AAA	AGC	AGC	CAG	AGC	CTG	CTG	AAC	AGC	GCC	AAC	CAG	AAA	AAC	TAT	CTG	GCC	
GTG	ACC	ATT	ACC	TGC	AAA	GCC	AGC	CAG	GAT	ATT	AAC	AAA							TAT	ATC	GCC	
GTG	ACC	ATT	AGC	TGC	AGC	AGC	AGC	CAG	AGC	CTG	GTT	CAT	AGC	GAT	GGC	AAA	AAA	ACC	TAT	CTG	AAG	
GTG	ACC	ATT	GAA	TGC	AAA	GCG	AGC	CAG	AAC	ATC	TAT	AAA							TAT	CTG	GCC	

Uno de los dos aminoácidos (T o V) se puede encontrar en estos residuos (solo Vk1 y Vk3)

 Región de la secuencia recuperada del „Sequence Analysis Software“ (SAS)

 Posición de la región diversificada

 Variabilidad potencial de la longitud

 No hay aminoácidos en esta posición

Figura 28a

		LCDR2										LCDR2										LCDR2																																																													
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62																																					
W	Y	L	Q	K	P	G	Q	S	P	K	L	L	I	Y	K	V	S	N	R	F	S	G	V	P	D	R	F	W	Y	L	Q	K	P	G	Q	S	P	K	L	L	I	Y	K	V	S	N	R	F	S	G	V	P	D	R	F	W	Y	L	Q	K	P	G	Q	S	P	K	L	L	I	Y	K	V	S	N	R	F	S	G	V	P	D	R	F
W	Y	Q	Q	K	P	G	Q	P	P	K	L	L	I	Y	L	A	S	N	L	E	S	G	V	P	A	R	F	W	Y	Q	Q	K	P	G	Q	P	P	K	L	L	I	Y	L	A	S	N	L	E	S	G	V	P	A	R	F	W	Y	Q	Q	K	P	E	Q	S	P	K	L	L	I	Y	G	A	S	N	R	Y	I	G	V	P	D	R	F
W	Y	Q	Q	K	P	G	Q	P	P	K	L	L	I	Y	A	A	S	N	L	E	S	G	V	P	A	R	F	W	Y	Q	Q	K	P	G	Q	P	P	K	L	L	I	Y	A	A	S	N	L	E	S	G	V	P	A	R	F	W	Y	Q	Q	K	P	G	Q	P	P	K	L	L	I	Y	A	A	S	N	L	E	S	G	V	P	A	R	F
W	Y	Q	H	K	P	G	K	G	P	R	L	L	I	H	Y	T	S	T	L	Q	P	G	V	P	S	R	F	W	Y	Q	H	K	P	G	K	G	P	R	L	L	I	H	Y	T	S	T	L	Q	P	G	V	P	S	R	F	W	Y	Q	H	K	P	G	K	G	P	R	L	L	I	H	Y	T	S	T	L	Q	P	G	V	P	S	R	F
W	Y	L	Q	K	P	G	Q	S	P	Q	L	L	I	Y	W	V	S	N	R	E	S	G	V	P	D	R	F	W	Y	L	Q	K	P	G	Q	S	P	Q	L	L	I	Y	W	V	S	N	R	E	S	G	V	P	D	R	F	W	Y	L	Q	K	P	G	Q	S	P	Q	L	L	I	Y	W	V	S	N	R	E	S	G	V	P	D	R	F
W	Y	Q	Q	K	P	G	K	A	P	K	L	L	I	Y	N	A	N	S	L	Q	T	G	V	P	S	R	F	W	Y	Q	Q	K	P	G	K	A	P	K	L	L	I	Y	N	A	N	S	L	Q	T	G	V	P	S	R	F	W	Y	Q	Q	K	P	G	K	A	P	K	L	L	I	Y	N	A	N	S	L	Q	T	G	V	P	S	R	F

Figura 28b

63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87
S	G	S	G	S	G	T	D	F	T	L	K	I	S	R	V	E	A	E	D	L	G	V	Y	Y
S	G	S	G	S	G	T	D	F	T	L	N	I	H	P	V	E	E	E	D	A	A	T	Y	Y
T	G	S	G	S	G	T	D	F	T	L	T	I	S	S	V	Q	V	E	D	L	T	H	Y	Y
S	G	S	G	S	G	T	D	F	T	L	N	I	H	P	V	E	E	E	D	A	A	T	Y	Y
T	G	S	G	S	G	T	D	F	T	L	T	I	S	S	V	Q	A	E	D	L	A	V	Y	Y
S	G	S	G	S	G	R	D	Y	S	F	S	I	S	N	L	E	P	E	D	I	A	T	Y	Y
S	G	S	G	S	G	T	D	F	T	L	K	I	S	R	V	E	A	E	D	L	G	V	Y	Y
S	G	S	G	S	G	T	D	Y	T	L	T	I	S	S	L	Q	P	E	D	V	A	T	Y	F

Figura 28c

35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62			
K:ent																														
T	CG	TAC	C	TG	CAA	AAA	CCG	GGC	CAG	AGC	CCG	AAA	CTG	CTG	ATC	TAT	AAA	GTT	AGC	AAC	CCG	TTT	AGC	GGC	GTG	CCG	GAT	CGC	TTC	
T	GG	TAC	C	AG	CAG	AAA	CCG	GGC	CAG	CCG	CCG	AAA	CTG	CTG	ATC	TAT	CTG	GCG	AGC	AAC	CTG	GAA	AGC	GGC	GTG	CCA	GCC	CGC	TTT	
T	CG	TAC	C	AG	AAA	AAA	CCG	GAA	CAG	AGC	CCG	AAA	CTG	CTG	ATC	TAT	GGC	GCC	AGC	AAC	CCG	ATC	ATT	GGC	GTG	CCG	GAT	CGC	TTT	
T	GG	TAC	C	AG	CAG	AAA	CCG	GGC	CAG	CCG	CCG	AAA	CTG	CTG	ATC	TAT	GCC	GCC	AGC	AAC	CTG	GAA	AGC	GGC	ATT	CCA	GCC	CGC	TTT	
T	GG	TAC	C	AG	CAG	AAA	CCG	GGC	CAG	CCG	CCG	AAA	CTG	CTG	ATC	TAT	GGC	GCC	AGC	ACC	CCG	GAA	AGC	GGC	GTG	CCG	GAT	CGC	TTT	
T	CG	TAC	C	AG	CAT	AAA	CCG	GGC	AAA	GGC	CCA	CCG	AAA	CTG	CTG	ATT	CAT	TAT	ACC	AGC	ACC	CTG	CAA	CCG	GGC	ATT	CCG	AGC	CGC	TTT
T	GG	TAC	C	TG	CAA	AAA	CCG	GGC	CAG	AGC	CCG	CAG	AGC	CTG	CTG	ATT	TAT	TGG	GTG	AGC	AAC	CGT	GAA	AGC	GGC	GTG	CCG	GAT	CGT	TTT
T	GG	TAC	C	AG	CAG	AAA	CCG	GGC	AAA	GGC	CCG	AAA	CTG	CTG	ATT	TAT	AAC	GCC	AAC	AGC	CTG	CAA	ACC	GGT	GTG	CCG	AGC	CGT	TTT	

Uno de los dos aminoácidos (T o V) se puede encontrar en estos residuos (solo Vx1 y Vx3)
 Región de la secuencia recuperada del „Sequence Analysis Software“ (SAS)
 Posición de la región diversificada
 Variabilidad potencial de la longitud
 No hay aminoácidos en esta posición

Figura 28d

63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87
				</																				

Figura 29a

74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
K	I	S	R	V	E	A	E	D	L	G	V	Y	Y	C
N	I	H	P	V	E	E	E	D	A	A	T	Y	Y	C
T	I	S	S	V	Q	V	E	D	L	T	H	Y	Y	C
N	I	H	P	V	E	E	E	D	A	A	T	Y	Y	C
T	I	S	S	V	Q	A	E	D	L	A	V	Y	Y	C
S	I	S	N	L	E	P	E	D	I	A	T	Y	Y	C
K	I	S	R	V	E	A	E	D	L	G	V	Y	Y	C
T	I	S	S	L	Q	P	E	D	V	A	T	Y	F	C

Figura 29b

LCDR3												
LCDR3												
LCDR3												
93	94	95	96	97	98	99	00	01	02	03	04	
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E
X	X	X	X	X	T	F	G	G	T	K	L	E

50%Q	100%Q	20%G	30%S	34%S	25%Y	100%P	30%Y
10%F	20%S	15%T	24%H	20%V	20%L		
10%S	10%Y	15%N	14%E	10%L	10%P		
1.6%all	2.6%all	10%Y	1.0%all	10%F	10%F		
		1.6%all	1.6%all	1.3%all	1.6%all		

sin variabilidad de la longitud

Figura 29c

74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
							Ebol							
AAA	ATC	AGC	CGC	GTG	GAA	GCC	GAA	GAC	CTG	GGC	GTG	TAT	TAT	TGC
AAC	ATC	CAC	CDC	GTG	GAA	GAA	BAA	GAC	GCC	GCC	ACC	TAT	TAT	TGC
ACC	ATC	AGC	AGC	GTG	CAG	GTG	GAA	GAC	CTG	ACC	CAT	TAT	TAC	TGC
AAC	ATC	CAC	CCG	GTG	GAA	GAA	GAA	GAC	GCC	GCC	ACC	TAT	TAT	TGC
ACC	ATT	AGC	AGC	GTG	CAG	GCC	GAA	GAC	CTG	GCC	GTG	TAT	TAC	TGC
AGC	ATT	AGC	AAC	CTG	GAA	CCG	BAA	GAC	ATC	GCC	ACC	TAT	TAT	TGC
AAA	ATT	AGC	CGT	GTG	GAA	GCC	GAA	GAC	CTG	GGC	GTG	TAT	TAT	TGC
AGG	ATT	AGC	TCT	CTG	GAA	CCG	GAA	GAC	GTG	GGC	ACC	TAT	TTT	TGC

Figura 30d

HCDR2			HCDR2			HCDR2																															
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70				
GTG AAA	CAG CCG	CCA CGT	CSC GGC	GTC GAG	TGC GAT	GGC GGC	ATT GAT	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA
GTG AAA	CAG CCG	CCA CGT	CSC GGC	GTC GAG	TGC GAT	GGC GGC	ATT GAT	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA
GTT CSC	CAG AGC	CGG GGC	AAA GGT	CTC GAG	TGG GTC	GCC GTC	ATT TGG	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA
ATT CSC	CAG TTT	CGG GGC	AAA GGT	CTC GAG	TGG GTC	GCC GTC	ATT TGG	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA
GTT CSC	CAG ACC	CGG GGC	AAA GGT	CTC GAG	TGG GTC	GCC GTC	ATT TGG	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA
GTG AAA	CAG GGC	CCA CGT	CSC GGC	GTC GAG	TGC GAT	GGC GGC	ATT GAT	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA
ATT AAA	CAG CBT	CGG GGC	AAA GGT	CTC GAG	TGG GTC	GCC GTC	ATT TGG	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA
GTT CSC	CAG GCG	CGG GGC	AAA GGT	CTC GAG	TGG GTC	GCC GTC	ATT TGG	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA	GGC GGC	GGC GGC	ACC AAA	TAT AAC	GAA AAA	TTC AAA

Figura 31a

66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91			
K	A	T	L	T	V	D	K	P	S	S	T	A	Y	M	Q	L	S	S	L	T	S	E	D	S	A	V	Y	Y
V	T	L	T	V	D	K	K	P	S	S	T	A	Y	M	Q	L	S	S	L	T	S	E	D	S	A	V	Y	Y
R	L	S	I	S	K	D	N	S	K	S	Q	V	F	F	K	M	N	S	L	Q	A	D	D	T	A	I	Y	Y
R	I	S	I	T	R	D	T	S	K	N	Q	F	F	L	K	L	N	S	V	T	T	E	D	T	A	T	Y	Y
R	F	T	I	S	R	D	N	A	K	N	T	L	Y	L	Q	M	S	S	L	R	S	E	D	T	A	M	Y	Y
R	F	A	F	S	L	E	T	S	A	S	T	A	Y	L	Q	I	N	N	L	K	N	E	D	T	A	T	Y	F
K	A	T	F	T	V	D	K	S	S	S	T	A	Y	M	Q	L	S	S	L	T	P	E	D	T	A	V	Y	Y
R	F	T	I	S	R	D	N	A	K	S	T	L	Y	L	Q	M	D	S	L	R	S	E	D	T	A	T	Y	Y

Figura 31b

HCDR3											
HCDR3											
HSDR3											
57	56	55	54	53	52	51	50	49	48	47	46
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X
C	A	R	X	X	(X)	(X)	(X)	X	X	X	X

20%Y 20%Y 40%Y 15%Y 15%Y 15%Y 15%Y 15%Y 30%Y 30%Y 60%F 80%D 80%Y
 20%G 20%G 10%G 15%G 15%G 15%G 15%G 10%S 20%A 20%M 20%A 20%V
 10%D 10%D 5%D 5%L 5%R 5%R 5%R 5%R 2.6% all
 10%S 10%S 5%L 5%R 5%D 5%D 5%D 5%D
 10%R 10%R 5%R 5%D 5%N 5%N 5%N
 1.6% all 1.6% all 5%S 5%N 5%N 5%N
 1.6% all 2.1% all 2.1% all 2.1% all 2.1% all

Figura 31c

	65	66	68	70	71	72	73	74	75	76	77	78	79	80	81	82	82A	82B	82C	83	84	85	86	87	88	89	90	91	
AAA	GCC	ACC	CTG	ACC	GTG	GAT	AAA	CCG	AGC	AGC	ACC	GCC	TAT	ATG	CAA	CTG	AGT	AGC	CTG	ACC	AGC	GAA	GAT	AGC	GCC	GTG	TAT	TAT	
CG	G	GTG	ACC	CTG	ACC	GTG	GAT	AAA	CCG	AGC	AGC	ACC	GCC	TAT	ATG	CAA	CTG	AGT	AGC	CTG	ACC	AGC	GAA	GAT	AGC	GCC	GTG	TAT	TAT
CGC	CTG	AGC	ATC	AGC	AAA	GAT	AAC	AGC	AAA	AGC	CAG	GTG	TTC	TTC	AAA	ATG	AAC	AGC	CTG	CAA	GCC	GAT	GAT	ACC	GCC	ATC	TAT	TAT	
CGC	ATT	AGC	ATC	ACC	CGC	GAT	ACC	AGC	AAA	AAC	CAG	TTT	TTT	CTG	AAA	CTG	AAC	AGC	GTG	ACC	ACC	GAA	GAT	ACC	GCC	ACC	TAT	TAT	
CGC	TTT	ACC	ATT	AGC	CGC	GAT	AAC	GCC	AAA	AAC	ACC	CTG	TAT	CTG	CAA	ATG	AGC	AGC	CTG	CGC	AGC	GAA	GAT	ACC	GCC	ATG	TAT	TAT	
CGC	TTT	GCC	TTT	AGC	CTG	GAA	ACC	AGC	GCC	AGC	ACC	GCC	IAT	CTG	CAA	ATT	AAC	AAC	CTG	AAA	AAC	GAA	GAT	ACC	GCC	ACC	TAT	TTT	
AAA	GCG	ACC	TTT	ACC	GTG	GAT	AAA	AGC	AGC	ACC	GCG	TAT	ATG	CAG	CTG	TCC	AGC	CTG	ACC	CCG	GAA	GAT	ACC	GCG	GTG	TAT	TAC		
CGT	TTT	ACC	ATT	AGC	CGT	GAT	AAC	GCG	AAA	AGC	ACC	CTG	TAT	CTG	CAA	ATG	GAT	AGC	CTG	CGT	AGC	GAA	GAT	ACC	GCG	ACC	TAT	TAT	

Figura 32

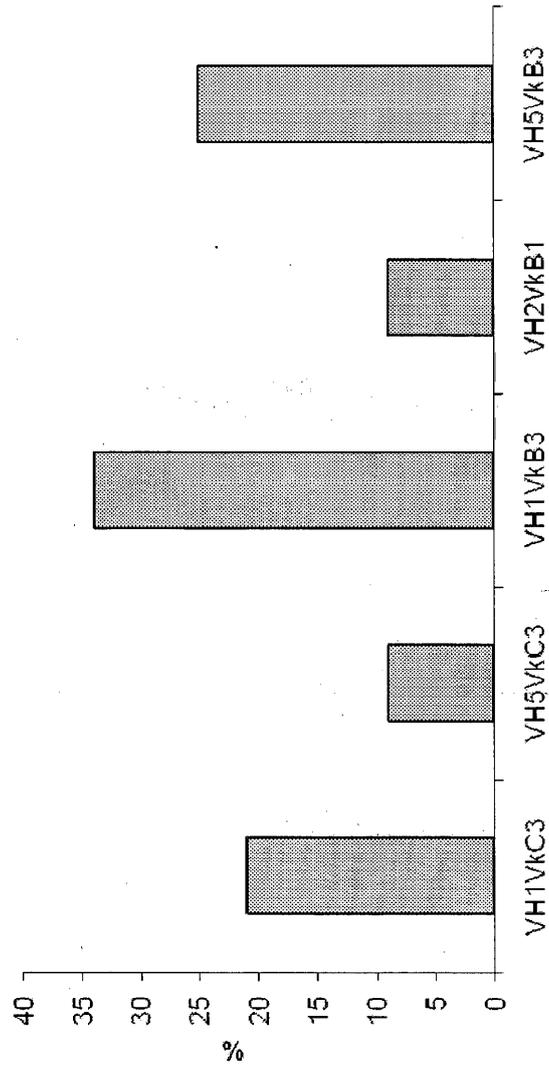


Figura 33

Intervalo de rendimiento de la expresión en *E. coli*

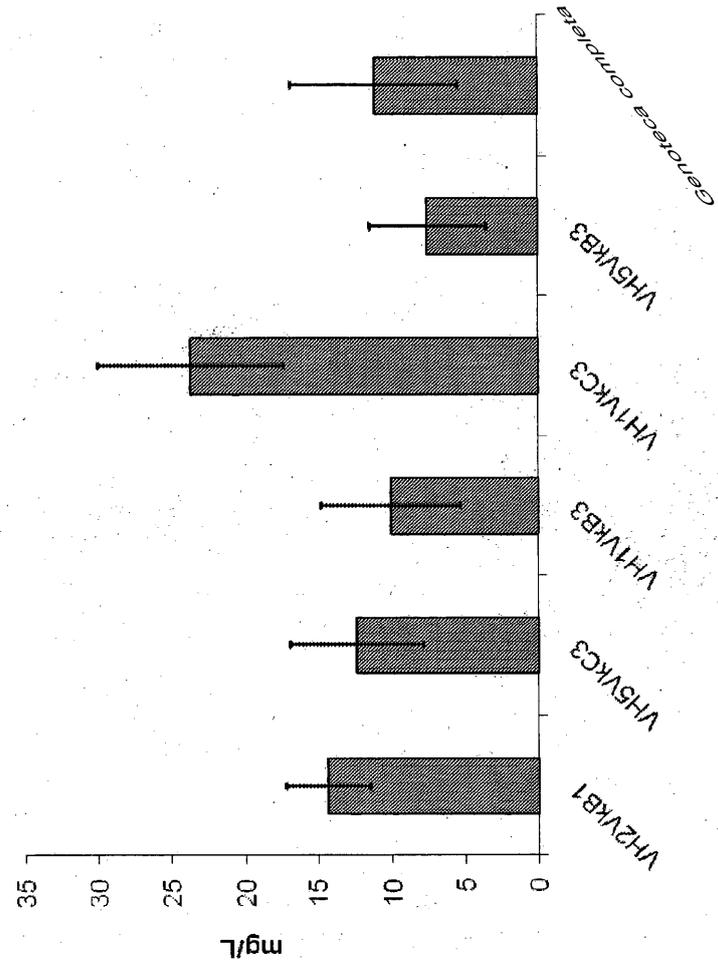


Figura 34

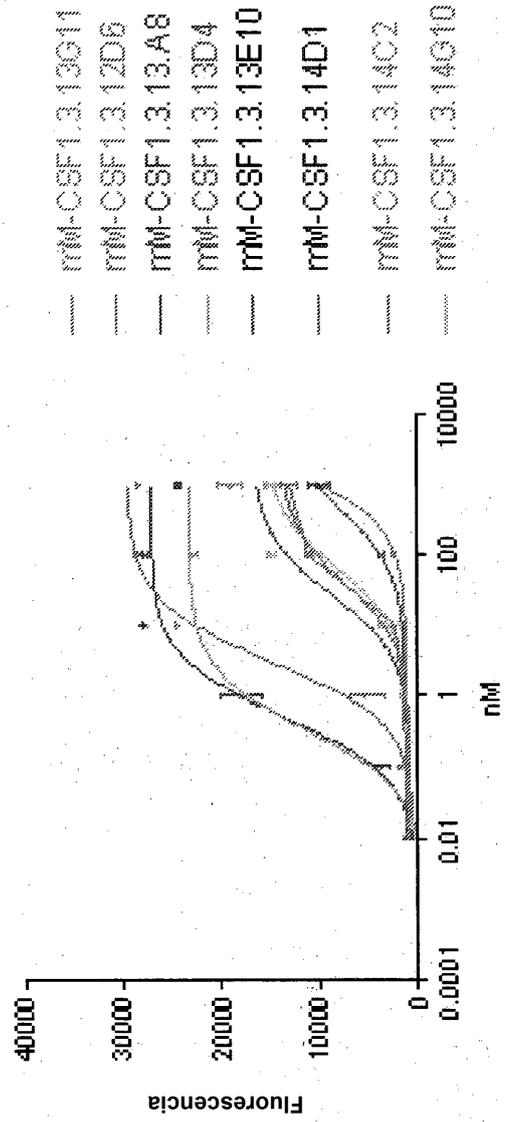


Figura 35

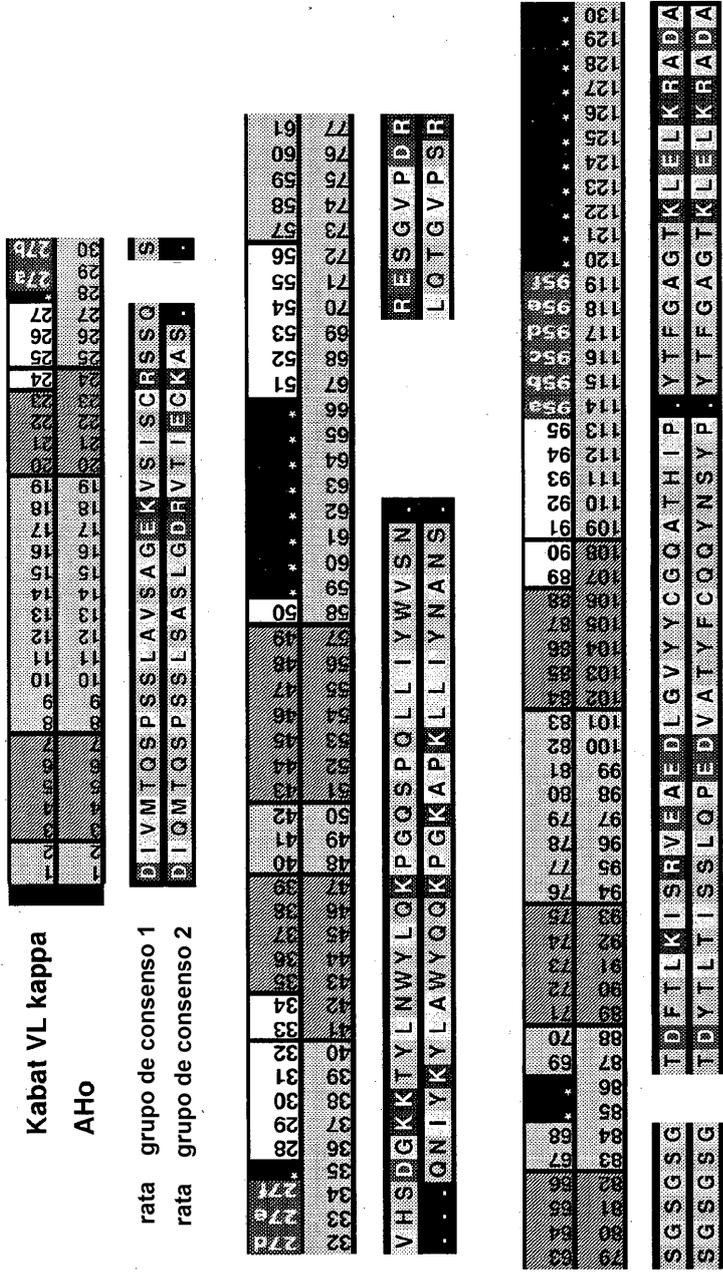


Figura 37

J00746|GKJ1*01|Rattus norvegicus_Louvain
 J00746|GKJ2-1*01|Rattus norvegicus_Louvain
 J00746|GKJ2-2*01|Rattus norvegicus_Louvain
 J00746|GKJ2-3*01|Rattus norvegicus_Louvain
 J00746|GKJ4*01|Rattus norvegicus_Louvain
 J00746|GKJ5*01|Rattus norvegicus_Louvain

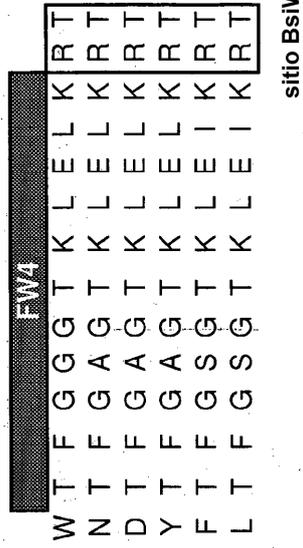


Figura 38

Presentación relativa normalizada de Fab de rata
sobre fagos
Datos procedentes de MOR y NVS incluidos

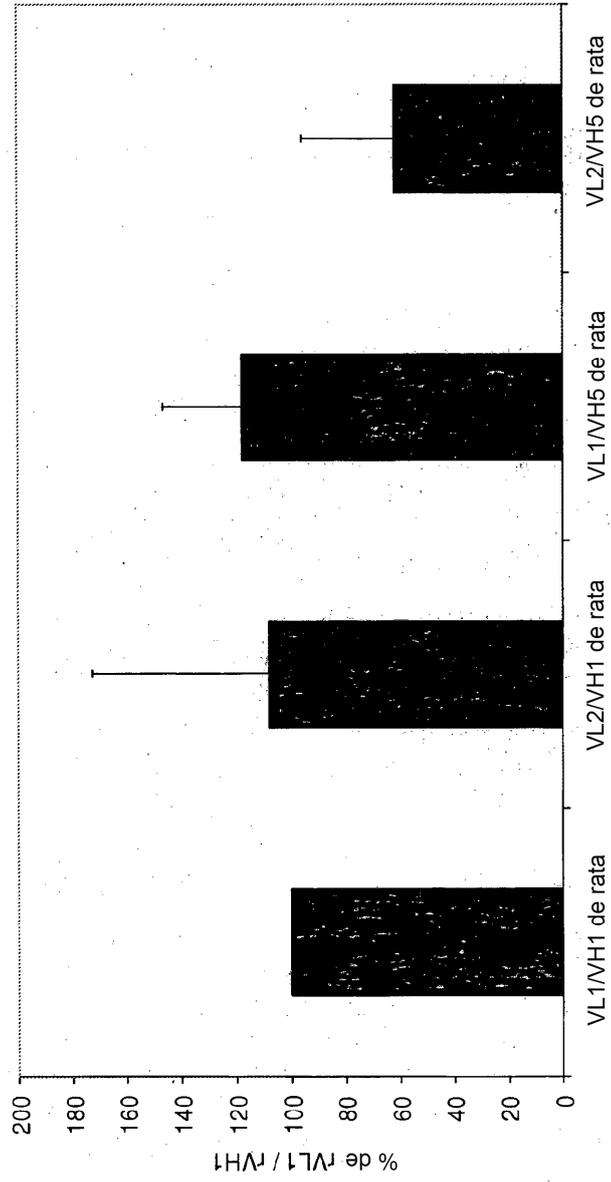


Figura 39

expresión de Fab de rata

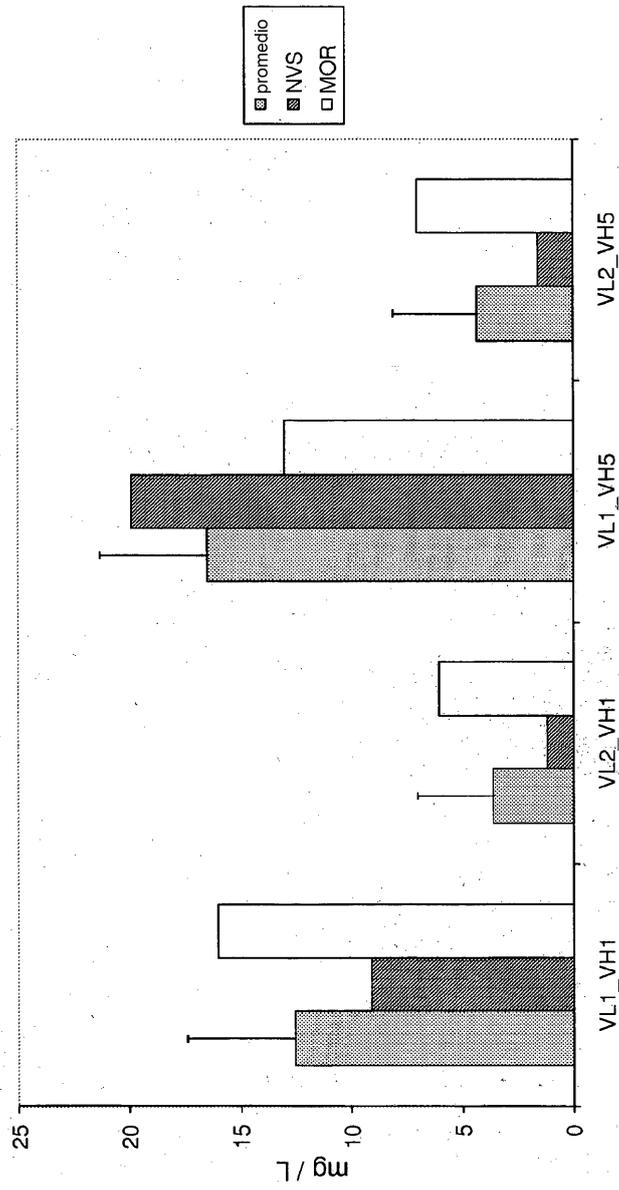


Figura 40

Expresión de IgG2b de rata

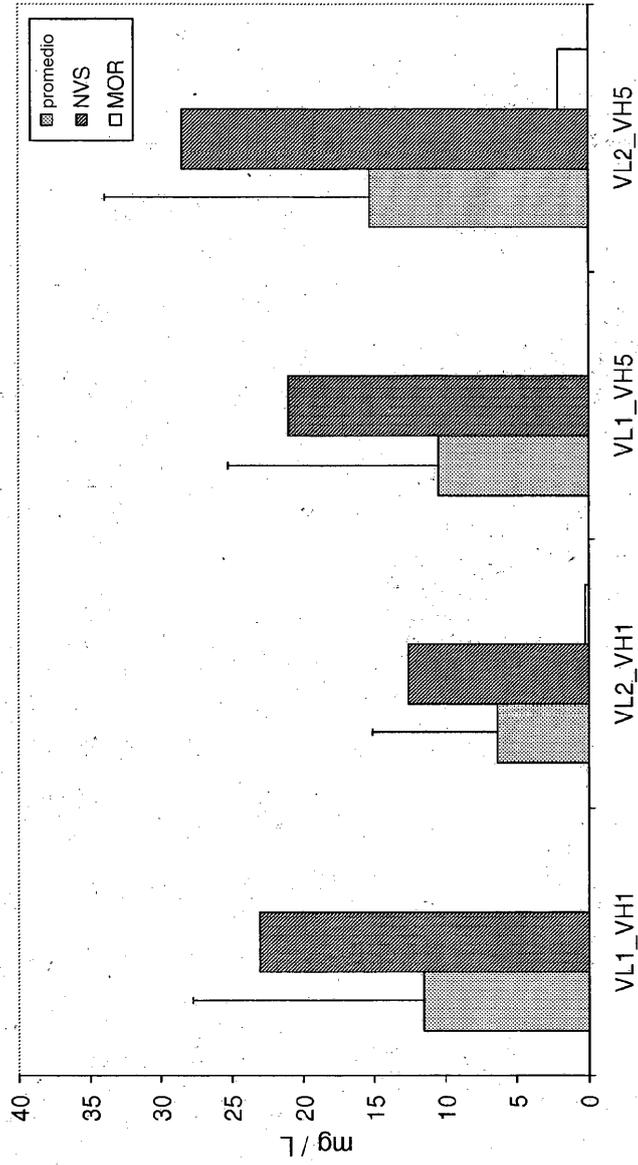


Figura 41

Expresión de IgG2c de rata

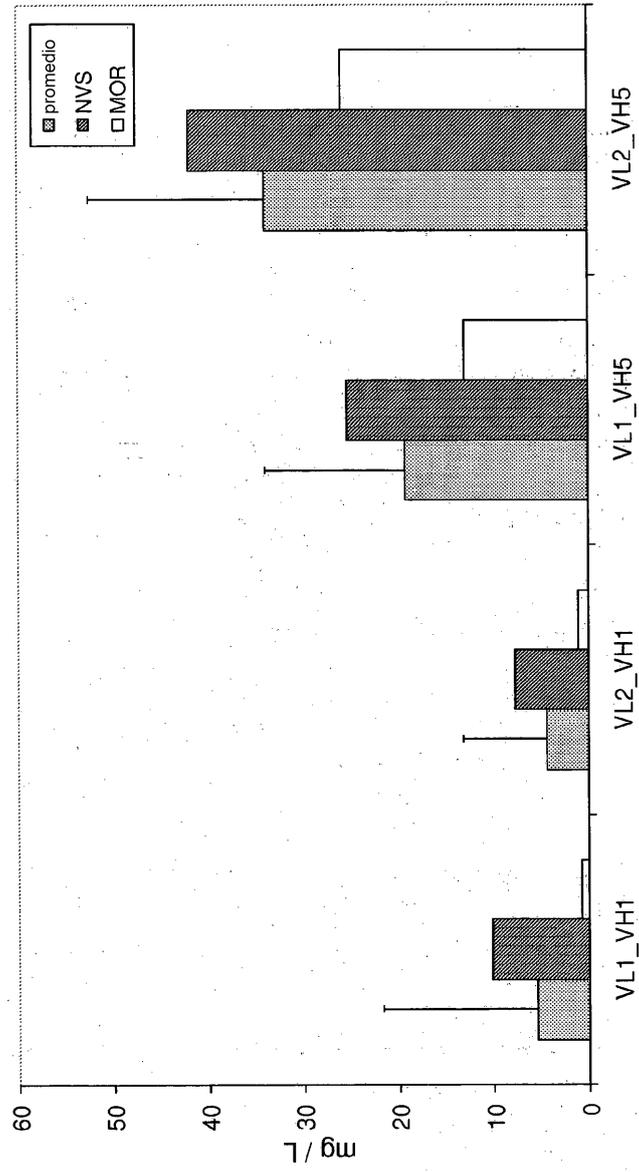


Figura 42

	formato pMx9 Fab_FH		formato pMx31_Displ.		Formato IgG							
	Tasa de present. MOR (si/no)	rendim. NVS [mg/L] con Iso	rendim. NVS [mg/L] sin Iso	rendim. MOR [mg/L] cult.	porción monómero SEC [%]	Tm [°C] con Iso	Tm [°C] sin Iso	rigG2b título MOR [mg/L] cult.	rigG2c título MOR [mg/L] cult.	rigG2b título IgG NVS	rigG2c título IgG NVS	rigG2b IgG DAS Tm [°C]
VL1_VH1		5,0	9,1	16	100	65,1	65,1	n.a.	0,8	23	66,5	70,2
VL2_VH1		0,0	1,2	6	100	-	29,3	0,2	1,1	12,6	67,3	73,9
VL1_VH5		12,6	19,9	13	100	67,7	67,9	0	13	21	67,9	72,5
VL2_VH5		0,0	1,6	7	100	-	19,0	2,1	2,6	28,5	70,7	74,7

Figura 43

Estabilidad térmica de los genes maestros de rata

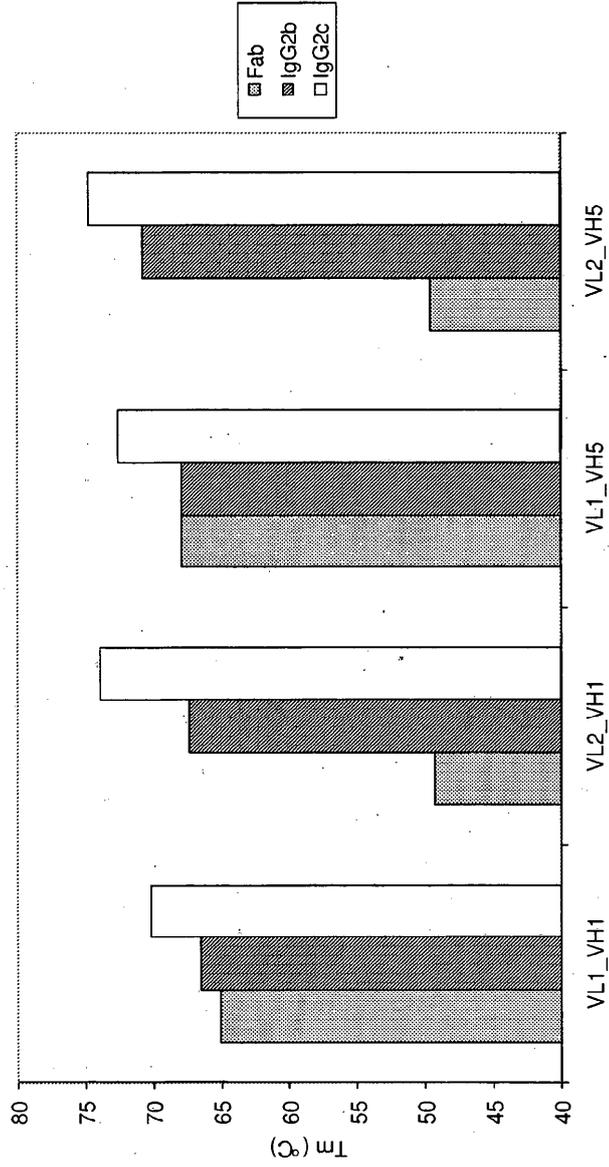


Figura 44

rVH5

50	51	52	52a	52b	52c	55	56	57	58
T	I	70S	Y	D	G	S	60S	T	Y
		30T					20N		
		W17					20T		
							W18		

Figura 45

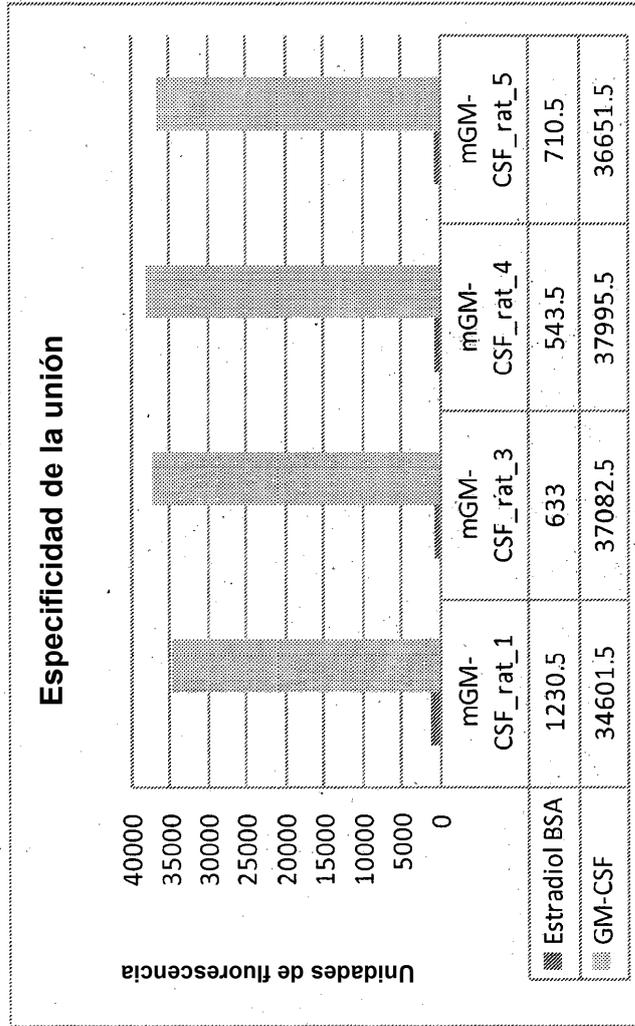
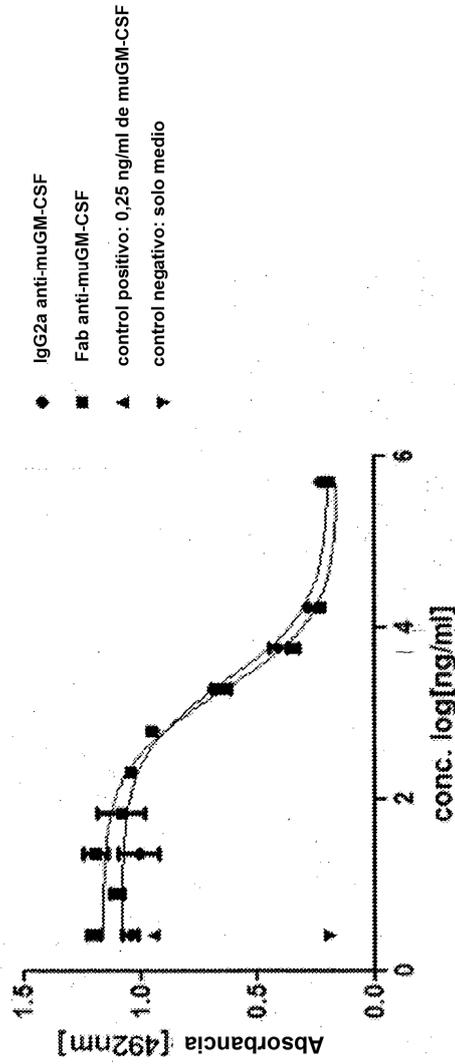


Figura 46

Ensayo de proliferación de FDCP-1 con un anticuerpo anti-mGM-CSF murino en dos formatos diferentes – Fab e IgG2a (n: 2:100806_14ORA39)



	IgG2a anti-muGM-CSF	Fab-MH anti-muGM-CSF
EC50[ng/ml]	2321	1732
EC50[nM]	15.4	346