



- (51) **International Patent Classification:**  
A01K 67/027 (2006.01) C07K 16/46 (2006.01)  
C12N 15/85 (2006.01)
- (21) **International Application Number:**  
PCT/GB2012/052298
- (22) **International Filing Date:**  
18 September 2012 (18.09.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**

1116122.1	19 September 2011 (19.09.2011)	GB
1116120.5	19 September 2011 (19.09.2011)	GB
1203257.9	24 February 2012 (24.02.2012)	GB
1204592.8	15 March 2012 (15.03.2012)	GB
1205702.2	29 March 2012 (29.03.2012)	GB
1208749.0	18 May 2012 (18.05.2012)	GB
1211692.7	2 July 2012 (02.07.2012)	GB

bridge, Cambridgeshire CB22 3AT (GB). **CLUBE, Jasper** [GB/GB]; c/o Kymab Limited, Meditrina (B260), Babraham Research Campus, Cambridge, Cambridgeshire CB22 3AT (GB). **ENGLAND, Nicholas** [GB/GB]; c/o Kymab Limited, Meditrina (B260), Babraham Research Campus, Cambridge CB22 3AT (GB).

(74) **Agents:** **STEPHEN, Robert** et al.; Olswang LLP, 90 High Holborn, London, Greater London WC1V 6XX (GB).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(71) **Applicant** (for all designated States except US): **KYMAB LIMITED** [GB/GB]; Meditrina (B260), Babraham Research Campus, Cambridge, Cambridgeshire CB22 3AT (GB).

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

- (72) **Inventors; and**
- (71) **Applicants** (for US only): **FRIEDRICH, Glenn** [GB/GB]; c/o Kymab Limited, Meditrina (B260), Babraham Research Campus, Cambridge, Cambridgeshire CB22 3AT (GB). **LEE, E-Chiang** [US/GB]; c/o Kymab Limited, Meditrina (B260), Babraham Research Campus, Cam-

[Continued on next page]

(54) **Title:** MANIPULATION OF IMMUNOGLOBULIN GENE DIVERSITY AND MULTI-ANTIBODY THERAPEUTICS

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
122582	IGHV1-69*01, HVL051	CAG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG	CTG
217366	IGHV1-69*02, y109617901	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
392340	IGHV1-69*03, ST95A5	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
393130	IGHV1-69*04, H31261	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
367892	IGHV1-69*05, H31261.2	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
122583	IGHV1-69*06, H31261a	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
628276	IGHV1-69*07, H31261b	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
614339	IGHV1-69*08	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
114301	IGHV1-69*09	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
214370	IGHV1-69*10	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
214376	IGHV1-69*11	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
214384	IGHV1-69*12	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
214216	IGHV1-69*13	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Alignment of 13 IGHV1-69 alleles showing the variable (V) coding region only. Nucleotides that differ from VH1-69 allele \*01 are indicated at the appropriate position whereas identical nucleotides are marked with a dash. Where nucleotide changes result in amino acid differences, the encoded amino acid is shown above the corresponding triplet. Boxed regions correspond to CDR1, CDR2 and CDR3 as indicated.

Figure 5 (part 1 of 4)

(57) **Abstract:** The invention provides improved non-human vertebrates and non-vertebrate cells capable of expressing antibodies comprising human variable region sequences. The present invention is directed to the provision of long HCDR3s from non-human vertebrates and cells. The present invention is also directed to the provision of novel V, D and J pairings in immunoglobulin heavy and light chain loci. Novel, biased antibody diversities and potentially expanded diversities are provided. The invention also provides for novel and potentially expanded diversity or diversity that is biased towards variable gene usage common to antibodies useful for treating and/or preventing certain diseases or conditions, such as infectious diseases. The invention also provides methods of generating antibodies using such vertebrates, as well as the antibodies per se, therapeutic compositions thereof and uses.

WO 2013/041846 A2

**Published:**

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

**MANIPULATION OF IMMUNOGLOBULIN GENE DIVERSITY AND**

**MULTI-ANTIBODY THERAPEUTICS**

5 **FIELD OF THE INVENTION**

The present invention relates to the provision of antibodies with long HCDR3 lengths. The present invention is also directed to the provision of novel V, D and J pairings in immunoglobulin heavy and light chain loci. Novel, biased antibody diversities and potentially expanded diversities are provided. The invention also provides for novel and potentially expanded diversity or diversity that is biased  
10 towards variable gene usage common to antibodies useful for treating and/or preventing certain diseases or conditions, such as infectious diseases. This ability to bias the antibody repertoire also provides methods of simplifying the production of antibody mixtures, such as polyclonal antibody therapeutics useful for the treatment and/or prevention of infectious diseases where a polyclonal approach to target multiple pathogen antigens is desirable. To this end, the present invention also  
15 provides bispecific antibodies that are capable of binding to more than one antigen (eg, multiple infectious antigens expressed by the same pathogen), thus providing advantages (such as manufacturing, dosing and administration advantages) not possible with polyclonal antibody mixtures.

20 The present invention provides vertebrates and cells, such as transgenic mice or rats or transgenic mouse or rat cells. Furthermore, the invention relates to methods of using the vertebrates to isolate antibodies or nucleotide sequences encoding antibodies. Antibodies, nucleotide sequences, pharmaceutical compositions and uses are also provided by the invention.

25 **BACKGROUND**

The state of the art provides non-human vertebrates (eg, mice and rats) and cells comprising transgenic immunoglobulin loci, such loci comprising human variable (V), diversity (D) and/or joining (J) segments, and optionally human constant regions. Alternatively, endogenous constant regions of the host vertebrate (eg, mouse or rat constant regions) are provided in the transgenic loci. Methods  
30 of constructing such transgenic vertebrates and use of these to generate antibodies and nucleic

acids thereof following antigen immunisation are known in the art, eg, see US7501552 (Medarex),  
US5939598 (Abgenix), US6130364 (Abgenix), WO02/066630 (Regeneron), WO2011004192 (Genome  
Research Limited), WO2009076464, WO2009143472 and WO2010039900 (Ablexis), the disclosures  
of which are explicitly incorporated herein. Such transgenic loci in the art include varying amounts  
5 of the human V(D) J repertoire.

Existing transgenic immunoglobulin loci tend to be based on a single human DNA source. The  
potential diversity of human antibody variable regions in non-human vertebrates bearing such  
transgenic loci is thus confined by the repertoire used.

It would be desirable to provide for novel and potentially expanded repertoire and diversity of  
10 human variable regions in transgenic immunoglobulin loci and non-human vertebrates harbouring  
these, as well as in antibodies produced following immunisation of such animals.

### **SUMMARY OF THE INVENTION**

The present inventors have discovered, by way of construction of transgenic non-human  
vertebrates, immunisation, antibody heavy chain collection and deep bioinformatics analysis, how to  
15 rationally design for VH domains, heavy chains and antibodies having long HCDR3s. These are useful  
for addressing antigens (such as infectious disease pathogen antigens, receptors and enzyme clefts)  
where a longer CDR better addresses the target.

The present inventors also realised the possibility of providing combinations of V, D and J gene  
20 segments in new ways to provide synthetic gene segment combinations in immunoglobulin loci that  
are not found in nature or in state-of-the-art loci. The inventors realised the importance of this to  
provide for novel and potentially expanded repertoire and diversity of human variable regions in  
transgenic immunoglobulin loci and non-human vertebrates harbouring these, as well as in  
antibodies produced following immunisation of such animals. In one aspect, the inventors realised  
25 that it would be desirable to bias the novel repertoire for the production of antibodies having  
improved affinity and/or biophysical characteristics, and/or wherein the range of epitope  
specificities produced by means of such repertoire is novel, provides for antibodies to epitopes that  
have hitherto been intractable by prior transgenic immunoglobulin loci or difficult to address. For  
example, the inventors envisaged a specific application to bias the novel repertoire for the  
30 production of antibodies useful in the therapy and/or prevention of infectious disease.

To this end, in a first configuration of the invention, there is provided

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

(a) An immunoglobulin heavy chain locus comprising one or more human V gene segments, one or more human D gene segments and one or more human J gene segments upstream of a constant region; optionally wherein the heavy chain locus is according to any configuration of the invention described below; and

(b) An immunoglobulin light chain locus comprising either

(i) one or more human VH gene segments and one or more human J gene segments upstream of a constant region (optionally a rearranged VHJLCL or VHJλCL, wherein the CL is Cλ or Cκ); or

(ii) one or more human VL gene segments, one or more human D gene segments and one or more human JH gene segments upstream of a constant region (optionally a rearranged VLDJHCL or VλDJHCL, wherein the CL is Cλ or Cκ); or

(iii) one or more human VL gene segments selected from the group consisting of: a VλII gene family member, VλVII 4A, VλII 2.1, VλVII 4A, a Vλ1 gene family member, a Vλ3 gene family member, IGLV1S2, Vλ3-cML70, Ialh2, Ialvl, Ia3h3, Kv325, a VκI gene family member, κI-15A (KL012), VκII family member, a VκIII family member, a VκI gene family member, κI-15A (KL012), VκII A2 (optionally the A2a allele), Vκ A27 (Humkv325) and a gene segment at least 80% identical thereto, and one or more human JL gene segments upstream of a constant region;

Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains produced by recombination of the light chain locus.

In one aspect,

in (b)(i) the V gene segment repertoire of the light chain locus comprises or consists of one or more VH gene segments selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1,

VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto; or

in (b)(iii) the light chain locus V gene segment repertoire consists of one VL gene segment type (optionally and one or mutants thereof), wherein the VL gene segment is selected from said group of

5 VL gene segments.

In a second configuration of the present invention, there is provided

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

10 (a) An immunoglobulin heavy chain locus comprising one or more human V gene segments, one or more human D gene segments and one or more human J gene segments upstream of a constant region; and

(b) (i) An unrearranged immunoglobulin light chain locus comprising one or more human VH gene segments and one or more human J gene segments upstream of a constant region, wherein  
15 each human VH gene segment is a human gene segment identical to (or mutant of) a human VH gene segment used to produce a rearranged VDJ encoding a heavy chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused  
20 or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism); or

(ii) An immunoglobulin light chain locus comprising a rearranged VJ region or VDJ region upstream of a constant region, wherein the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a  
25 human D gene segment with a human VH gene segment that is identical to (or mutant of) the human VH gene segment used to produce a rearranged VDJ encoding a heavy chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition  
30 caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);

(c) Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments or VJ or VDJ in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains derived from the light chain locus;

(d) Optionally when (b)(i) applies, each said VH gene segment in the light chain locus is selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto;

(e) Optionally when (b)(ii) applies, the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto.

In one aspect, the V gene segment repertoire of the light chain locus comprises or consists of one human VH gene segment; optionally germline VH and one or more polymorphic variants thereof, eg, where each polymorphic variant differs from the germline VH nucleotide sequence by 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 positions. In one aspect, the V gene segment repertoire of the light chain locus comprises or consists of human VH1-69 gene segment; optionally germline VH1-69 and one or more polymorphic variants thereof, eg, where each polymorphic variant differs from the germline VH1-69 nucleotide sequence by 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 positions. An example of constructing an immunoglobulin locus comprising VH1-69 and polymorphic variants is given below. By using a particular gene segment (eg, one commonly found in human antibodies raised in humans against an infection or other condition) and one or more polymorphic variants thereof, it is possible to provide a repertoire of genes and yet still bias the antibody gene repertoire to a gene segment that is relevant to the disease (eg, an infectious disease, such as a bacterial or viral disease, eg, influenza). This provides a useful pool of genes from which to ultimately generate and isolate a lead antibody for a therapeutic/prophylactic against the disease in question. In an example, the polymorphic

variants are natural variants seen in human beings or human populations. The skilled person will know of sources of human antibody gene sequences, such as IMGT ([www.imgt.org](http://www.imgt.org)), GenBank ([www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)) and the 1000 Genomes databases ([www.1000genomes.org](http://www.1000genomes.org)). Bioinformatics tools for database manipulation are also readily available and known to the skilled person, eg, as publicly available from the 1000 Genomes Project/EBI ([www.1000genomes.org](http://www.1000genomes.org))

In another aspect, the genome of said vertebrate or cell is homozygous for light chain locus (b)(i) or (ii); optionally wherein:

- 10 - the V gene segment repertoire of the light chain loci consists of one or more human VH gene segments selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto; or
- 15 - the recombined VJ or VDJ repertoire of the light chain loci consists of sequences identical to one or more nucleotide sequences produced by the recombination of a human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 20 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto, with a human J gene segment and optionally a human D gene segment.

In another aspect, each immunoglobulin light chain locus of said vertebrate or cell is according to (b)(i) and comprises only a single human VH gene segment selected from the group consisting of: a 25 VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain so that all light chain loci comprise the 30 same, single human VH gene segment.



The invention provides a first method of isolating an antibody that binds a predetermined antigen, the method comprising

- (a) providing a vertebrate (optionally a mouse or rat) according to any preceding configuration or aspect;
- 5 (b) immunising said vertebrate with said antigen (optionally wherein the antigen is an antigen of an infectious disease pathogen);
- (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes expressing antibodies that bind to the antigen;
- (d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by  
10 producing hybridomas therefrom; and
- (e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes.

A second method is provided comprising carrying out the first method and the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally  
15 exchanging the heavy chain constant region nucleotide sequence of the antibody with a nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host.

20 An aspect provides method of producing a polyclonal antibody mixture, the method comprising carrying out the first method by separately immunising first and second vertebrates (optionally first and second mice or first and second rats) with antigen and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or  
25 (i) and (iii)):

- (i) the vertebrates are immunised with the same antigen or different antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));

(ii) prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene) and optionally the identical J repertoire; optionally the light chain loci of the mammals are identical prior to immunisation;

(iii) prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.

An aspect provides method of producing a polyclonal antibody mixture, the method comprising carrying out the first method by immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):

- (i) the antigens are expressed by the same pathogenic organism (or a family member thereof);
- (ii) prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene) and optionally the identical J repertoire; optionally the light chain loci of the mammals are identical prior to immunisation;
- (iii) prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.

20

An aspect provides method of producing host cells capable of expressing a polyclonal antibody mixture, the method comprising, in the second method:-

- (a) immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));
- (b) isolating nucleic acid encoding first and second anti-antigen antibodies from B lymphocytes from said vertebrates;
- (c) determining the nucleotide sequences of the heavy and light chain variable regions of the first antibody;

- (d) determining the nucleotide sequence of the heavy variable region and optionally the light chain variable region of the second antibody;
- (e) inserting the heavy chain variable region coding sequence of each antibody into a heavy chain expression vector; optionally wherein the constant region coding sequence of each heavy chain is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (f) inserting the light chain variable region coding sequence of the first antibody into a light chain expression vector; optionally wherein the constant region coding sequence of the light chain of the first antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (g) optionally inserting the light chain variable region coding sequence of the second antibody into a light chain expression vector; optionally wherein the constant region coding sequence of the light chain of the second antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region; and
- (h) introducing each expression vector into a host cell and co-expressing antibody chains in a mixture of said host cells to produce antibodies, each antibody comprising one or both of said heavy chain variable regions and a light chain; optionally wherein the expression vectors are introduced together into the same host cell (eg, a CHO or HEK293 cell) so that the cell is capable of expressing antibody light chains and heavy chains, such that the cell or a plurality of the host cells express antibodies, each comprising one or both of said heavy chain variable regions and a light chain;
- (i) optionally:
- prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene segment) and optionally the identical J repertoire (optionally a single J gene segment); optionally the light chain loci of the vertebrates are identical prior to immunisation; or
- prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.

The invention also provides a method of producing a monoclonal or polyclonal antibody mixture, optionally for use in medicine, optionally for the treatment and/or prevention of an infectious

disease, wherein optionally wherein each antibody binds an antigen of an infectious disease pathogen, preferably the same antigen. The invention also provides the use of an isolated, monoclonal or polyclonal antibody, or a mutant or derivative antibody thereof in the manufacture of a medicament for the treatment and/or prevention of an infectious disease, optionally wherein the  
5 infectious disease is a disease caused by a bacterial or viral pathogen.

The invention further provides an isolated antibody (eg, IgG-type antibody) obtainable or obtained by a method of the invention, or a mutant or derivative antibody thereof wherein (i) the isolated antibody comprises two copies of the heavy chain variable region of said first antibody paired with  
10 two copies of the light chain variable region of said first antibody; or (ii) the isolated antibody comprises two copies of the heavy chain variable region of said second antibody paired with two copies of the light chain variable region of said first antibody; or (iii) the isolated antibody is a bispecific antibody comprising one copy of the heavy chain variable region of said first antibody paired with a copy of the light chain variable region of the first antibody, and one copy of the heavy  
15 chain variable region of said the antibody paired with a copy of the light chain variable region of the first antibody, optionally wherein the bispecific antibody binds to said first and second antigens are expressed by the same pathogenic organism (or a family member thereof; optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease.

20 In an aspect, there is provided a nucleotide sequence encoding an antibody of the invention, optionally wherein the nucleotide sequence is part of a vector.

In an aspect, there is provided a pharmaceutical composition comprising the antibody or antibodies of the invention and a diluent, excipient or carrier.

25

In a third configuration of the invention, there is provided

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

(a) An immunoglobulin heavy chain locus comprising either:-

(i) one or more human VL gene segments, one or more human D gene segments and one or more human J gene segments upstream of a constant region (optionally a rearranged VLDJHCH or V $\lambda$ DJHCH); or

(ii) one or more human VH gene segments selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical; one or more human D gene segments and one or more human JH gene segments upstream of a constant region; and

10 (b) An immunoglobulin light chain locus comprising one or more human V gene segments and one or more human J gene segments upstream of a constant region, optionally wherein the light chain locus is according to (b)(i) or (b)(ii) of the first configuration of the invention;

Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains produced by recombination of the light chain locus.

In a fourth configuration of the invention, there is provided

20 A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

(a) (i) An unrearranged immunoglobulin heavy chain locus comprising one or more human VL gene segments, one or more human D gene segments and one or more JH gene segments upstream of a constant region, wherein each human VL gene segment is a human gene segment identical to (or mutant of) a human VL gene segment used to produce a rearranged VJ encoding a light chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism); or

30 (ii) An immunoglobulin heavy chain locus comprising a rearranged VJ region or VDJ region upstream of a constant region, wherein the nucleotide sequence of the recombined region is identical to a

nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VL gene segment that is identical to (or mutant of) the human VL gene segment used to produce a rearranged VJ encoding a light chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);

(b) An immunoglobulin light chain locus comprising one or more human V gene segments and one or more human J gene segments upstream of a constant region; and

(c) Wherein the gene segments in the light chain locus are operably linked to the constant region thereof, and the gene segments or VJ or VDJ in the heavy chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising light chains produced by recombination of the light chain locus and heavy chains derived from the heavy chain locus;

(d) Optionally when (a)(i) applies, each said VL gene segment in the heavy chain locus is selected from the group consisting of a VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ I gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, IalH2, IalVl, Ial3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto;

(e) Optionally when (a)(ii) applies, the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VL gene segment selected from the group consisting of a VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ I gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, IalH2, IalVl, Ial3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto.

In one aspect of the fourth configuration of the invention, the genome of said vertebrate or cell is homozygous for heavy chain locus (a)(i) or (ii); optionally wherein:

- the V gene segment repertoire of the heavy chain loci consists of one or more human VL gene segments selected from the group consisting of a VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ I gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, I $\lambda$ h2, I $\lambda$ l $\nu$ l, I $\lambda$ 3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto; or

- the recombined VJ or VDJ repertoire of the heavy chain loci consists of sequences identical to one or more nucleotide sequences produced by the recombination of a human VL gene segment selected from the group consisting of a VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ I gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, I $\lambda$ h2, I $\lambda$ l $\nu$ l, I $\lambda$ 3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto with a human J gene segment and optionally a human D gene segment.

The invention provides a monoclonal or polyclonal antibody composition prepared by immunisation of at least one vertebrate (eg, mouse or rat) according to any preceding configuration or aspect with an antigen, optionally wherein the antigen is an antigen of an infectious disease pathogen, optionally wherein the same antigen is used to immunise all the vertebrates; optionally wherein the antibody or antibodies are IgG-type.

The invention also provides an isolated chimaeric antibody for treating and/or preventing an infectious disease or condition, the antibody comprising a non-human vertebrate (optionally a mouse or rat) heavy chain constant regions and human variable regions that bind an antigen of an infectious disease pathogen, wherein the antibody is obtainable or obtained in a method comprising immunisation of a non-human vertebrate of the invention with said antigen.

The invention also provides an isolated human antibody for treating and/or preventing an infectious disease or condition, the antibody comprising human heavy chain constant regions and human variable regions that bind an antigen of an infectious disease pathogen, wherein the antibody is

obtainable or obtained in a method comprising affinity maturation of antibody variable regions in vivo in a transgenic non-human vertebrate (eg, mouse or rat) when said variable regions are operably linked to heavy chain constant regions of said vertebrate (eg, mouse or rat heavy chain constant regions) by (a) immunisation of a vertebrate of the invention with said antigen, (b) isolation of nucleic acid encoding a chimaeric antibody according to the invention, (c) replacing the nucleotide sequences of the nucleic acid that encode the non-human vertebrate heavy chain constant regions with nucleotide sequence encoding human heavy chain constant regions to produce nucleic acid encoding a human antibody; (d) expressing the human antibody in vitro (optionally from CHO or HEK293 cells harbouring the human nucleic acid) and (e) isolating the human antibody (optionally with further affinity maturation of the antibody and/or producing a derivative thereof).

An aspect provides a mixture of first and second human antibodies, each antibody being capable of binding to an antigen of an infectious disease pathogen (optionally wherein the first antibody binds a first antigen and the second antibody binds a second antigen, said antigens being from the same pathogen; or wherein the antigens are the same). In an embodiment, a common light chain is used which enables simplified manufacture of the antibody mixture. Thus, there is provided in the mixture, the light chain amino acid sequence of the first antibody that is identical to the light chain amino acid sequence of the second antibody, or has up to 15 amino acid changes therefrom.

The invention further provides a host cell comprising one or more expression vectors encoding 3 or more first and second antibody heavy and light chains.

In a fifth configuration of the invention, there is provided

A synthetic immunoglobulin locus comprising one or more variable and J gene segments (and optionally one or more D gene segments) operably linked 5' of a constant region, wherein the locus comprises a 5' to 3' V(D)J arrangement selected from the group consisting of immunoglobulin locus can be constructed with one or more of the following arrangements (5' to 3'):-

(a) [V (heavy, lambda or kappa)] – [two-turn RSS] --- [one-turn RSS] – [D] – [JH], wherein said RSSs are in an opposite orientation;



- (b) [VH] – [D] – [two-turn RSS] --- [one-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (c) [VH] – [D] – [one-turn RSS] --- [two-turn RSS] – [J kappa], wherein said RSSs are in an opposite orientation;
- 5 (d) [VH or V kappa] - [two-turn RSS] --- [one-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (e) [V kappa] - [one-turn RSS] --- [two-turn RSS] – [JH or J lambda], wherein said RSSs are in an opposite orientation;
- (f) [V (heavy, lambda or kappa)] – [one-turn RSS] --- [two-turn RSS] – [D] – [JH], wherein said RSSs  
10 are in an opposite orientation;
- (g) [VH] – [D] – [one-turn RSS] --- [two-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (h) [VH] – [D] – [two-turn RSS] --- [one-turn RSS] – [J kappa], wherein said RSSs are in an opposite orientation;
- 15 (i) [VH or V kappa] - [one-turn RSS] --- [two-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (j) [V kappa] - [two-turn RSS] --- [one-turn RSS] – [JH or J lambda], wherein said RSSs are in an opposite orientation.

20 In a sixth configuration, the invention also provides means for generating VH domains, heavy chains and antibodies having a long HCDR3 length. In this context, the invention provides:-

A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin D gene segment repertoire that is  
25 biased to the human D2 and/or D3 family or biased to one, more or all human D gene segments selected from the group D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.

A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell), optionally according to any preceding claim, whose genome comprises a human  
30 immunoglobulin VH gene segment repertoire that is biased to one, more or all of gene segments selected from the group VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.

A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin JH gene segment repertoire that is  
 5 biased to human JH6.

A monoclonal or polyclonal antibody composition or a population of antibody-producing cells for producing such composition, wherein the composition or population is prepared by immunising at least one vertebrate according to any preceding claim with an antigen, wherein the antibody or  
 10 antibodies have human heavy chain variable regions comprising non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences; wherein the composition comprises at least one antigen-  
 15 specific antibody having a HCDR3 length of at least 20 amino acids (according to IMGT).

A repertoire of antibody heavy chains (eg, provided by antibodies) comprising one or more heavy chains whose variable domain HCDR3 has a length of at least 20 amino acids (according to IMGT) and derived from the recombination of a human VH, D and JH, wherein  
 20 the VH is selected from the group  
 VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01 and  
 the D is selected from the group  
 D2-2\*02, D3-9\*01, D3-10\*01 and D3-22\*01, or  
 25 D2-2\*02, D3-9\*01 and D3-10\*01, or  
 D3-9\*01 and D3-10\*01, or  
 D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or  
 D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01, or  
 D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-  
 30 17\*01, D6-13\*01 and D6-19\*01, or  
 D1-26, D2-2, D3-10 and D6-19, or  
 D2-2, D3-9 and D3-10;  
 and optionally the JH is JH6 (eg, JH6\*02);

Wherein

(a) the heavy chain variable domain has been produced *in vivo* in a non-human vertebrate (eg, a mouse or a rat); and/or

- (b) the heavy chain variable domain comprises non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences.

#### 10 **BRIEF DESCRIPTION OF THE FIGURES**

Figures 1 to 3: Schematic illustrating a protocol for producing recombineered BAC vectors to add V gene segments into a mouse genome;

Figure 4: Schematic illustrating a protocol for adding V gene segments to a mouse genome using sequential recombinase mediated cassette exchange (sRMCE);

- 15 Figure 5 (in 4 parts): Alignment of 13 IGHV1-69 alleles showing the variable (V) coding region only. Nucleotides that differ from VH1-69 allele \*01 are indicated at the appropriate position whereas identical nucleotides are marked with a dash. Where nucleotide changes result in amino acid differences, the encoded amino acid is shown above the corresponding triplet. Boxed regions correspond to CDR1, CDR2 and CDR3 as indicated; and

- 20 Figure 6: RSS structure and recombination schematic.

#### **DETAILED DESCRIPTION OF THE INVENTION**

A source for human V, D and J gene segments is Bacterial Artificial Chromosomes (RPCI-11 BACs) obtained from Roswell Park Cancer Institute (RPCI)/Invitrogen. See

- 25 <http://bacpac.chori.org/hmale11.htm>, which describes the BACs as follows:-

#### **"RPCI - 11 Human Male BAC Library**

*The RPCI-11 Human Male BAC Library (Osoegawa et al., 2001) was constructed using improved cloning techniques (Osoegawa et al., 1998) developed by Kazutoyo Osoegawa. The library was*

generated by Kazutoyo Osoegawa. Construction was funded by a grant from the National Human Genome Research Institute (NHGRI, NIH) (#1R01RG01165-03). This library was generated according to the new NHGRI/DOE "Guidance on Human Subjects in Large-Scale DNA Sequencing..."

- 5 "Male blood was obtained via a double-blind selection protocol. Male blood DNA was isolated from one randomly chosen donor (out of 10 male donors)".
- Osoegawa K, Mammoser AG, Wu C, Frengen E, Zeng C, Catanese JJ, de Jong PJ; Genome Res. 2001 Mar;11(3):483-96; "A bacterial artificial chromosome library for sequencing the complete human genome";
  - 10 • Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J., and de Jong, P.J. (1998); "An Improved Approach for Construction of Bacterial Artificial Chromosome Libraries"; Genomics 52, 1-8.

As a source of antibody gene segment sequences, the skilled person will also be aware of the following available databases and resources (including updates thereof):-

15

1.1. The Kabat Database (G. Johnson and T. T.Wu, 2002; <http://www.kabatdatabase.com>). Created by E. A. Kabat and T. T. Wu in 1966, the Kabat database publishes aligned sequences of antibodies, T-cell receptors, major histocompatibility complex (MHC) class I and II molecules, and other proteins of immunological interest. A searchable interface is provided by the SeqhuntII tool, and a range of utilities is available for sequence alignment, sequence subgroup classification, and the generation of variability plots. See also Kabat, E. A., Wu, T. T., Perry, H., Gottesman, K., and Foeller, C. (1991) *Sequences of Proteins of Immunological Interest*, 5th ed., NIH Publication No. 91-3242, Bethesda, MD, which is incorporated herein by reference, in particular with reference to human gene segments for use in the present invention.

20

25

1.2. KabatMan (A. C. R. Martin, 2002; <http://www.bioinf.org.uk/abs/simkab.html>). This is a web interface to make simple queries to the Kabat sequence database.

1.3. *IMGT, the International ImMunoGeneTics Information System®; M.-P. Lefranc, 2002; <http://imgt.cines.fr>*. IMGT is an integrated information system that specializes in antibodies, T cell receptors, and MHC molecules of all vertebrate species. It provides a common portal to standardized data that include nucleotide and protein sequences, oligonucleotide primers, gene maps, genetic polymorphisms, specificities, and two-dimensional (2D) and three-dimensional (3D) structures. IMGT includes three sequence databases (*IMGT/LIGM-DB*, *IMGT/MHC-DB*, *IMGT/PRIMERDB*), one genome database (*IMGT/GENE-DB*), one 3D structure database (*IMGT/3Dstructure-DB*), and a range of web resources (“*IMGT Marie-Paule page*”) and interactive tools.

10 1.4. *V-BASE (I. M. Tomlinson, 2002; <http://www.mrc-cpe.cam.ac.uk/vbase>)*. V-BASE is a comprehensive directory of all human antibody germline variable region sequences compiled from more than one thousand published sequences. It includes a version of the alignment software DNAPLOT (developed by Hans-Helmar Althaus and Werner Müller) that allows the assignment of rearranged antibody V genes to their closest germline gene segments.

15

1.5. *Antibodies—Structure and Sequence (A. C. R. Martin, 2002; <http://www.bioinf.org.uk/abs>)*. This page summarizes useful information on antibody structure and sequence. It provides a query interface to the Kabat antibody sequence data, general information on antibodies, crystal structures, and links to other antibody-related information. It also distributes an automated summary of all antibody structures deposited in the Protein Databank (PDB). Of particular interest is a thorough description and comparison of the various numbering schemes for antibody variable regions.

20

1.6. *AAAAA—AHO’s Amazing Atlas of Antibody Anatomy (A. Honegger, 2001; <http://www.unizh.ch/~antibody>)*. This resource includes tools for structural analysis, modeling, and engineering. It adopts a unifying scheme for comprehensive structural alignment of antibody and T-cell-receptor sequences, and includes Excel macros for antibody analysis and graphical representation.

25

1.7. *WAM—Web Antibody Modeling (N. Whitelegg and A. R. Rees, 2001;*

30 *<http://antibody.bath.ac.uk>*). Hosted by the Centre for Protein Analysis and Design at the University

of Bath, United Kingdom. Based on the AbM package (formerly marketed by Oxford Molecular) to construct 3D models of antibody Fv sequences using a combination of established theoretical methods, this site also includes the latest antibody structural information.

5 1.8. *Mike's Immunoglobulin Structure/Function Page (M. R. Clark, 2001;*  
*<http://www.path.cam.ac.uk/~mrc7/mikeimages.html>)* These pages provide educational materials on  
immunoglobulin structure and function, and are illustrated by many colour images, models, and  
animations. Additional information is available on antibody humanization and Mike Clark's  
Therapeutic Antibody Human Homology Project, which aims to correlate clinical efficacy and anti-  
10 immunoglobulin responses with variable region sequences of therapeutic antibodies.

1.9. *The Antibody Resource Page (The Antibody Resource Page, 2000;*  
*<http://www.antibodyresource.com>)*. This site describes itself as the "complete guide to antibody  
research and suppliers." Links to amino acid sequencing tools, nucleotide antibody sequencing tools,  
15 and hybridoma/cell-culture databases are provided.

1.9. *Humanization by Design (J. Saldanha, 2000; <http://people.cryst.bbk.ac.uk/~ubcq07s>)*. This  
resource provides an overview on antibody humanization technology. The most useful feature is a  
searchable database (by sequence and text) of more than 40 published humanized antibodies  
20 including information on design issues, framework choice, framework back-mutations, and binding  
affinity of the humanized constructs.

See also Antibody Engineering Methods and Protocols, Ed. Benny K C Lo, Methods in Molecular  
Biology™, Human Press. Also at [http://www.blogsua.com/pdf/antibody-engineering-methods-and-  
25 protocolsantibody-engineering-methods-and-protocols.pdf](http://www.blogsua.com/pdf/antibody-engineering-methods-and-protocolsantibody-engineering-methods-and-protocols.pdf)

In one embodiment throughout the present text, "germline" refers to the canonical germline gene  
segment sequence.

The present invention is directed to the provision of novel V, D and J pairings in immunoglobulin heavy and light chain loci. Novel, biased antibody diversities and potentially expanded diversities are provided. One aspect of the invention exploits the natural pairing of compatible recombination signal sequences (RSSs) during antibody V(D)J recombination *in vivo*, and this aspect of the invention provides new, synthetic combinations of V, D and J gene segments using the observation of RSS compatibility.

Another aspect of the invention is based on the observation of V, D and J usage bias in naturally-occurring human antibodies raised against infectious disease pathogens. The invention is useful for manipulating the antibody gene diversity in transgenic non-human animals, thus providing for novel and potentially expanded diversity or diversity that is biased towards variable gene usage common to antibodies useful for treating and/or preventing certain diseases or conditions, such as infectious diseases. This ability to bias the antibody repertoire also provides methods of simplifying the production of antibody mixtures, such as polyclonal antibody therapeutics useful for the treatment and/or prevention of infectious diseases where a polyclonal approach to target multiple pathogen antigens is desirable. To this end, the present invention also provides bispecific antibodies that are capable of binding to more than one antigen (eg, multiple infectious antigens expressed by the same pathogen), thus providing advantages (such as manufacturing, dosing and administration advantages) not possible with polyclonal antibody mixtures.

The present invention provides vertebrates and cells, such as transgenic mice or rats or transgenic mouse or rat cells. Furthermore, the invention relates to methods of using the vertebrates to isolate antibodies or nucleotide sequences encoding antibodies. Antibodies, nucleotide sequences, pharmaceutical compositions and uses are also provided by the invention.

To this end, the present invention provides, in a first configuration

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

- (a) An immunoglobulin heavy chain locus comprising one or more human V gene segments (optionally a plurality of VH), one or more human D gene segments and one or more human J gene segments upstream of a constant region; optionally wherein the heavy chain locus is according to (a) of the second configuration described below; and
- 5 (b) An immunoglobulin light chain locus comprising either
- (i) one or more human VH gene segments and one or more human J gene segments upstream of a constant region (optionally a rearranged  $V_HJ_LC_L$  or  $V_HJ_\lambda C_L$ , wherein the  $C_L$  is  $C_\lambda$  or  $C_k$ ); or
- (ii) one or more human VL gene segments, one or more human D gene segments and one or more human  $J_H$  gene segments upstream of a constant region (optionally a rearranged  $V_LDJ_HC_L$  or  $V_\lambda DJ_HC_L$ , wherein the  $C_L$  is  $C_\lambda$  or  $C_k$ ); or
- 10 (iii) one or more human VL gene segments selected from the group consisting of: a  $V_\lambda$ II gene family member,  $V_\lambda$ VII 4A,  $V_\lambda$ II 2.1,  $V_\lambda$ VII 4A, a  $V_\lambda$ 1 gene family member, a  $V_\lambda$ 3 gene family member, IGLV1S2,  $V_\lambda$ 3-cML70, Ialh2, Ialvl, Ia3h3, Kv325, a  $V_{kl}$  gene family member,  $\kappa$ I-15A (KL012),  $V_\kappa$ II family member, a  $V_\kappa$ III family member, a  $V_{kl}$  gene family member,  $\kappa$ I-15A (KL012),  $V_\kappa$ II A2 (optionally the A2a allele),  $V_\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto, and one or more human  $J_L$  gene segments upstream of a constant region; optionally the one or more VL gene segments are selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below.
- 20
- Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and
- 25 light chains produced by recombination of the light chain locus.

This configuration of the invention, thus, provides for the possibility of novel, synthetic antibody and gene repertoires in a transgenic non-human vertebrate, such as a mouse or rat. Such new repertoires are desirable, since they provide for the possibility of a novel pool of antibodies from which lead antibodies can be selected following immunisation of the vertebrate with a

30 predetermined antigen. This, therefore, provides for a pool from which antibodies with desirable characteristics can be isolated, for example, antibodies with relatively high affinity for specific target antigen binding. It is desirable to isolate high affinity antibodies directly from the immunised vertebrate, since this can provide for an antibody lead that is potentially useful as a therapeutic



and/or prophylactic medicament without the need for further extensive affinity maturation (eg, by *in vitro* antibody display such as ribosome display or phage display). Modification of the effector portions of the antibody can be made as desired (eg, humanisation of the constant region), without the need to manipulate the sequences of the variable regions. Alternatively, or additionally, the pool of antibodies may allow for selection of a lead antibody with desirable biophysical characteristics and/or epitope specificity. The latter may be important for finding lead antibodies against epitopes that have not previously raised therapeutic and/or prophylactic antibodies or epitopes that are difficult to reach by antibodies generated by antibody gene diversities generated by prior non-human vertebrates bearing transgenic immunoglobulin loci, eg, those based on the single human genome represented by the RPCI-11 BACs.

The cells of the invention (according to any aspect or configuration) is, for example, a B-cell, hybridoma or a stem cell, optionally an embryonic stem cell or haematopoietic stem cell. In one aspect the ES cell is derived from the mouse C57BL/6N, C57BL/6J, 129S5 or 129Sv strain. In one aspect the non-human vertebrate is a rodent, suitably a mouse, and cells of the invention, are rodent cells or ES cells, suitably mouse ES cells. The ES cells of the present invention can be used to generate animals using techniques well known in the art, which comprise injection of the ES cell into a blastocyst followed by implantation of chimaeric blastocysts into females to produce offspring which can be bred and selected for homozygous recombinants having the required insertion. In one aspect the invention relates to a transgenic animal comprised of ES cell-derived tissue and host embryo derived tissue. In one aspect the invention relates to genetically-altered subsequent generation animals, which include animals having a homozygous recombinants for the VDJ and/or VJ regions.

Vertebrates bearing one or more light chain loci according to (b)(i) and (ii) provide for novel and potentially expanded antibody and gene repertoires by exploiting synthetic, non-naturally-occurring, combinations of immunoglobulin gene segments (V, D, J, C). In this respect, the present inventors have realised the desirability and possibility of providing for antibody and gene repertoires that mix heavy chain gene segments with those of light chain loci. This is based on observations of the inventors: Firstly, nature suggests the possibility of functional antibodies having VH-VH or VL-VL pairings (as opposed to more classical VH-VL pairings). For example, reference is made to heavy chain antibodies of *Camelidae* which produce antibodies with paired VH domains and is devoid of

light chain VL domains (eg, see Nature. 1993 Jun 3;363(6428):446-8; Naturally occurring antibodies devoid of light chains; Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R). These antibodies function to specifically bind antigen, such antibodies being found in the blood of such *Camelidae* (eg, llamas, camels, alpacas). Such antibodies with VH pairs can also be synthetically produced to provide therapeutic and prophylactic medicaments (eg, see WO1994004678, WO2004041862, WO2004041863). Transgenic mice also can produce such heavy chain antibodies and the *in vivo* production of the antibodies allows the mouse's immune system to select for VH-VH pairings, sometimes selecting for such pairings in which mutations have been introduced *in vivo* by the mouse to accommodate the pairing (WO2010109165A2). Thus, the inventors realised that the adoption of an *in vivo* antibody production system (rather than an *in vitro* system such as phage or ribosome display of antibodies) is desirable to accommodate the synthetic immunoglobulin gene segment combinations that are now contemplated by the present invention.

A second observation of the present inventors lies in the architecture of naturally-occurring immunoglobulin loci, and in particular the arrangement of recombination signal sequences (RSSs) that mediate V(D)J recombination *in vivo* (see, eg, Cell. 2002 Apr;109 Suppl:S45-55. The mechanism and regulation of chromosomal V(D)J recombination; Bassing CH, Swat W, Alt FW, the disclosure of which is incorporated herein by reference). As illustrated in Fig. 6, two types of RSS element have been identified: a one-turn RSS (12-RSS) and a two-turn RSS (23-RSS). In natural VJ recombination in the lambda light chain locus, recombination if effected between a two-turn RSS that lies 3' of a V lambda and a one-turn RSS that lies 5' of a J lambda, the RSSs being in opposite orientation. In natural VJ recombination in the kappa light chain locus, recombination if effected between a one-turn RSS that lies 3' of a V kappa and a two-turn RSS that lies 5' of a J kappa, the RSSs being in opposite orientation. In natural VD recombination in the heavy chain locus, recombination if effected between a two-turn RSS that lies 3' of a VH and a one-turn RSS that lies 5' of a D, the RSSs being in opposite orientation. In natural DJ recombination in the heavy chain locus, recombination if effected between a one-turn RSS that lies 3' of a D and a two-turn RSS that lies 5' of a JH, the RSSs being in opposite orientation. Thus, generally a two-turn RSS is compatible with a one-turn RSS in the opposite orientation. The inventors realised that they could use this observation in constructing transgenic immunoglobulin loci such that a 5' gene segment can recombine with a 3' gene segment (eg, a V with a J; or a V with a D) when there is provided a two-turn RSS and a one-turn RSS in the opposite orientation, with each RSS adjacent a respective one of the gene segments. Thus, the

inventors realised in one embodiment that an immunoglobulin locus can be constructed with one or more of the following arrangements (5' to 3'):-

- 5 (k) [V (heavy, lambda or kappa)] – [two-turn RSS] --- [one-turn RSS] – [D] – [JH], wherein said RSSs are in an opposite orientation;
- (l) [VH] – [D] – [two-turn RSS] --- [one-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (m) [VH] – [D] – [one-turn RSS] --- [two-turn RSS] – [J kappa], wherein said RSSs are in an opposite orientation;
- 10 (n) [VH or V kappa] - [two-turn RSS] --- [one-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (o) [V kappa] - [one-turn RSS] --- [two-turn RSS] – [JH or J lambda], wherein said RSSs are in an opposite orientation;
- 15 (p) [V (heavy, lambda or kappa)] – [one-turn RSS] --- [two-turn RSS] – [D] – [JH], wherein said RSSs are in an opposite orientation;
- (q) [VH] – [D] – [one-turn RSS] --- [two-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (r) [VH] – [D] – [two-turn RSS] --- [one-turn RSS] – [J kappa], wherein said RSSs are in an opposite orientation;
- 20 (s) [VH or V kappa] - [one-turn RSS] --- [two-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (t) [V kappa] - [two-turn RSS] --- [one-turn RSS] – [JH or J lambda], wherein said RSSs are in an opposite orientation.

25

The skilled person will realise that standard molecular biology techniques can be used to provide vectors comprising synthetic combinations of RSS with V, D or J for use in this aspect of the invention, such that the vectors can be used to build a transgenic immunoglobulin locus (eg, using homologous recombination and/or recombinase mediated cassette exchange as known in the art, eg, see US7501552 (Medarex), US5939598 (Abgenix), US6130364 (Abgenix), WO02/066630 (Regeneron), WO2011004192 (Genome Research Limited), WO2009076464, WO2009143472 and WO2010039900 (Ablexis), the disclosures of which are explicitly incorporated herein. For example, such synthetic combinations with RSS and gene segments can be made using standard recombineering techniques in E coli to construct BAC vectors harbouring the synthetic combination

30

prior to insertion in embryonic stem cells using homologous recombination or RMCE (eg, using cre/lox site-specific recombination). Details of recombineering can be found at [www.genebridges.com](http://www.genebridges.com) and in EP1034260 and EP1204740 the disclosures of which are explicitly incorporated herein.

5

In one embodiment of (b)(i), all of the light chain locus V gene segments are human VH gene segments (optionally with one or more human V lambda gene segments).

10

In one embodiment of (b)(i), the constant region is a mouse, rat or human CL, eg, Cλ. In one embodiment, the J and constant regions are provided by one or more human JλCλ.

15

20

Although having utility generally to any antigen and disease setting, vertebrates bearing one or more light chain loci according to (b)(iii) are useful, in particular, for generating antibody leads against infectious disease pathogens. In this respect, the present inventors have realised the desirability and possibility of providing for antibody and gene repertoires that are biased to immunoglobulin gene segments commonly found in natural antibody reactions of humans to infectious disease pathogens. The inventors realised that it would be desirable to provide for vertebrates, cells, methods etc for the production of therapeutic and/or prophylactic antibodies based on natural human immune responses to antigens, such as antigens of infectious disease pathogens. In this respect, the literature observes frequently used immunoglobulin gene segments to raise anti-infective responses in humans (Table 1).



**Table 1: Immunoglobulin Gene Usage in Human Antibody Responses to Infectious Disease Pathogens**

GENE	ANTIGEN	ORGANISM	REFERENCES [Human Ab Source]
<b>BACTERIAL PATHOGENS</b>			
<p><u>KAPPA V GENES</u></p> <ul style="list-style-type: none"> <li>• Vk II germ line gene A2 + JK3</li> <li>• Vk II family gene + JK4</li> <li>• 94% identical to the A27 (Humkv325) germ line gene</li> <li>• a Vkl gene family member; κI-15A (KL012)</li> </ul> <p><u>LAMBDA V GENES</u></p> <ul style="list-style-type: none"> <li>• Four Vλ VII family members that are 96-98% identical to each other</li> <li>• Vλ II family members (82, 89 and 91% homologous to Vλ2.1 gene)</li> </ul>	<p>Haemophilus influenzae type b polysaccharide (Hib PS)</p>	<p>Haemophilus influenzae</p>	<ol style="list-style-type: none"> <li>1. Lonberg, Nat Biotech 2005; [human PBMCs]</li> <li>2. Adderson et al, J Clin Invest 1992; [Human PBLs]</li> <li>3. Chung et al, J Immunol 1993</li> <li>4. Nadel et al, J Immunol 1998</li> <li>5. Feeney et al, J Clin Invest 1996</li> <li>6. Lucas et al, Infect Immun 1994; [Human PBLs]</li> <li>7. Adderson et al, J Clin Invest 1993; [Human PBLs]</li> <li>8. Granoff et al, J Clin Invest 1993; [human PBLs]</li> <li>9. Azmi et al, Infect Immun 1994; [human tonsil cells]</li> </ol>

<p>+ VHIII segments closely homologous to germline gene 9.1  <math>V_{\lambda}</math>VII 4A</p> <ul style="list-style-type: none"> <li>All with J<math>\lambda</math> homologous to germline J<math>\lambda</math>2 and J<math>\lambda</math>3</li> </ul> <p><u>VH GENES</u></p> <ul style="list-style-type: none"> <li>VH 96% identical to the VH germ line gene segment DP77 (V3-21)</li> <li>LSG6.1, LSG12.1, V<sub>H</sub>III VH26, V<sub>H</sub>III 9.1</li> </ul> <p><u>VH and VL COMBINATIONS</u></p> <ul style="list-style-type: none"> <li>V<sub>H</sub>III 9.1+ V<sub><math>\lambda</math></sub>VII 4A</li> <li>V<sub>H</sub>III 9.1 + V<sub><math>\lambda</math></sub>II 2.1</li> <li>V<sub>H</sub>III 9.1 + V<sub><math>\kappa</math></sub>II A2</li> <li>V<sub>H</sub>III VH26 + V<sub><math>\lambda</math></sub>II 2.1</li> </ul> <ul style="list-style-type: none"> <li>V<sub>H</sub>III 9.1; V<sub>H</sub>III H11; V<sub>H</sub>III VH26</li> </ul>	<ul style="list-style-type: none"> <li>Polysaccharide capsule of E coli</li> </ul>	<ul style="list-style-type: none"> <li>E coli K1</li> </ul>	<p>9. Azmi et al, Infect Immun 1994</p>
---	--	---	---

<p>kl15A Vλ2.1</p>	<p>K1</p> <ul style="list-style-type: none"> <li>• Meningococcal B polysaccharide; Poly[α(2→8)-N-acetylneuramic acid</li> </ul>	<ul style="list-style-type: none"> <li>• Neisseria meningitidis Group B</li> </ul>	
<p><b>VIRAL PATHOGENS</b></p>			
<ul style="list-style-type: none"> <li>• VHIII or VHIV family member</li> <li>• VλI or Vλ3 member</li> </ul>	<ul style="list-style-type: none"> <li>• HSV 120-kD glycoprotein</li> </ul>	<ul style="list-style-type: none"> <li>• Herpes family virus</li> <li>• Herpes simplex virus (HSV); HSV-1; HSV-2</li> </ul>	<p>10. Huang et al, J Clin Invest 1992; [human tonsils]</p>
<ul style="list-style-type: none"> <li>• VH26 + Dk1 + JH6 with IGLV1S2 + Jλ2</li> </ul>	<ul style="list-style-type: none"> <li>• 116-, 105-, 64-kD glycoproteins of VZV</li> </ul>	<ul style="list-style-type: none"> <li>• Varicella zoster virus (VZV)</li> </ul>	
<ul style="list-style-type: none"> <li>• VH4.18</li> </ul>			
<ul style="list-style-type: none"> <li>• VH2-1 (VH3) + D region Dxp'1 + JH5 with Vλ3 cML70 + Jλ3</li> </ul>			
<ul style="list-style-type: none"> <li>• VH1GRR+ JH3 + Dn4r or D2r with IGLV1S2 + Jλ2</li> </ul>			



<p>For VZV Abs: ha3h2 (VH3) with la1h2 (VA); or ha1c1 (VH1) with la1l (VA1)</p> <p>For VZV Abs: ha4h3 (VH4) with la3h3 (VA3)</p>			
<ul style="list-style-type: none"> <li>• Hv1051 (VH)</li> <li>• Kv325 (Vk)</li> </ul>		<ul style="list-style-type: none"> <li>• Cytomegalovirus (CMV)</li> </ul>	<p>10. Huang et al, J Clin Invest 1992;</p>
<ul style="list-style-type: none"> <li>• 71-2 (VH)</li> <li>• Hv1f10 (VH)</li> <li>• VH4.11</li> <li>• 71-4 (VH)</li> <li>• VH251</li> <li>• VH1-69</li> </ul>		<ul style="list-style-type: none"> <li>• HIV</li> </ul>	<p>10. Huang et al, J Clin Invest 1992;</p> <p>11. Wang &amp; Palese, Science 2011</p>
<ul style="list-style-type: none"> <li>• VH1-69</li> </ul>	<p>Haemagglutinin (HA)</p>	<ul style="list-style-type: none"> <li>• Influenza virus, eg, Group 1 and/or Group 2 Influenza A virus; eg, H1N1, H2N2, or H3N2 or H7N2 or H7N7 influenza virus</li> </ul>	<p>12. Ekiert et al, Science 2009</p> <p>13. Throsby et al, PLoS One 2008</p> <p>14. Sui et al, Nat Struct Mol Biol 2009</p> <p>15. Ekiert et al, Science 2011</p>



References:

1. Nat Biotechnol. 2005 Sep;23(9):1117-25; Human antibodies from transgenic animals; Lonberg N.
2. J Clin Invest. 1992 Mar;89(3):729-38; Immunoglobulin light chain variable region gene sequences for human antibodies to Haemophilus influenzae type b capsular polysaccharide are dominated by a limited number of V kappa and V lambda segments and VJ combinations; Adderson EE, Shackelford PG, Insel RA, Quinn A, Wilson PM, Carroll WL.
3. J Immunol. 1993 Oct 15;151(8):4352-61; Clonal characterization of the human IgG antibody repertoire to Haemophilus influenzae type b polysaccharide. V. In vivo expression of individual antibody clones is dependent on Ig CH haplotypes and the categories of antigen; Chung GH, Scott MG, Kim KH, Kearney J, Siber GR, Ambrosino DM, Nahm MH.
4. J Immunol. 1998 Dec 1;161(11):6068-73; Decreased frequency of rearrangement due to the synergistic effect of nucleotide changes in the heptamer and nonamer of the recombination signal sequence of the V kappa gene A2b, which is associated with increased susceptibility of Navajos to Haemophilus influenzae type b disease; Nadel B, Tang A, Lugo G, Love V, Escuro G, Feeney AJ.
5. J Clin Invest. 1996 May 15;97(10):2277-82; A defective Vkappa A2 allele in Navajos which may play a role in increased susceptibility to haemophilus influenzae type b disease; Feeney AJ, Atkinson MJ, Cowan MJ, Escuro G, Lugo G.
6. Infect Immun. 1994 Sep;62(9):3873-80; Variable region sequences of a protective human monoclonal antibody specific for the Haemophilus influenzae type b capsular polysaccharide; Lucas AH, Larrick JW, Reason DC.
7. J Clin Invest. 1993 Jun;91(6):2734-43; Restricted immunoglobulin VH usage and VDJ combinations in the human response to Haemophilus influenzae type b capsular polysaccharide. Nucleotide sequences of monospecific anti-Haemophilus antibodies and polyspecific antibodies cross-reacting with self antigens; Adderson EE, Shackelford PG, Quinn A, Wilson PM, Cunningham MW, Insel RA, Carroll WL.

8. J Clin Invest. 1993 Mar;91(3):788-96; Variable region expression in the antibody responses of infants vaccinated with Haemophilus influenzae type b polysaccharide-protein conjugates. Description of a new lambda light chain-associated idiootype and the relation between idiootype expression, avidity, and vaccine formulation. The Collaborative Vaccine Study Group; Granoff DM, Shackelford PG, Holmes SJ, Lucas AH.
9. Infect Immun. 1994 May;62(5):1776-86; Variable region sequences and idiotypic expression of a protective human immunoglobulin M antibody to capsular polysaccharides of Neisseria meningitidis group B and Escherichia coli K1; Azmi FH, Lucas AH, Raff HV, Granoff DM.
10. J Clin Invest. 1992 Dec;90(6):2197-208; Sequence analyses of three immunoglobulin G anti-virus antibodies reveal their utilization of autoantibody-related immunoglobulin Vh genes, but not V lambda genes; Huang DF, Olee T, Masuho Y, Matsumoto Y, Carson DA, Chen PP.
11. Science. 2011 Aug 12;333(6044):834-5, Biochemistry. Catching a moving target, Wang TT, Palese P
12. Science. 2009 Apr 10;324(5924):246-51. Epub 2009 Feb 26; Antibody recognition of a highly conserved influenza virus epitope; Ekiert DC, Bhabha G, Elsliger MA, Friesen RH, Jongeneelen M, Throsby M, Goudsmit J, Wilson IA.
13. PLoS One. 2008;3(12):e3942. Epub 2008 Dec 16; Heterosubtypic neutralizing monoclonal antibodies cross-protective against H5N1 and H1N1 recovered from human IgM+ memory B cells; Throsby M, van den Brink E, Jongeneelen M, Poon LL, Alard P, Cornelissen L, Bakker A, Cox F, van Deventer E, Guan Y, Cinatl J, ter Meulen J, Lasters I, Carsetti R, Peiris M, de Kruif J, Goudsmit J.
14. Nat Struct Mol Biol. 2009 Mar;16(3):265-73. Epub 2009 Feb 22; Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses, Sui J, Hwang WC, Perez S, Wei G, Aird D, Chen LM, Santelli E, Stec B, Cadwell G, Ali M, Wan H, Murakami A, Yammanuru A, Han T, Cox NJ, Bankston LA, Donis RO, Liddington RC, Marasco WA.

15. Science. 2011 Aug 12;333(6044):843-50. Epub 2011 Jul 7, A highly conserved neutralizing epitope on group 2 influenza A viruses, Ekiert DC, Friesen RH, Bhabha G, Kwaks T, Jongeneelen M, Yu W, Ophorst C, Cox F, Korse HJ, Brandenburg B, Vogels R, Brakenhoff JP, Kompier R, Koldijk MH, Cornelissen LA, Poon LL, Peiris M, Koudstaal W, Wilson IA, Goudsmit J.

\*\*\*

In one embodiment, in (b)(i) the J gene segments of the light chain locus are  $J_{\lambda}$  gene segments and optionally the constant region of the light chain locus is a lambda constant region; or in (b)(ii) the VL is a  $V_{\lambda}$  and optionally the constant region of the light chain locus is a lambda constant region. Alternatively, the constant region is C kappa.

In one embodiment, in (b)(i) the V gene segment repertoire of the light chain locus comprises or consists of one or more VH gene segments selected from the group consisting of: a  $V_{HIII}$  gene family member (optionally, a  $V_{HIIIa}$  or  $V_{HIIIb}$  family member), a  $V_{HIV}$  gene family member,  $V_{HIII}$  9.1 (VH3-15),  $V_{HIII}$  VH26 (VH3-23),  $V_{H3}$ -21, LSG6.1, LSG12.1, DP77 (V3-21),  $V_H$  H11, VH1GRR, ha3h2,  $V_{HI}$ -ha1c1,  $V_{HIII}$ -VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto. These gene segments are useful because they expand the repertoire *in vivo* to VH gene segments that are found in natural human immune responses to antigens, such as antigens of infectious disease pathogens. This is useful, for example, when the vertebrate is immunised with an antigen of an infectious disease pathogen, for generation and isolation of an antibody for treating and/or preventing a disease or condition mediated by said pathogen. In one example, in (b)(i) the V gene segment repertoire of the light chain locus comprises or consists of only VH gene segment selected from the group consisting of: a  $V_{HIII}$  gene family member (optionally, a  $V_{HIIIa}$  or  $V_{HIIIb}$  family member), a  $V_{HIV}$  gene family member,  $V_{HIII}$  9.1 (VH3-15),  $V_{HIII}$  VH26 (VH3-23),  $V_{H3}$ -21, LSG6.1, LSG12.1, DP77 (V3-21),  $V_H$  H11, VH1GRR, ha3h2,  $V_{HI}$ -ha1c1,  $V_{HIII}$ -VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto. This is useful to bias the immune response of the vertebrate (and thus resultant lead antibodies) to a predetermined gene segment, eg, one known to be commonly used in natural human immune responses to antigens, such as antigens of infectious disease pathogens. For example, VH1-69 is commonly used to produce antibodies in humans against Influenza virus (see Table 1); it is possible, therefore, to confine the single VH segment to VH1-69 in

embodiment (b)(i) of the invention.

In one embodiment, in (b)(iii) the light chain locus V gene segment repertoire consists of only one (optionally only two, three or four) VL gene segment type (optionally and one or mutants thereof), wherein the VL gene segment is selected from said group of VL gene segments. This is useful to bias the immune response of the vertebrate (and thus resultant lead antibodies) to a predetermined gene segment, eg, one known to be commonly used in natural human immune responses to antigens, such as antigens of infectious disease pathogens.

In one embodiment, in (a) said constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region and/or in (b) said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

In one embodiment in any configuration of the invention, the genome has been modified to prevent or reduce the expression of fully-endogenous antibody. Examples of suitable techniques for doing this can be found in PCT/GB2010/051122, US7501552, US6673986, US6130364, WO2009/076464, EP1399559 and US6586251, the disclosures of which are incorporated herein by reference. In one embodiment, the non-human vertebrate VDJ region of the endogenous heavy chain immunoglobulin locus, and optionally VJ region of the endogenous light chain immunoglobulin loci (lambda and/or kappa loci), have been inactivated. For example, all or part of the non-human vertebrate VDJ region is inactivated by inversion in the endogenous heavy chain immunoglobulin locus of the mammal, optionally with the inverted region being moved upstream or downstream of the endogenous Ig locus (see, eg, WO2011004192, the disclosure of which is incorporated herein by reference). For example, all or part of the non-human vertebrate VJ region is inactivated by inversion in the endogenous kappa chain immunoglobulin locus of the mammal, optionally with the inverted region being moved upstream or downstream of the endogenous Ig locus. For example, all or part of the non-human vertebrate VJ region is inactivated by inversion in the endogenous lambda chain immunoglobulin locus of the mammal, optionally with the inverted region being moved upstream or downstream of the endogenous Ig locus. In one embodiment the endogenous heavy chain locus is inactivated in this way as is one or both of the endogenous kappa and lambda loci.

Additionally or alternatively, the vertebrate has been generated in a genetic background which prevents the production of mature host B and T lymphocytes, optionally a RAG-1-deficient and/or RAG-2 deficient background. See US5859301 for techniques of generating RAG-1 deficient animals.

Thus, in one embodiment of any configuration or aspect of the invention herein, endogenous heavy and light chain expression has been inactivated.

In a second configuration of the invention, there is provided

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

- (a) An immunoglobulin heavy chain locus comprising one or more human V gene segments (eg, a plurality of VH), one or more human D gene segments and one or more human J gene segments upstream of a constant region; and
- (b) (i) An unrearranged immunoglobulin light chain locus comprising one or more human VH gene segments and one or more human J gene segments upstream of a constant region, wherein each human VH gene segment is a human gene segment identical to (or mutant of, eg, having up to 15 or 10 nucleotide changes from the human gene segment) a human VH gene segment (eg, a germline VH gene segment; eg, a gene segment selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below.) used to produce a rearranged VDJ encoding a heavy chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism); or  
or  
(ii) An immunoglobulin light chain locus comprising a rearranged VJ region or VDJ region upstream of a constant region, wherein the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VH gene segment that is identical to (or mutant of; eg, having up to 15 or 10 nucleotide changes from the human

- gene segment)) the human VH gene segment (eg, germline VH gene segment; eg, a gene segment selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below.) used to produce a rearranged VDJ encoding a heavy chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
- (c) Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments or VJ or VDJ in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains derived from the light chain locus;
- (d) Optionally when (b)(i) applies, each said VH gene segment in the light chain locus is selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto; optionally each VH gene segment is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below.
- (e) Optionally when (b)(ii) applies, the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto; optionally each VH gene segment is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below.



In one embodiment, the antigen is an antigen expressed by a bacterial or viral infectious disease pathogen, eg, any of the pathogens listed in Table 1. For example, the antigen is an antigen selected from the antigens listed in Table 1.

In one embodiment of any aspect, configuration or embodiment of the invention herein, the “human individual harbouring said organism” is a patient that has natural resistance to the pathogen and produces antibodies that bind to the pathogen or an antigen expressed thereby.

In one embodiment of the second configuration, the J gene segments of the light chain locus are J<sub>λ</sub> gene segments and optionally the constant region of the light chain locus is a lambda constant region. Alternatively, the constant region is C kappa.

In one embodiment of the second configuration, the V gene segment repertoire of the light chain locus comprises or consists of one or more VH gene segments selected from the group consisting of: a V<sub>H</sub>III gene family member (optionally, a VHIIIa or VHIIIb family member), a V<sub>H</sub>IV gene family member, V<sub>H</sub>III 9.1 (VH3-15), V<sub>H</sub>III VH26 (VH3-23), V<sub>H</sub>3-21, LSG6.1, LSG12.1, DP77 (V3-21), V<sub>H</sub> H11, VH1GRR, ha3h2, V<sub>H</sub>I-ha1c1, V<sub>H</sub>III-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto. These gene segments are useful because they expand the repertoire *in vivo* to VH gene segments that are found in natural human immune responses to antigens, such as antigens of infectious disease pathogens. This is useful, for example, when the vertebrate is immunised with an antigen of an infectious disease pathogen, for generation and isolation of an antibody for treating and/or preventing a disease or condition mediated by said pathogen. In one example, in (b)(i) the V gene segment repertoire of the light chain locus comprises or consists of only VH gene segment selected from the group consisting of: a V<sub>H</sub>III gene family member (optionally, a VHIIIa or VHIIIb family member), a V<sub>H</sub>IV gene family member, V<sub>H</sub>III 9.1 (VH3-15), V<sub>H</sub>III VH26 (VH3-23), V<sub>H</sub>3-21, LSG6.1, LSG12.1, DP77 (V3-21), V<sub>H</sub> H11, VH1GRR, ha3h2, V<sub>H</sub>I-ha1c1, V<sub>H</sub>III-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto. This is useful to bias the immune response of the vertebrate (and thus resultant lead antibodies) to a predetermined gene segment, eg, one known to be commonly used in natural human immune responses to antigens, such as antigens of infectious disease pathogens. For example, VH1-69 is commonly used to produce antibodies in humans against

Infuenza virus (see Table 1); it is possible, therefore, to confine the single VH segment to VH1-69 in embodiment (b)(i) of the invention.

In one embodiment of the second configuration, in (a) said constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

In one embodiment of the second configuration, in (b) said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

In one embodiment of the second configuration, the genome of said vertebrate or cell is homozygous for light chain locus (b)(i) or (ii); optionally wherein:

- the V gene segment repertoire of the light chain loci consists of one or more human VH gene segments selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto; or

- the recombined VJ or VDJ repertoire of the light chain loci consists of sequences identical to one or more nucleotide sequences produced by the recombination of a human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto, with a human J gene segment and optionally a human D gene segment. In one embodiment, all of the light chain locus V gene segments are from this group.

In one embodiment of the second configuration, endogenous heavy and light chain expression has been inactivated, and wherein light chain loci according to the second configuration are the only

functional light chain loci in the genome of the vertebrate or cell.

In one embodiment of the second configuration, each immunoglobulin light chain locus of said vertebrate or cell is according to (b)(i) and comprises only a single human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain so that all light chain loci comprise the same, single human VH gene segment. In this embodiment (and generally in other embodiments, configurations and aspects of the invention), confinement of heavy and/or light chain locus architecture is useful for biasing or controlling the antibody and gene repertoire, eg, to mirror human immune responses as mentioned above. Provision of a single light or heavy chain variable (and optionally D and/or J) gene segment (or only this with closely related mutants thereof) – or confinement in embodiments below to a single rearranged V(D)J region or single heavy or light chain – is advantageous for simplifying the expression and production of therapeutic/prophylactic antibodies since this restricts the number of antibody species produced during downstream manufacture. A common heavy or light chain is advantageous to enable co-expression of a plurality (eg, two, three or more) different antibodies in the same expression medium, for example from the same host cell. See, eg, EP1523496 (Merus BV) and WO2011097603 (Regeneron Pharmaceuticals, Inc).

In one embodiment of the second configuration, each immunoglobulin light chain locus of said vertebrate or cell is according to (b)(ii) and comprises only a single rearranged VJ or VDJ region, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain so that all light chain loci comprise the same, single rearranged VJ or VDJ region.

In one embodiment of the second configuration, each immunoglobulin light chain locus further comprises a VH gene segment or rearranged region that is a mutant (eg, having up to 15 or 10 nucleotide changes from the VH gene segment) respectively of said selected human VH gene segment or rearranged region, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain mutant VH gene segment or rearranged region.

In one embodiment of the second configuration, each immunoglobulin light chain locus comprises only two or three human VH gene segments selected from said group, optionally wherein the genome of the vertebrate or cell is homozygous for said two or three light chain human VH gene segments.

In one embodiment of the second configuration, each immunoglobulin light chain locus comprises only two or three of said rearranged VJ or VDJ regions, optionally wherein the genome of the vertebrate or cell is homozygous for said two or three light chain rearranged VJ or VDJ regions.

The invention provides a monoclonal or polyclonal antibody composition prepared by immunisation of at least one vertebrate (eg, mouse or rat) according to any configuration, aspect or embodiment of the invention, optionally wherein the antigen is an antigen of an infectious disease pathogen (eg, a bacterial or viral pathogen antigen or an antigen listed in Table 1), optionally wherein the same antigen is used to immunise all the vertebrates; optionally wherein the antibody or antibodies are IgG-type (eg, IgG1).

The invention provides a first method of isolating an antibody that binds a predetermined antigen (eg, a bacterial or viral pathogen antigen or an antigen listed in Table 1), the method comprising

- (a) providing a vertebrate (optionally a mouse or rat) according to any configuration, aspect or embodiment of the invention;
- (b) immunising (eg, using a standard prime-boost method) said vertebrate with said antigen (optionally wherein the antigen is an antigen of an infectious disease pathogen);
- (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes expressing antibodies that bind to the antigen;
- (d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by producing hybridomas therefrom; and
- (e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes.

In a first embodiment of the first method of the invention, the method comprises the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally exchanging the heavy chain constant region nucleotide sequence of the antibody with a

nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host. The skilled person will be aware of standard molecular biology techniques to do this. For example, see Harlow, E. & Lane, D. 1998, 5<sup>th</sup> edition, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Lab. Press, Plainview, NY; and Pasqualini and Arap, *Proceedings of the National Academy of Sciences* (2004) 101:257-259 for standard immunisation. Joining of the variable regions of an antibody to a human constant region can be effected by techniques readily available in the art, such as using conventional recombinant DNA and RNA technology as will be apparent to the skilled person. See e.g. Sambrook, J and Russell, D. (2001, 3<sup>d</sup> edition) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).

In one embodiment of the first method of the invention, the method comprises further making a mutant or derivative of the antibody.

A method of producing a polyclonal antibody mixture is provided, the method comprising carrying out the first method of the invention by separately immunising first and second vertebrates (optionally first and second mice or first and second rats) with antigen and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):

- (i) the vertebrates are immunised with the same antigen or different antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or by family members or different strains of the organism));
- (ii) prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene) and optionally the identical J repertoire; optionally the light chain loci of the mammals are identical prior to immunisation;
- (iii) prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.

A method of producing a polyclonal antibody mixture is provided, the method comprising carrying out the first method of the invention by immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):

- (i) the antigens are expressed by the same pathogenic organism (or by family members or different strains of the organism));
- (ii) prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene) and optionally the identical J repertoire; optionally the light chain loci of the mammals are identical prior to immunisation;
- (iii) prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.

The invention provides a second method:

A method of producing host cells (eg, Chinese Hamster Ovary (CHO) or HEK293 cells) capable of expressing a polyclonal antibody mixture is provided, the method comprising, in a method according to said first embodiment of the first method of the invention:-

- (a) immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));
- (b) isolating nucleic acid encoding first and second anti-antigen antibodies from B lymphocytes from said vertebrates;
- (c) determining the nucleotide sequences of the heavy and light chain variable regions (optionally the entire heavy and/or light chain sequences) of the first antibody;
- (d) determining the nucleotide sequence of the heavy variable region and optionally the light chain variable region of the second antibody;

- (e) inserting the heavy chain variable region coding sequence of each antibody into a heavy chain expression vector; optionally wherein the constant region coding sequence of each heavy chain is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (f) inserting the light chain variable region coding sequence of the first antibody into a light chain expression vector; optionally wherein the constant region coding sequence of the light chain of the first antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (g) optionally inserting the light chain variable region coding sequence of the second antibody into a light chain expression vector; optionally wherein the constant region coding sequence of the light chain of the second antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region; and
- (h) introducing each expression vector into a host cell and co-expressing antibody chains in a mixture of said host cells to produce antibodies, each antibody comprising one or both of said heavy chain variable regions and a light chain; optionally wherein the expression vectors are introduced together into the same host cell (eg, a CHO or HEK293 cell) so that the cell is capable of expressing antibody light chains and heavy chains, such that the cell or a plurality of the host cells express antibodies (eg, two, three, four or more different antibodies), each comprising one or both of said heavy chain variable regions and a light chain;
- (i) optionally:
  - prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene segment) and optionally the identical J repertoire (optionally a single J gene segment); optionally the light chain loci of the vertebrates are identical prior to immunisation; or
  - prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.
- (j) optionally:
  - producing a monoclonal or polyclonal antibody mixture, by expressing a monoclonal antibody or polyclonal mixture of said antibodies; optionally followed by isolating an antibody comprising the heavy chain variable region of the first and/or second antibodies.

The invention also provides a monoclonal or polyclonal antibody mixture so produced or a derivative antibody or mixture thereof, eg, where one or more constant region has been changed (eg, replaced

with a different constant region such as a human constant region; or mutated to enhance or ablate Fc effector function). (optionally the entire heavy and/or light chain sequences)

In any of the methods of the invention, optionally each vertebrate used for immunisation is provided by

- (a) isolating from a human blood or tissue (eg, B lymphocytes (PBLs), peripheral blood mononuclear cells (PBMCs), bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
- (b) determining which human germline VH gene segment was recombined in the human to produce the nucleotide sequence of said B lymphocyte that encodes the heavy chain variable region of the antibody;
- (c) constructing a transgenic vertebrate wherein said human germline VH gene segment is provided in a light chain locus thereof according the first or second configuration of the invention; and
- (d) providing said transgenic vertebrate for immunisation in the first method of the invention.

The term "Human blood" herein includes a human blood product minus one or more non-B lymphocyte cellular populations, provided that the product retains antibody-producing cells, eg, PBLs.

In an embodiment of the first method of the invention, each vertebrate used for immunisation is provided by

- (a) isolating from a human blood or tissue (eg, B lymphocytes, PBMCs, bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
- (b) determining a nucleotide sequence of said B lymphocyte that encodes a rearranged VDJ or



VJ region of the antibody;

- (c) constructing a transgenic vertebrate wherein said rearranged VDJ or VJ region is provided in a light chain locus thereof according to the first or second configuration of the invention; and
- (d) providing said transgenic vertebrate for immunisation in the first method of the invention.

Common Light Chain Antibodies & Bispecifics (eg, to two pathogen antigens for infectious diseases)

The invention provides an isolated antibody (eg, IgG-type, such as IgG1-type, antibody) obtainable or obtained by the second method of the invention (including step (j)), or a mutant or derivative antibody thereof wherein (i) the isolated antibody comprises two copies of the heavy chain variable region of said first antibody paired with two copies of the light chain variable region of said first antibody; or (ii) the isolated antibody comprises two copies of the heavy chain variable region of said second antibody paired with two copies of the light chain variable region of said first antibody; or (iii) the isolated antibody is a bispecific antibody comprising one copy of the heavy chain variable region of said first antibody paired with a copy of the light chain variable region of the first antibody, and one copy of the heavy chain variable region of said the antibody paired with a copy of the light chain variable region of the first antibody, optionally wherein the bispecific antibody binds to said first and second antigens recited in claim 24; optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease.

In an aspect of the invention, there is provided a monoclonal or polyclonal antibody mixture (eg, IgG-type antibody or antibodies), wherein the monoclonal antibody or mixture is according to any configuration, aspect, embodiment or example of the invention, or a mutant or derivative antibody thereof optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease, wherein optionally wherein each antibody binds an antigen of an infectious disease pathogen, preferably the same antigen.

In an aspect of the invention, there is provided the use of an isolated, monoclonal or polyclonal antibody according to any configuration, aspect, embodiment or example of the invention, or a mutant or derivative antibody thereof in the manufacture of a medicament for the treatment and/or prevention of an infectious disease, optionally wherein the infectious disease is a disease caused by a bacterial or viral pathogen.

An example of a mutant antibody is one that bears up to 15 or 10 amino acid mutations in its variable regions relative to an isolated antibody (eg, IgG-type, such as IgG1-type, antibody) obtainable or obtained by the second method of the invention (including step (j)). An example of a derivative is one that has been modified to replace a constant region with a different constant region such as a human constant region; or mutated to enhance or ablate Fc effector function.

Examples of infectious diseases are diseases caused or mediated by a bacterial or viral pathogen, eg, a pathogen listed in Table 1. Examples of antigens are those listed in Table 1.

For example, the infectious disease is selected from the group consisting of a disease caused by a pathogen selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus.

The invention further provides a nucleotide sequence encoding an antibody according to any configuration, aspect, embodiment or example of the invention, optionally wherein the nucleotide sequence is part of a vector.

The invention further provides a pharmaceutical composition comprising the antibody or antibodies of any configuration, aspect, embodiment or example of the invention and a diluent, excipient or carrier.

In a third configuration of the invention, there is provided

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

- (a) An immunoglobulin heavy chain locus comprising either:-
  - (i) one or more human VL gene segments, one or more human D gene segments and one or more human J gene segments upstream of a constant region (optionally a rearranged  $V_L D J_H C_H$  or  $V_\lambda D J_H C_H$ ); or
  - (ii) one or more human VH gene segments selected from the group consisting of: a VHIII

gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical; one or more human D gene segments and one or more human J<sub>H</sub> gene segments upstream of a constant region; optionally each VH gene segment (and optionally each D) is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below; and

- (b) An immunoglobulin light chain locus comprising one or more human V gene segments (eg, a plurality of VL) and one or more human J gene segments upstream of a constant region, optionally wherein the light chain locus is according to (b)(i) or (b)(ii) of the first configuration of the invention;

Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains produced by recombination of the light chain locus.

In one example, in (a)(i) all of the heavy chain locus V gene segments are human VL gene segments.

In one embodiment of the third configuration, the V gene segment repertoire of the light chain locus comprises or consists of one or more VL gene segments selected from the group consisting of a V<sub>λ</sub>II gene family member, V<sub>λ</sub>VII 4A, V<sub>λ</sub>II 2.1, V<sub>λ</sub>VII 4A, a V<sub>λ</sub>1 gene family member, a V<sub>λ</sub>3 gene family member, IGLV1S2, V<sub>λ</sub>3-cML70, Ialh2, Ialvl, Ia3h3, Kv325, a V<sub>κ</sub>I gene family member, κI-15A (KL012), V<sub>κ</sub>II family member, a V<sub>κ</sub>III family member, a V<sub>κ</sub>I gene family member, κI-15A (KL012), V<sub>κ</sub>II A2 (optionally the A2a allele), V<sub>κ</sub> A27 (Humkv325) and a gene segment at least 80% identical thereto; optionally each VL gene segment is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below; and/or wherein or in (a)(ii) the heavy chain locus V gene segment repertoire consists of only one (or two, three or four) VH gene segment type (optionally and one or mutants thereof), wherein the VH gene segment is selected from said group of VH gene segments. This is useful to bias the immune response of the vertebrate (and thus resultant lead antibodies) to a predetermined gene segment, eg, one known to be commonly used in natural human immune responses to antigens, such as antigens of infectious disease pathogens.

In one embodiment of the third configuration, in (a) said constant region is a heavy chain

endogenous non-human vertebrate (optionally host mouse or rat) constant region and/or in (b) said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

In one embodiment of the third configuration, endogenous heavy and light chain expression has been inactivated.

A fourth configuration of the present invention provides

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

(a) (i) An unrearranged immunoglobulin heavy chain locus comprising one or more human VL gene segments, one or more human D gene segments and one or more  $J_H$  gene segments upstream of a constant region, wherein each human VL gene segment is a human gene segment identical to (or mutant of, eg, having up to 15 or 10 nucleotide changes from the human gene segment) a human VL gene segment (eg, a germline VL gene segment; eg, a VL gene segment selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below) used to produce a rearranged VJ encoding a light chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism); or

(ii) An immunoglobulin heavy chain locus comprising a rearranged VJ region or VDJ region upstream of a constant region, wherein the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human gene segment identical to (or mutant of, eg, having up to 15 or 10 nucleotide changes from the human gene segment) a human VL gene segment (eg, a germline VL gene segment; eg, a VL gene segment selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below) used to produce a rearranged VJ encoding a light chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition

caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);

- (b) An immunoglobulin light chain locus comprising one or more human V gene segments (eg, a plurality of VL) and one or more human J gene segments upstream of a constant region; and
- (c) Wherein the gene segments in the light chain locus are operably linked to the constant region thereof, and the gene segments or VJ or VDJ in the heavy chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising light chains produced by recombination of the light chain locus and heavy chains derived from the heavy chain locus;
- (d) Optionally when (a)(i) applies, each said VL gene segment in the heavy chain locus is selected from the group consisting of a VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ I gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, Ialh2, Ialvl, Ia3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto; optionally each VL gene segment is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4 below;
- (e) Optionally when (a)(ii) applies, the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ I gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, Ialh2, Ialvl, Ia3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto.

The group of VL gene segments is useful to bias the immune response of the vertebrate (and thus resultant lead antibodies) to a predetermined gene segment, eg, one known to be commonly used in natural human immune responses to antigens, such as antigens of infectious disease pathogens.

In an embodiment of the fourth configuration, the VL gene segments of the heavy chain locus are V $\lambda$  gene segments.

In an embodiment of the fourth configuration, in (a) said constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

In an embodiment of the fourth configuration, in (b) said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

In an embodiment of the fourth configuration, the genome of said vertebrate or cell is homozygous for heavy chain locus (a)(i) or (ii); optionally wherein:

- the V gene segment repertoire of the heavy chain loci consists of one or more (or consists only of) human VL gene segments selected from the group consisting of a VL gene segment selected from the group consisting of a  $V_{\lambda}II$  gene family member,  $V_{\lambda}VII$  4A,  $V_{\lambda}II$  2.1,  $V_{\lambda}VII$  4A, a  $V_{\lambda}1$  gene family member, a  $V_{\lambda}3$  gene family member, IGLV1S2,  $V_{\lambda}3$ -cML70, Iah2, Ialvl, Ia3h3, Kv325, a  $V_{\kappa}I$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa}II$  family member, a  $V_{\kappa}III$  family member, a  $V_{\kappa}I$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa}II$  A2 (optionally the A2a allele),  $V_{\kappa}$  A27 (Humkv325) and a gene segment at least 80% identical thereto; or

- the recombined VJ or VDJ repertoire of the heavy chain loci consists of sequences identical to one or more nucleotide sequences produced by the recombination of a human VL gene segment selected from the group consisting of a VL gene segment selected from the group consisting of a  $V_{\lambda}II$  gene family member,  $V_{\lambda}VII$  4A,  $V_{\lambda}II$  2.1,  $V_{\lambda}VII$  4A, a  $V_{\lambda}1$  gene family member, a  $V_{\lambda}3$  gene family member, IGLV1S2,  $V_{\lambda}3$ -cML70, Iah2, Ialvl, Ia3h3, Kv325, a  $V_{\kappa}I$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa}II$  family member, a  $V_{\kappa}III$  family member, a  $V_{\kappa}I$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa}II$  A2 (optionally the A2a allele),  $V_{\kappa}$  A27 (Humkv325) and a gene segment at least 80% identical thereto with a human J gene segment and optionally a human D gene segment.

In an embodiment of the fourth configuration, endogenous heavy and light chain expression has been inactivated, and wherein heavy chain loci according to the fourth configuration are the only functional heavy chain loci in the genome of the vertebrate or cell.

In an embodiment of the fourth configuration, each immunoglobulin heavy chain locus of said vertebrate or cell is according to (a)(i) and comprises only a single human VL gene segment selected

from the group consisting of a VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ I gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, IalH2, IalVl, Ial3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto, optionally wherein the genome of the vertebrate or cell is homozygous for said heavy chain so that all heavy chain loci comprise the same, single human VL gene segment.

In an embodiment of the fourth configuration, each immunoglobulin heavy chain locus of said vertebrate or cell is according to (a)(ii) and comprises only a single rearranged VJ or VDJ region, optionally wherein the genome of the vertebrate or cell is homozygous for said heavy chain so that all heavy chain loci comprise the same, single rearranged VJ or VDJ region.

In an embodiment of the fourth configuration, each immunoglobulin heavy chain locus further comprises a VL gene segment or rearranged region that is a mutant respectively of said selected human VL gene segment or rearranged region, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain mutant VL gene segment or rearranged region.

In all configurations, aspects, examples and embodiments of the invention, where a "mutant" is mentioned, this includes a mutant sequence that is identical to a reference sequence (eg, reference VH, VL, VJ or VDJ) but with 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotide or amino acid changes therefrom.

In an embodiment of the fourth configuration, each immunoglobulin heavy chain locus comprises only two or three human VL gene segments selected from said group, optionally wherein the genome of the vertebrate or cell is homozygous for said two or three heavy chain human VL gene segments.

In an embodiment of the fourth configuration, each immunoglobulin heavy chain locus comprises only two or three of said rearranged VJ or VDJ regions, optionally wherein the genome of the vertebrate or cell is homozygous for said two or three heavy chain rearranged VJ or VDJ regions.

The invention provides a monoclonal or polyclonal antibody composition prepared by immunisation of at least one vertebrate (eg, mouse or rat) according to the third or fourth embodiment of the invention with an antigen, optionally wherein the antigen is an antigen of an infectious disease pathogen, optionally wherein the same antigen is used to immunise all the vertebrates; optionally wherein the antibody or antibodies are IgG-type.

The invention provides a third method: A method of isolating an antibody (eg, IgG-type, such as IgG1) that binds a predetermined antigen, the method comprising

- (a) providing a vertebrate (optionally a mouse or rat) according to the third or fourth embodiment of the invention;
- (b) immunising (eg, using standard prime-boost) said vertebrate with said antigen (optionally wherein the antigen is an antigen of an infectious disease pathogen);
- (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes expressing antibodies that bind to the antigen;
- (d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by producing hybridomas therefrom; and
- (e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes;
- (f) Optionally, the third method comprises the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally exchanging the heavy chain constant region nucleotide sequence of the antibody with a nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host.

Optionally, the third method further comprises making a mutant or derivative of the antibody.

The invention provides a fourth method: A method of producing a polyclonal antibody mixture, the method comprising carrying out the third method by separately immunising first and second vertebrates (optionally first and second mice or first and second rats) with antigen (eg, any antigen disclosed herein) and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally



wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):

- (i) the vertebrates are immunised with the same antigen or different antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or different family members thereof or different strains of the organism));
- (ii) prior to immunisation the heavy chain loci of the vertebrates contain the identical VL gene repertoire (optionally a single VL gene) and optionally the identical D and/or J repertoire; optionally the heavy chain loci of the mammals are identical prior to immunisation;
- (iii) prior to immunisation the heavy chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the heavy chain loci of the vertebrates are identical prior to immunisation.

The invention provides a fifth method: A method of producing a polyclonal antibody mixture, the method comprising carrying out the third method by immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):

- (i) the antigens are expressed by the same pathogenic organism (or different family members thereof or different strains of the organism);
- (ii) prior to immunisation the heavy chain loci of the vertebrates contain the identical VL gene repertoire (optionally a single VL gene) and optionally the identical D and/or J repertoire; optionally the heavy chain loci of the mammals are identical prior to immunisation;
- (iii) prior to immunisation the heavy chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the heavy chain loci of the vertebrates are identical prior to immunisation.

The invention provides a sixth method: A method of producing host cells capable of expressing a polyclonal antibody mixture, the method comprising, in the third method wherein step (f) is carried out:-

- (a) immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));
- (b) isolating nucleic acid encoding first and second anti-antigen antibodies from B lymphocytes from said vertebrates;
- (c) determining the nucleotide sequences of the heavy and light chain variable regions (optionally the entire heavy and/or light chain sequences) of the first antibody;
- (d) determining the nucleotide sequence of the light variable region and optionally the heavy chain variable region of the second antibody;
- (e) inserting the light chain variable region coding sequence of each antibody into a light chain expression vector; optionally wherein the constant region coding sequence of each light chain is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (f) inserting the heavy chain variable region coding sequence of the first antibody into a heavy chain expression vector; optionally wherein the constant region coding sequence of the heavy chain of the first antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (g) optionally inserting the heavy chain variable region coding sequence of the second antibody into a heavy chain expression vector; optionally wherein the constant region coding sequence of the heavy chain of the second antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region; and
- (h) introducing each expression vector into a host cell and co-expressing antibody chains in a mixture of said host cells to produce antibodies, each antibody comprising one or both of said light chain variable regions and a heavy chain; optionally wherein the expression vectors are introduced together into the same host cell (eg, a CHO or HEK293 cell) so that the cell is capable of expressing antibody light chains and heavy chains, such that the cell or a plurality of the host cells express antibodies (eg, two, three or four different antibodies), each comprising one or both of said light chain variable regions and a heavy chain;
- (i) optionally:  
prior to immunisation the heavy chain loci of the vertebrates contain the identical VL gene

repertoire (optionally a single VL gene segment) and optionally the identical D and/or J repertoire (optionally a single D and J gene segment); optionally the heavy chain loci of the vertebrates are identical prior to immunisation; or

prior to immunisation the heavy chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the heavy chain loci of the vertebrates are identical prior to immunisation.

The invention also provides a monoclonal or polyclonal antibody mixture so produced or a derivative antibody or mixture thereof, eg, where one or more constant region has been changed (eg, replaced with a different constant region such as a human constant region; or mutated to enhance or ablate Fc effector function).

The invention provides a seventh method: A method of producing a monoclonal antibody or polyclonal antibody mixture, the method comprising carrying out the sixth method and expressing a monoclonal antibody or polyclonal mixture of said antibodies; optionally followed by isolating an antibody comprising the light chain variable region of the first and/or second antibodies.

Optionally, each vertebrate used for immunisation is provided by

- (a) isolating from a human blood or tissue (eg, B lymphocytes, PBMCs, bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
- (b) determining which human germline VL gene segment was recombined in the human to produce the nucleotide sequence of said B lymphocyte that encodes the light chain variable region of the antibody;
- (c) constructing a transgenic vertebrate wherein said human germline VL gene segment is provided in a heavy chain locus thereof according to the third or fourth configuration of the invention; and
- (d) providing said transgenic vertebrate for immunisation in the fourth, fifth or sixth method of

the invention.

In another embodiment, each vertebrate used for immunisation is provided by

- (a) isolating from a human blood or tissue (eg, B lymphocytes, PBMCs, bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
- (b) determining a nucleotide sequence of said B lymphocyte that encodes a rearranged VDJ or VJ region of the antibody;
- (c) constructing a transgenic vertebrate wherein said rearranged VDJ or VJ region is provided in a heavy chain locus thereof according the third or fourth configuration of the invention; and
- (d) providing said transgenic vertebrate for immunisation in the method of the fourth, fifth or sixth method of the invention.

#### Common Heavy Chain Antibodies & Bispecifics (eg, to two pathogen antigens for infectious diseases)

The invention provides an isolated antibody (eg, IgG-type antibody) obtainable or obtained by the seventh method, or a mutant or derivative antibody thereof wherein (i) the isolated antibody comprises two copies of the heavy chain variable region of said first antibody paired with two copies of the light chain variable region of said first antibody; or (ii) the isolated antibody comprises two copies of the heavy chain variable region of said second antibody paired with two copies of the light chain variable region of said first antibody; or (iii) the isolated antibody is a bispecific antibody comprising one copy of the heavy chain variable region of said first antibody paired with a copy of the light chain variable region of the first antibody, and one copy of the heavy chain variable region of said the antibody paired with a copy of the light chain variable region of the first antibody, optionally wherein the bispecific antibody binds to said first and second antigens described above; optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease.

The invention provides a monoclonal or polyclonal antibody mixture (eg, IgG-type antibody or antibodies), wherein the monoclonal antibody or mixture comprises or consists of antibodies produced by the fourth, fifth, sixth or seventh method, or a mutant or derivative antibody thereof optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease, wherein optionally wherein each antibody binds an antigen of an infectious disease pathogen, preferably the same antigen.

The following embodiments relate to antibodies, host cells, nucleic acids and compositions and apply to such elements obtained or obtainable by any previous configuration or method of the invention:-

The invention provides an isolated chimaeric antibody for treating and/or preventing an infectious disease or condition, the antibody comprising a non-human vertebrate (optionally a mouse or rat) heavy chain constant regions and human variable regions that bind an antigen of an infectious disease pathogen, wherein the antibody is obtainable or obtained in a method comprising immunisation of a vertebrate according to of any one of the first to seventh methods of the invention with said antigen. The antigen is, for example, any antigen mentioned above. The disease or condition is, for example, any disease or condition mentioned above.

The invention provides an isolated human antibody for treating and/or preventing an infectious disease or condition, the antibody comprising human heavy chain constant regions and human variable regions that bind an antigen of an infectious disease pathogen, wherein the antibody is obtainable or obtained in a method comprising affinity maturation of antibody variable regions *in vivo* in a transgenic non-human vertebrate (eg, mouse or rat) when said variable regions are operably linked to heavy chain constant regions of said vertebrate (eg, mouse or rat heavy chain constant regions) by (a) immunisation of a vertebrate of any configuration of the invention with said antigen, (b) isolation of nucleic acid encoding a chimaeric antibody as described above, (c) replacing the nucleotide sequences of the nucleic acid that encode the non-human vertebrate heavy chain constant regions with nucleotide sequence encoding human heavy chain constant regions to produce nucleic acid encoding a human antibody; (d) expressing the human antibody *in vitro* (optionally from CHO or HEK293 cells harbouring the human nucleic acid) and (e) isolating the human antibody (optionally with further affinity maturation of the antibody and/or producing a

derivative thereof). The invention provides a mixture of first and second such human antibodies (an optionally also third and optionally fourth antibodies), each antibody being capable of binding to an antigen of an infectious disease pathogen (optionally wherein the first antibody binds a first antigen and the second antibody binds a second antigen, said antigens being from the same pathogen; or wherein the antigens are the same). Optionally, the light chain amino acid sequence of the first antibody is identical to the light chain amino acid sequence of the second antibody, or has up to 15 amino acid changes therefrom. The advantages of such a common (or closely-related) chain are explained above, and include relative ease of manufacture.

The antigen is, for example, any antigen mentioned above. The disease or condition is, for example, any disease or condition mentioned above. The pathogen is, for example, any pathogen mentioned above.

The invention provides an antibody comprising human variable domains that bind a predetermined antigen (eg, an antigen expressed by a bacterial or viral pathogen), wherein the variable domain sequences are encoded by rearranged VDJ and VJ regions, each of the VDJ and/or VJ being a hybrid region produced by the *in vivo* rearrangement of human heavy and light chain variable region gene segments (V and J and optionally D segments); optionally wherein the antibody comprises human constant regions.

The invention provides a method of producing an isolated human antibody for treating and/or preventing an infectious disease or condition, the antibody comprising human heavy chain constant regions and human variable regions that bind an antigen of an infectious disease pathogen, wherein the method comprises affinity maturing antibody variable regions *in vivo* in a transgenic non-human vertebrate (eg, mouse or rat) when said variable regions are operably linked to heavy chain constant regions of said vertebrate (eg, mouse or rat heavy chain constant regions) by (a) immunisation of a vertebrate of any configuration of the invention with said antigen, (b) isolation of nucleic acid encoding a chimaeric antibody as described above, (c) replacing the nucleotide sequences of the nucleic acid that encode the non-human vertebrate heavy chain constant regions with nucleotide sequence encoding human heavy chain constant regions to produce nucleic acid encoding a human antibody; (d) expressing the human antibody *in vitro* (optionally from CHO or HEK293 cells harbouring the human nucleic acid) and (e) isolating the human antibody (optionally with further

affinity maturation of the antibody and/or producing a derivative thereof). The antigen is, for example, any antigen mentioned above. The disease or condition is, for example, any disease or condition mentioned above. The pathogen is, for example, any pathogen mentioned above.

The invention provides the use of any isolated, monoclonal or polyclonal antibody or mixture of the invention as described above, in the manufacture of a medicament for the treatment and/or prevention of an infectious disease, optionally wherein the infectious disease is a disease caused by a bacterial or viral pathogen. The disease or condition is, for example, any disease or condition mentioned above. The pathogen is, for example, any pathogen mentioned above. For example, the infectious disease is selected from the group consisting of a disease caused by a pathogen selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus.

The invention provides first and second nucleotide sequences (eg , DNA, RNA, mRNA, cDNA) encoding the heavy and light chains of an antibody according to any configuration, aspect, example or embodiment of the invention or at least the variable regions thereof, optionally wherein each nucleotide sequence is part of a vector.

The invention provides a host cell comprising one or more expression vectors encoding the heavy chains of the first and second antibodies mentioned above, and the light chain of the first antibody mentioned above (and optionally also the light chain of the second antibody). Again, reference is made to the discussion above about the advantages of having a common antibody chain for the production of antibody mixtures.

The invention provides a pharmaceutical composition comprising the antibody or antibodies of any configuration, aspect, example or embodiment of the invention and a diluent, excipient or carrier; optionally wherein the composition is provided in a container connected to an IV needle or syringe or in an IV bag. The skilled person will know standard diluents, excipients and carriers suitable for pharmaceutical application.

Throughout this description, where it is mentioned “at least 80% identical”, there is contemplated in the alternative one of the following identities: at least 85%, 90, 95, 96, 97, 98 or 99 identical and the disclosure herein contemplates that one or more of these identities may be recited in a claim herein in place of “at least 80% identical”.

### **Tailoring V(D)J Incorporation Into Immunoglobulin Loci For The Generation of Antibodies Against Infectious Disease**

In the various configurations, aspects, embodiments and examples above, the invention provides the skilled addressee with the possibility of choosing immunoglobulin gene segments in a way that tailors or biases the repertoire for application to generating antibodies to treat and/or prevent infectious diseases. The inventors have categorised the following groups of gene segments for use in the invention according to the desired application of resultant antibodies.

#### List A:

#### Immunoglobulin gene segments for antibodies that bind an antigen expressed by a Pathogen

- (a) a VL gene segment selected from the group consisting of a V<sub>λ</sub>II gene family member, V<sub>λ</sub>VII 4A, V<sub>λ</sub>II 2.1, V<sub>λ</sub>VII 4A, a V<sub>λ</sub>1 gene family member, a V<sub>λ</sub>3 gene family member, IGLV1S2, V<sub>λ</sub>3-cML70, Iah2, Ialvl, Ia3h3, Kv325, a VκI gene family member, κI-15A (KL012), VκII family member, a VκIII family member, a VκI gene family member, κI-15A (KL012), VκII A2 (optionally the A2a allele), Vκ A27 (Humkv325) and a gene segment at least 80% identical thereto.
- (b) a V<sub>λ</sub> gene segment selected from a V<sub>λ</sub>II gene family member, V<sub>λ</sub>VII 4A, V<sub>λ</sub>II 2.1, V<sub>λ</sub>VII 4A, a V<sub>λ</sub>1 gene family member, a V<sub>λ</sub>3 gene family member, IGLV1S2, V<sub>λ</sub>3-cML70, Iah2, Ialvl, Ia3h3 and a gene segment at least 80% identical thereto.
- (c) a Vκ gene segment selected from Kv325, a VκI gene family member, κI-15A (KL012), VκII family member, a VκIII family member, a VκI gene family member, κI-15A (KL012), VκII A2 (optionally the A2a allele), Vκ A27 (Humkv325) and a gene segment at least 80% identical thereto.
- (d) a V<sub>H</sub> gene segment a V<sub>H</sub>III gene family member (optionally, a VHIIIa or VHIIIb family member), a V<sub>H</sub>IV gene family member, V<sub>H</sub>III 9.1 (VH3-15), V<sub>H</sub>III VH26 (VH3-23), V<sub>H</sub>3-21, LSG6.1, LSG12.1,



DP77 (V3-21), V<sub>H</sub> H11, VH1GRR, ha3h2, V<sub>H</sub>I-ha1c1, V<sub>H</sub>III-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto.

(e) a J<sub>λ</sub> gene segment selected from J<sub>λ</sub>2, J<sub>λ</sub>3 and a gene segment at least 80% identical thereto.

(f) a D gene segment selected from Dk1, Dxp'1, Dn4r, D2r and a gene segment at least 80% identical thereto.

List A1:

Immunoglobulin gene segments for antibodies that bind an antigen expressed by a Bacterial Pathogen

(a) a V<sub>λ</sub> gene segment selected from a V<sub>λ</sub>II gene family member, V<sub>λ</sub>VII 4A, V<sub>λ</sub>II 2.1, V<sub>λ</sub>VII 4A and a gene segment at least 80% identical thereto.

(b) a V<sub>κ</sub> gene segment selected from a VκI gene family member, κI-15A (KL012), V<sub>κ</sub>II family member, a V<sub>κ</sub>III family member, a VκI gene family member, κI-15A (KL012), V<sub>κ</sub>II A2 (optionally the A2a allele), V<sub>κ</sub> A27 (Humkv325) and a gene segment at least 80% identical thereto.

(c) a V<sub>H</sub> gene segment a VH3 gene family member (optionally, a VHIIIa or VHIIIb family member), V<sub>H</sub>III 9.1 (VH3-15), V<sub>H</sub>III VH26 (VH3-23), V<sub>H</sub>3-21, LSG6.1, LSG12.1, DP77 (V3-21), V<sub>H</sub> H11 and a gene segment at least 80% identical thereto.

(d) a J<sub>λ</sub> gene segment selected from J<sub>λ</sub>2, J<sub>λ</sub>3 and a gene segment at least 80% identical thereto.

(e) a J<sub>H</sub> gene segment selected from J<sub>H</sub>2, J<sub>H</sub>3, J<sub>H</sub>4 and a gene segment at least 80% identical thereto.

List A1.1:

Immunoglobulin gene segments for antibodies that bind an antigen expressed by *H influenza*

- (a) a  $V_{\lambda}$  gene segment selected from a  $V_{\lambda}$ II gene family member,  $V_{\lambda}$ VII 4A,  $V_{\lambda}$ II 2.1,  $V_{\lambda}$ VII 4A and a gene segment at least 80% identical thereto.
- (b) a  $V_{\kappa}$  gene segment selected from a  $V_{\kappa}$ II family member, a  $V_{\kappa}$ III family member, a  $V_{\kappa}$ I gene family member,  $\kappa$ I-15A (KL012),  $V_{\kappa}$ II A2 (optionally the A2a allele),  $V_{\kappa}$  A27 (Humkv325) and a gene segment at least 80% identical thereto.
- (c) a  $V_H$  gene segment a VH3 gene family member (optionally, a VHIIIb family member),  $V_H$ III 9.1 (VH3-15),  $V_H$ III VH26 (VH3-23),  $V_H$ 3-21, LSG6.1, LSG12.1, DP77 (V3-21) and a gene segment at least 80% identical thereto.
- (d) a  $J_{\lambda}$  gene segment selected from  $J_{\lambda}2$ ,  $J_{\lambda}3$  and a gene segment at least 80% identical thereto.

List A1.2:

Immunoglobulin gene segments for antibodies that bind an antigen expressed by E Coli or Neisseria meningitidis

- (a) a  $V_H$  gene segment a VH3 gene family member (optionally a VHIIIa or VHIIIb member),  $V_H$ III 9.1 (VH3-15),  $V_H$  H11,  $V_H$ III VH26 (VH3-23) a gene segment at least 80% identical thereto, eg,  $V_H$ III 9.1 +  $J_H3$ ; or  $V_H$  H11 +  $J_H4$ ; or  $V_H$ III VH26 +  $J_H2$ .
- (b) a  $V_{\kappa}$  gene segment selected from a  $V_{\kappa}$ I gene family member,  $\kappa$ I-15A (KL012) and a gene segment at least 80% identical thereto.
- (c) a  $V_{\lambda}$  gene segment selected from a  $V_{\lambda}$ II gene family member,  $V_{\lambda}$ II 2.1 and a gene segment at least 80% identical thereto.
- (d) a  $J_H$  gene segment selected from  $J_H2$ ,  $J_H3$ ,  $J_H4$  and a gene segment at least 80% identical thereto.

A2:

Immunoglobulin gene segments for antibodies that bind an antigen expressed by a viral Pathogen

- (a) a  $V_H$  gene segment selected from a  $V_{HIII}$  gene family member, a  $V_{HIV}$  gene family member,  $V_{HIII}$ -VH26 (VH3-23), VH1GRR, ha3h2,  $V_{HI}$ -ha1c1,  $V_{HIII}$ -VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto.
- (b) a  $V_\lambda$  gene segment selected from a  $V_{\lambda 1}$  gene family member, a  $V_{\lambda 3}$  gene family member, IGLV1S2,  $V_{\lambda 3}$ -cML70, lalh2, lalvl, la3h3 and a gene segment at least 80% identical thereto.
- (c) a  $V_k$  gene segment selected from Kv325 and a gene segment at least 80% identical thereto.
- (d) a  $J_H$  gene segment selected from  $J_{H3}$ ,  $J_{H5}$ ,  $J_{H6}$  and a gene segment at least 80% identical thereto.
- (e) a D gene segment selected from Dk1, Dxp'1, Dn4r, D2r and a gene segment at least 80% identical thereto.
- (f) a  $J_\lambda$  gene segment selected from  $J_{\lambda 2}$ ,  $J_{\lambda 3}$  and a gene segment at least 80% identical thereto.

#### A2.1:

#### Immunoglobulin gene segments for antibodies that bind an antigen expressed by Herpes Virus

#### Family (eg, VZV or HSV)

- (a) a  $V_H$  gene segment selected from a  $V_{HIII}$  gene family member, a  $V_{HIV}$  gene family member,  $V_{HIII}$ -VH26 (VH3-23), VH1GRR, ha3h2,  $V_{HI}$ -ha1c1,  $V_{HIII}$ -VH2-1, VH4.18, ha4h3, and a gene segment at least 80% identical thereto.
- (b) a  $V_\lambda$  gene segment selected from a  $V_{\lambda 1}$  gene family member, a  $V_{\lambda 3}$  gene family member, IGLV1S2,  $V_{\lambda 3}$ -cML70, lalh2, lalvl, la3h3 and a gene segment at least 80% identical thereto.
- (c) a  $J_H$  gene segment selected from  $J_{H3}$ ,  $J_{H5}$ ,  $J_{H6}$  and a gene segment at least 80% identical thereto.
- (d) a D gene segment selected from Dk1, Dxp'1, Dn4r, D2r and a gene segment at least 80% identical thereto.
- (e) a  $J_\lambda$  gene segment selected from  $J_{\lambda 2}$ ,  $J_{\lambda 3}$  and a gene segment at least 80% identical thereto.

A2.2:Immunoglobulin gene segments for antibodies that bind an antigen expressed by CMV

- (a) a V<sub>H</sub> gene segment selected from Hv1051 and a gene segment at least 80% identical thereto.
- (b) a V<sub>K</sub> gene segment selected from Kv325 and a gene segment at least 80% identical thereto.

A2.3:Immunoglobulin gene segments for antibodies that bind an antigen expressed by HIV

- (a) a V<sub>H</sub> gene segment selected from 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto.

A2.4:Immunoglobulin gene segments for antibodies that bind an antigen expressed by Influenza Virus

- (a) a V<sub>H</sub> gene segment selected from VH1-69 and a gene segment at least 80% identical thereto.

Thus,

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease, one or more V, D and/or or all J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A1. Thus, for example in (a) of the first configuration of the invention, the recited heavy chain V gene segment is selected from the VH gene segments in List A, optionally with a D in that list.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by a bacterial pathogen, one or more or all V, D and/or J gene

segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A1.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by a viral pathogen, one or more or all V, D and/or J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A2.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by *H influenza*, one or more or all V, D and/or J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A1.1.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by *E Coli* or *Neisseria meningitidis*, one or more or all V, D and/or J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A1.2.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by Herpes Virus Family (eg, VZV or HSV), one or more or all V, D and/or J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A2.1.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by CMV, one or more or all V, D and/or J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A2.2.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by HIV, one or more or all V, D and/or J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A2.3.

Where one wishes to generate an antibody or antibody mixture to treat and/or prevent an infectious disease caused or mediated by Influenza Virus, one or more or all V, D and/or J gene segments used in any configuration, aspect, method, example or embodiment of the invention can be selected from List A2.4.

Optionally each VH segment in the locus of the invention is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4.

Optionally each VL segment in the locus of the invention is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4

Optionally each D segment in the locus of the invention is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4.

Optionally each J<sub>L</sub> segment in the locus of the invention is selected from List A1, A2, A1.1, A1.2, A2.1, A2.2, A2.3 or A2.4.

### **Long HCDR3 Binding sites & Tailoring Gene Segments to Pathogens & Other Antigens**

This aspect of the invention relates to the development of vertebrates, cells, methods and antibodies with relatively long HCDR3 binding sites. There is also provided embodiments in which genomes and antibodies are tailored in terms of their gene segments usage to address infectious disease pathogen antigens or other antigens which are advantageously addressed with a longer HCDR3 length for binding or neutralisation. Antibodies may be raised in the vertebrates by

immunisation with a non-pathogen target antigen, eg, an antigen bearing an epitope in a cleft requiring a long CDR for contact, or an antigen from a pathogen that causes or is implicated in harmful human disease or conditions. Examples are bacterial or viral pathogens and the target antigen may be a bacterial cell-surface antigen or a viral surface-exposed antigen (eg, coat protein). Additionally or alternatively, the antigen may be an antigen that is released (eg, secreted) from a pathogenic bacterium or virus. The invention is not limited to addressing pathogen antigens, but is also useful for addressing other antigens where a long CDR3 would be useful for binding (eg, an enzyme active site or receptor cleft).

Antibodies with long HCDR3 (at least 20 amino acids according to IMGT nomenclature) have been shown to effectively neutralise a variety of pathogens including HIV, Influenza virus, malaria and Africa trypanosomes. Reference is also made to naturally-occurring Camelid (eg, llama or camel) heavy chain-only antibodies which bear long HCDR3s for reaching relatively inaccessible epitopes (see, eg, EP0937140). Long HCDR3s can form unique stable subdomains with extended loop structure that towers above the antibody surface to confer fine specificity. In some cases, the long HCDR3 itself is sufficient for epitope binding and neutralization (Liu, L *et al*; Journal of Virology. 2011. 85: 8467-8476, incorporated herein by reference). The unique structure of the long HCDR3 allows it to bind to cognate epitopes within inaccessible structure or extensive glycosylation on a pathogen surface. In human peripheral blood, there is around 3.5% of naïve B antibodies or 1.9% of memory B IgG antibodies containing the HCDR3s with lengths of more than 24 amino acids (Briney, BS *et al*, referenced given below) (Fig. 1 of Briney, BS *et al*). The usage analysis indicates that these antibodies have the preference to use human VH1-69, D2-2, D3-3, D2-15 and JH6 segments (Figs. 2-5 of Briney, BS *et al*). There are around 20% of all HCDR3 length antibodies using JH6. However, in those antibodies with more than 24 amino acids of HCDR3, there are 70% using JH6 (Fig.2 of Briney, BS *et al*). Human VH5-51 is also commonly used for anti-HIV antibodies (see Gorny *et al*, PLoS One. 2011;6(12):e27780. Epub 2011 Dec 2.

Human anti-V3 HIV-1 monoclonal antibodies encoded by the VH5-51/VL lambda genes define a conserved antigenic structure, incorporated herein by reference).

Supplementing these observations, the inventors have found (see examples) that other selected human heavy chain variable region gene segments (V, D, J) recombine in transgenic non-human vertebrates to produce long HCDR3 (at least 20 amino acids).

Thus, as explained further in the examples, the inventors constructed transgenic IgH loci in ES cells, wherein the loci purposely included selected human heavy chain variable region gene segments (V, D, J) that recombine to produce long HCDR3 (at least 20 amino acids). From the ES cells, the inventors generated transgenic non-human vertebrates (both naïve and immunised with a range of different target antigen types – disease pathogen and human antigenic species), isolated antibodies and heavy chain sequences based on the selected gene segments as well as B-cells expressing these and made hybridomas expressing antigen-specific antibodies that are based on the selected gene segments.

There is a need in the art for genetically modified non-human animals that prefer to make human antibodies that have long HCDR3s, as well as antibodies that can be selected from such animals wherein the antibodies can address target epitopes more easily accessed by long HCDR3s. Long CDRH3 is also useful for penetrating highly glycan-covered epitope sites (eg, virus epitopes or any glycoprotein targets, eg, see Nature. 2011 Dec 14;480(7377):324-5. doi: 10.1038/480324a; Vaccinology: “A sweet cleft in HIV's armour”, Sattentau QJ, incorporated herein by reference), and the target antigen can comprise such a target epitope.

The present invention provides vertebrates that can artificially simulate those naturally-occurring human long HCDR3 antibodies, and can provide antibody, heavy chain and variable domain repertoires from which can be selected an antibody, heavy chain or variable domain having a long HCDR3 (eg, having a HCDR3 length of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more amino acids (according to IMGT). The invention provides for the combination of human VH, D and J gene repertoires upstream of non-human vertebrate (eg, mouse or rat, eg, endogenous mouse or rat) constant region in heavy chain loci comprised by the vertebrate genomes. This enables the recombination, maturation and selection of the human gene segments in the context of endogenous or other non-vertebrate constant regions which enhances the development of good sized antibody, heavy chain and variable domain repertoires from which to select long HCDR3-type binding sites. Thus, in an example of any configuration of the invention, the human gene segments are provided in a heavy chain locus upstream of a non-human vertebrate (eg, endogenous) constant region. Similarly any antibody of the invention comprises human variable domains and non-human



vertebrate (eg, endogenous) domains. The latter can be replaced by human constant domains after selection and isolation.

For example, the following antibodies of the invention are contemplated (eg, produced in a vertebrate of this aspect of the invention by a method disclosed herein) or a copy or derivative of an antibody so produced:-

An isolated, synthetic or recombinant antibody comprising human heavy chain variable domains having a HCDR3 length of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more amino acids (according to IMGT), the heavy chain variable domains being derived from the recombination of a human VH gene segment selected from a VH group disclosed herein with a human D gene segment and a human JH gene segment (optionally a JH6), wherein the antibody binds a target antigen; wherein the heavy chain variable domains have non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern junctional mutations when compared to corresponding human germline V, D and J sequences. In an example, the antibody of the invention has a HCDR3 length of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 amino acids (according to IMGT). In an example, the antigen is an antigen of a pathogen that causes or is implicated in a human infectious disease or condition, eg, a pathogen listed in Table 1. In an example, the antibody specifically binds an active site or cleft of an antigen (eg, an enzyme active site or receptor cleft). This can be determined, eg, using standard X-ray crystallography of a complex of the antibody (or heavy chain or VH domain) with the cognate antigen, as is known to the skilled person.

Mouse AID-pattern somatic hypermutations and/or mouse dTdT-pattern mutations can be provided, for example, wherein VH domain is produced in a mouse comprising mouse AID and/or mouse TdT (eg, endogenous AID or TdT). See also Annu. Rev. Biochem. 2007. 76:1–22; Javier M. Di Noia and Michael S. Neuberger, “Molecular Mechanisms of Antibody Somatic Hypermutation” (in particular figure 1 and associated discussion on AID hotspots in mouse); and Curr Opin Immunol. 1995 Apr;7(2):248-54, “Somatic hypermutation”, Neuberger MS and Milstein C (in particular, discussion on hotspots in mouse), the disclosures of which are incorporated herein by reference. Such mice can be made using corresponding mouse ES cell technology.

In an example, the antibody specifically binds to a HIV antigen. Several naturally-occurring human antibodies are known to be neutralising of HIV and have rather long HCDR3 lengths (20 amino acids or more according to IMGT; see Breden *et al*, PLoS One. 2011 Mar 30;6(3):e16857; "Comparison of antibody repertoires produced by HIV-1 infection, other chronic and acute infections, and systemic autoimmune disease" (incorporated herein by reference) - VH1-69 preferred for long HCDR3). See also PLoS One. 2012;7(5):e36750. Epub 2012 May 9; "Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes"; Briney BS *et al* (incorporated herein by reference). Thus, it is desirable to provide antibodies of the invention that have similarly long HCDR3 lengths. The antibody of the invention is, in one example, provided for treating and/or preventing HIV infection, eg, chronic HIV infection, in a human. The invention also provides a method of treating and/or preventing HIV infection, eg, chronic HIV infection, in a human, the method comprising administering a pharmaceutically acceptable dose of the antibody of the invention. The dose can be split into one or more administration aliquots, eg, administered over a time course according to a medically-determined regimen, as the skilled person will be able to determine.

In an example, the antibody specifically binds to *Hemophilus influenzae* type b polysaccharide. The antibody of the invention is, in one example, provided for treating and/or preventing *Hemophilus influenzae* infection, eg, chronic *Hemophilus influenzae* infection, in a human. The invention also provides a method of treating and/or preventing *Hemophilus influenzae* infection, eg, chronic *Hemophilus influenzae* infection, in a human, the method comprising administering a pharmaceutically acceptable dose of the antibody of the invention. The dose can be split into one or more administration aliquots, eg, administered over a time course according to a medically-determined regimen, as the skilled person will be able to determine.

In an example, the antibody specifically binds to a rotavirus antigen (eg, protein 6 or 7). The antibody of the invention is, in one example, provided for treating and/or preventing rotavirus infection, eg, chronic rotavirus infection, in a human. The invention also provides a method of treating and/or preventing rotavirus infection, eg, chronic rotavirus infection, in a human, the method comprising administering a pharmaceutically acceptable dose of the antibody of the invention. The dose can be split into one or more administration aliquots, eg, administered over a

time course according to a medically-determined regimen, as the skilled person will be able to determine.

In an example, the antibody specifically binds to a cytomegalovirus antigen (eg, cytomegalovirus gB antigen). The antibody of the invention is, in one example, provided for treating and/or preventing cytomegalovirus infection, eg, chronic cytomegalovirus infection, in a human. The invention also provides a method of treating and/or preventing cytomegalovirus infection, eg, chronic cytomegalovirus infection, in a human, the method comprising administering a pharmaceutically acceptable dose of the antibody of the invention. The dose can be split into one or more administration aliquots, eg, administered over a time course according to a medically-determined regimen, as the skilled person will be able to determine.

The invention also provides a vertebrate or cell for expressing such an antibody; thus the invention provides a non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin VH gene segment repertoire that is biased to one, more or all human VH gene segments selected from a VH group disclosed herein.

The invention also provides

a method of isolating an antibody that binds a HIV antigen, *Hemophilus influenza* type b polysaccharide, cytomegalovirus antigen or rotavirus antigen, the method comprising

- (a) providing the human VH biased vertebrate of the invention;
- (b) immunising said vertebrate with said HIV antigen, *Hemophilus influenza* type b polysaccharide, cytomegalovirus antigen or rotavirus antigen;
- (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes expressing antibodies that bind to the antigen;
- (d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by producing hybridomas therefrom; and
- (e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes, wherein the antibody has a HCDR3 length of 20 amino acids or more.

Optionally, the method further comprises the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally exchanging the heavy chain constant region nucleotide sequence of the antibody with a nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host.

Optionally, the method further comprises making a copy, mutant or derivative (eg, humanised version) of the antibody produced by the method.

This aspect of the invention also provides

A pharmaceutical composition comprising the anti-HIV antibody, for treating and/or preventing HIV in a human (eg, an infant human).

A pharmaceutical composition comprising the anti-*Hemophilus influenza* type b polysaccharide antibody, for treating and/or preventing Haemophilus influenza in a human (eg, an infant human).

A pharmaceutical composition comprising the anti-rotavirus antibody, for treating and/or preventing rotavirus in a human (eg, an infant human).

A pharmaceutical composition comprising the anti-cytomegalovirus antibody, for treating and/or preventing cytomegalovirus in a human (eg, an infant human).

The invention also provides a method of generating such an antibody (eg, any one of embodiments (i) *et seq* above) by immunising a vertebrate of the invention with the target antigen and isolating the antibody from the vertebrate, optionally also making a copy or derivative of the antibody. In a further step, a B-cell capable of expressing the antibody is isolated from the vertebrate. In a further step, a nucleic acid encoding the antibody (or a VH domain thereof) is isolated from the vertebrate (eg, a nucleic acid PCR cloned from a B-cell isolated from the vertebrate).

In an example, the antibody of the invention is a neutralising antibody. In an example, the antibody of the invention has a HCDR3 length of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more amino acids (according to IMGT). In an example, the antibody of the invention has a HCDR3 length of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 amino acids (according to IMGT). In an example, the antibody of the invention is isolated from a non-human vertebrate (eg, a mouse or a rat), for example a vertebrate of the invention; or the antibody is a copy or derivative (eg, humanised version) thereof. In an example, the antibody of the invention has non-human vertebrate constant regions (eg, mouse or rat constant regions); these may be replaced using standard recombinant DNA technology with human constant regions, so the invention also provides for human versions of the antibodies recited above, wherein the human antibody comprises human variable and constant regions, wherein the variable regions bind the antigen. In an example, the antibody of the invention has lambda-type human light chain variable domains. In another example, the antibody of the invention has kappa-type human light chain variable domains.

Antibody competition can be determined, for example, by ELISA or surface plasmon resonance (SPR; eg, by competition Biacore™ or Proteon) as is standard.

The invention also provides the following embodiments (recited below as numbered clauses):-

#### D Bias

1. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin D gene segment repertoire that is biased to the human D2 and/or D3 family or biased to one, more or all human D gene segments selected from the group D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.

For example, the repertoire consists of only human D gene segments from the D2 and/or D3 family.

Optionally the repertoire is biased to one or more of human D2-2, D2-15, D3-3, D3-9, D3-10 and D3-22, or the repertoire consists of one, more or all of these D gene segments. These produce long

HCDR3 lengths (eg, see Table 2 and references cited herein).

For example, the repertoire is biased to one or more of human D2-2\*02, D3-9\*01, D3-10\*01 and D3-22\*01, or the repertoire consists of one, more or all of these D gene segments.

For example, the repertoire is biased to one or more of human D2-2\*02, D3-9\*01 and D3-10\*01, or the repertoire consists of one, more or all of these D gene segments.

For example, the repertoire is biased to D3-9\*01 and D3-10\*01, or consists of one, more or all of these D gene segments.

Optionally the repertoire consists of one, more or all of human D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19. These produce long HCDR3 lengths (eg, see Table 2).

Optionally the repertoire is biased to one or more of human D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01, or the repertoire consists of one, more or all of these D gene segments.

Optionally the repertoire is biased to one or more of human D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or the repertoire consists of one, more or all of these D gene segments. Optionally the repertoire is biased to one or more of human D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01, or the repertoire consists of one, more or all of these D gene segments. These produce long HCDR3 lengths in naïve repertoires (eg, see Table 2).

Optionally the repertoire is biased to one or more of human D1-26, D2-2, D3-10 and D6-19, or the repertoire consists of one, more or all of these D gene segments. Optionally the repertoire is biased to one or more of human D1-26\*01, D2-2\*02, D3-10\*01 and D6-19\*01, or the repertoire consists of one, more or all of these D gene segments. These produce long HCDR3 lengths in immunised repertoires (eg, see Table 2).

Optionally the repertoire is biased to one or more of human D2-2, D3-9 and D3-10, or the repertoire consists of one, more or all of these D gene segments. Optionally the repertoire is biased to one or more of human D2-2\*02, D3-9\*01 and D3-10\*01, or the repertoire consists of one, more or all of these D gene segments. These produce long HCDR3 lengths in antigen-specific repertoires (eg, see

Table 2).

IMGT nomenclature is used for all gene segments.

Throughout this text, Genbank is a reference to Genbank release number 185.0 or 191.0; the 1000 Genomes database is Phase 1, release v3, 16<sup>th</sup> March 2012; the Ensembl database is assembly GRCh37.p8 (10/04/2012); the IMGT database is available at [www.imgt.org](http://www.imgt.org). The sequences of all VH gene segments explicitly mentioned herein are disclosed herein in their entirety (for possible inclusion in claims in conjunction with any aspect of the invention as claimed), such sequences being those in the IMGT and 1000 Genomes databases.

In one embodiment, the genome comprises an IgH locus comprising a targeted insertion of said human D gene segments. In an example, the IgH locus comprises (in 5' to 3' order) one or more human VH gene segments, said D gene segment repertoire, one or more human JH gene segments and a constant region (eg, wherein the constant region is a human constant region or a non-human (eg, endogenous, eg, mouse or rat) constant region).

In another embodiment, the genome comprises said human D gene segments randomly inserted therein. This can be effected, eg, by incorporating human DNA borne by YACS into the genome of ES cells (followed optionally by generation of a non-human vertebrate therefrom, as is standard).

Optionally, the human D gene segment repertoire further comprises no more than 5 additional human D gene segments, for example, the repertoire includes 1, 2, 3, 4 or 5 additional human D gene segments.

2. The vertebrate or cell of clause 1, wherein the D gene segment repertoire consists of or substantially consists of one, two or three human gene segments selected from the group D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.
3. The vertebrate or cell of clause 1 or 2, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising (in 5' to 3' order) human VH, D and JH gene segments and said human D gene segments recited in clause 1 are spaced from the VH gene segment(s) by no more than four other D gene segments (eg, by no D gene segments).

This provides for bias wherein proximal D gene segments (those more 3', ie, closer to the constant region) are likely to be more frequently used than those segments from distal (ie, 5' or further away from the constant region).

4. The vertebrate or cell of any preceding clause, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising said human D gene segments and there are no other D gene segments in the locus between said human D gene segments.

This is another way of biasing the repertoire of D gene segments. Thus, the desired Ds are provided in tandem, aimed to promote use in recombination.

5. The vertebrate or cell of any preceding clause, wherein the genome comprises three or more copies of a human D gene segment selected from D1-26, D2-2, D2-15, D3-3, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.

For example, the genome comprises three or more copies of a human D gene segment selected from D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.

This is another way of biasing the repertoire of D gene segments.

6. The vertebrate or cell of clause 5, wherein the genome comprises first and second human D gene segments selected from D1-26, D2-2, D2-15, D3-3, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19 when the first D gene segment is present as three or more copies and wherein the second D gene segment is present as three or more copies.

For example, the first and second gene segments are selected from D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.

The various gene segment biasing techniques described herein can be performed using conventional DNA manipulation in the construction of transgenic vertebrates or cells of the invention, which techniques (eg, recombineering and recombinant DNA technology) will be known to the skilled person. For example, BACs can be constructed using these techniques in



which the desired combination of human gene segments is provided, and these BACs can be introduced into ES cells for incorporation of the human gene segments into the genomes thereof (eg, by targeted insertion into Ig loci). The ES cells can be used to generate transgenic vertebrates as is standard and cells (eg, B-cells) can be isolated from these wherein the genome is as per the invention.

In one embodiment, the biased D gene segment(s) are selected from the IMGT database of variants or the 1000 Genomes database.

7. The vertebrate or cell of any preceding clause, wherein the D gene segments are selected from D2-2\*02, D3-9\*01, D3-10\*01 and D3-22\*01, or selected from D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01.

#### VH Bias

8. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell), optionally according to any preceding clause, whose genome comprises a human immunoglobulin VH gene segment repertoire that is biased to one, more or all of gene segments selected from the group VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.

These produce long HCDR3 lengths (see Table 2 and references cited herein).

For example, the VH repertoire is biased to one, more or all of VH1-2, VH1-3, VH1-8, VH1-18, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1. These produce long HCDR3 lengths (see Table 2), or the repertoire consists of one, more or all of these VH gene segments. For example, the VH repertoire is biased to one, more or all of VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01, or the repertoire consists of one, more or all of these VH gene segments.

For example, the VH repertoire is biased to one, more or all of VH1-2\*02, VH1-8\*01, VH1-18\*01, VH1-3\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01, or the repertoire consists of one, more or all of these VH gene segments. These produce long HCDR3 lengths in naïve repertoires (see Table 2).

For example, the VH repertoire is biased to one, more or all of VH4-4\*02, VH3-11\*01 and VH3-7\*01, or the repertoire consists of one, more or all of these VH gene segments. These produce long HCDR3 lengths in immunised repertoires (see Table 2).

For example, the VH repertoire is biased to one, more or all of VH1-3\*01, VH1-8\*01, VH3-7\*01, VH3-9\*01, VH3-11\*01 and VH4-4\*02, or the repertoire consists of one, more or all of these VH gene segments. These produce long HCDR3 lengths in antigen-specific repertoires (see Table 2).

Optionally, the human VH gene segment repertoire further comprises no more than 5 additional human VH gene segments, for example, the repertoire includes 1, 2, 3, 4 or 5 additional human VH gene segments.

In one embodiment, the genome comprises an IgH locus comprising a targeted insertion of said human VH gene segments. In an example, the IgH locus comprises (in 5' to 3' order) said VH gene segment repertoire, one or more human D gene segments, one or more human JH gene segments and a constant region (eg, wherein the constant region is a human constant region or a non-human (eg, endogenous, eg, mouse or rat) constant region).

In another embodiment, the genome comprises said human VH gene segments randomly inserted therein. This can be effected, eg, by incorporating human DNA borne by YACS into the genome of ES cells (followed optionally by generation of a non-human vertebrate therefrom, as is standard).

9. The vertebrate or cell of clause 8, wherein the VH gene segment repertoire substantially consists of or substantially consists of one, two or three human gene segments selected from VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.

For example, the VH gene segment repertoire substantially consists of or substantially consists of one, two or three human gene segments selected from the group consisting of VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01.

10. The vertebrate or cell of clause 8 or 9, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising (in 5' to 3' order) human VH, D and JH gene segments and said human VH gene segments are spaced from the D gene segment(s) by no more than four other VH gene segments (eg, by no VH gene segments).

This provides for bias wherein proximal VH gene segments (those more 3', ie, closer to the constant region) are likely to be more frequently used than those segments from distal (ie, 5' or further away from the constant region).

11. The vertebrate or cell of any one of clauses 8 to 10, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising said human VH gene segments and there are no other VH gene segments in the locus between said human VH gene segments.

This is another way of biasing the repertoire of VH gene segments.

12. The vertebrate or cell of any one of clauses 8 to 11, wherein the genome comprises three or more copies of a human VH gene segment selected from the group consisting of VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.

For example, the genome comprises three or more copies of a human VH gene segment selected from the group consisting of VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01.

This is another way of biasing the repertoire of VH gene segments.

13. The vertebrate or cell of clause 12, wherein the genome comprises first and second human VH gene segments selected from the group consisting of VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1 when the first VH gene segment is present as three or more copies and wherein the second VH gene segment is present as three or more copies.

For example, the genome comprises first and second human VH gene segments selected from the group consisting of VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01 when the first VH gene segment is present as three or more copies and wherein the second VH gene segment is present as three or more copies.

In an embodiment, all or substantially all of VH gene segments are present as three or more copies each.

The various gene segment biasing techniques described herein can be performed using conventional DNA manipulation in the construction of transgenic vertebrates or cells of the invention, which techniques (eg, recombineering and recombinant DNA technology) will be known to the skilled person. For example, BACs can be constructed using these techniques in which the desired combination of human gene segments is provided, and these BACs can be introduced into ES cells for incorporation of the human gene segments into the genomes thereof (eg, by targeted insertion into Ig loci). The ES cells can be used to generate transgenic vertebrates as is standard and cells (eg, B-cells) can be isolated from these wherein the genome is as per the invention.

In one embodiment, the biased D gene segment(s) are selected from the IMGT database of variants or the 1000 Genomes database.

14. The vertebrate or cell of any one of clauses 8 to 13, wherein the VH gene segments are selected from the group consisting of VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01.

In an embodiment, the genome comprises a human immunoglobulin VH gene segment repertoire that is biased to VH1-69.

In an embodiment, the human immunoglobulin VH gene segment repertoire substantially consists of one or more human VH1-69 gene segments.

The gene segments are provided in one or more immunoglobulin loci. For example, the gene segment repertoire (D and/or VH) is provided in both IgH loci (ie, in a homozygous state).

15. The vertebrate or cell of any one of clauses 8 to 16, comprising an immunoglobulin heavy chain locus comprising two or more copies of a VH gene segment selected from said group.

Thus, at least one of said copies is closer to the constant region of the locus than the germline distance in a human from a human constant region. The aim is to provide by bias by providing more than one copy on the same locus. Also as at least one of the copies is closer (more proximal to) the constant region and J-C intron (which includes regulatory elements such as the Emu enhancer region), this may favour use of the gene segment, thus contributing to the desired bias.

Optionally, the genome is homozygous for the heavy chain locus.

Optionally the two or more copies of gene segments are identical (eg, all VH1-69\*01, using IMGT nomenclature). In another example, copies are variants of each other, eg, naturally-occurring human variants. Alternatively, synthetic variants may be used with or without a naturally-occurring variant.

In any embodiment of the invention, the vertebrate is naïve or immunised with a target antigen.

16. The vertebrate or cell of any clause, wherein the genome comprises a human JH gene segment repertoire consisting of one or more human JH6 gene segments.

This biases the JH repertoire for the production of long HCDR3, since this is the longest naturally-occurring human JH gene segment type and is commonly found in naturally-occurring human

antibodies having long HCDR3.

For example, the repertoire comprises two or more different JH6 variants. In an example, the repertoire comprises two or more JH6\*02 variants (IMGT nomenclature).

17. The vertebrate or cell of any preceding clause, wherein the genome comprises a human immunoglobulin JH gene segment repertoire that is biased to JH6, optionally JH6\*02.
18. The vertebrate or cell of clause 17, wherein the JH gene segment repertoire consists or substantially consists of three or more human JH6 gene segments.
19. The vertebrate or cell of any preceding clause, wherein the sequence of each of said human gene segments is a human germline gene segment sequence.
20. The vertebrate or cell of any preceding clause, wherein one, more or all of the selected gene segments are present in the genome as two or more copies, the copies being variants of each other.

Thus, one, more or all of the human V, D and JH gene segments of said genome is present in two or more variant versions, such as naturally-occurring human variants, eg, variants found in the 1000 Genomes database and/or IMGT database. In another example, one or more of the variants may be a synthetic variant.

21. The vertebrate or cell of any preceding clause, wherein said human gene segments are provided by homozygous immunoglobulin heavy chain loci.

In an example, no other (non-human) active heavy chain VH, D or JH gene segments are present in heavy chain loci of the genome. Additionally, in an example no active non-human light chain VL or JL gene segments are present in the genome.

This is useful for ensuring that endogenous (non-human) variable region expression is

inactivated. Thus, all heavy chains produced by the vertebrate or cell will have human variable regions, which is useful for producing drugs for administration to humans.

22. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin VH gene segment repertoire, one or more human D gene segments and one or more human JH gene segments, wherein the VH repertoire does not comprise one, more or all VH gene segments selected from the group VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.
  
23. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin D gene segment repertoire, one or more human VH gene segments and one or more human JH gene segments, wherein the D repertoire does not comprise one, more or all D gene segments selected from the group D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.

In instances it has been observed in that the art that certain human gene usage may dominate the immune response to infectious disease pathogen antigens or other antigens. While this may yield many specific antibodies, typically these may not be neutralising and thus the immune response is relatively ineffective. This may happen, for example, where the antigen is a decoy antigen expressed by the pathogen. The present embodiments of the invention where specific gene segments are omitted are useful for avoiding dominance of certain human gene segments, such as those omitted from the genome. In this way, the genome human gene segment repertoire is biased away from the dominance and this enables better use and sampling of the remaining human gene segment sequence space, thereby providing the chance of producing antibodies that may not be normally raised in a natural setting. Antigen specific antibodies can be selected from vertebrates and cells with such genomes. In some examples, this may yield neutralising antibodies.

It is advantageous to include a plurality of different human VH gene segments, making up the human VH gene segment repertoire. This provides for good diversities of rearranged human variable regions from which to select leads. It is possible, for example, to include an otherwise complete, functional repertoire of human VH gene segments. To this end, the human VH gene

segment repertoire comprises, in one example, a plurality of human VH gene segments, eg, at least 7, 10, 15, 20, 15, 30, 35, 40 or 45 different human VH gene segments. This can be achieved, for example, using BACs harbouring stretches of unrearranged human variable region DNA comprising VH gene segments - homologous recombination and/or sRMCE being used to insert several stretches of such DNA from serial BACs into an endogenous heavy chain locus upstream of the constant region thereof in the genome of a non-human vertebrate ES cell (eg, mouse or rat ES cell), followed by development of one or more progeny vertebrates from such cells (and optional breeding to homozygosity of the heavy chain locus). In one embodiment, human DNA is inserted that includes a first human VH (eg, VH1-69 and/or VH1-2) and flanking VH gene segments upstream and downstream of these. In a second ES cell genomic manipulation, the first VH is deleted from the genome, eg, using standard homologous recombination techniques as is known in the art. In this way, one or more VH gene segments usually upstream and/or downstream of the deleted gene segment(s) in a wild-type human germline genome are retained so that they can be available to contribute to the subsequent rearranged human V region repertoire that is used for selection of leads. In another example, the initial insertion of human DNA is made using stretches of DNA that already omit the first VH (eg, by deleting such stretches using recombineering of BACs in E coli, as is known in the art). Similar techniques can be used (with appropriate BACs) for the omission of human D and/or J gene segments.

Thus, in an embodiment, VH gene segments that normally are upstream and/or downstream of the omitted human VH gene segments (or omitted D or J as per other embodiments) in a wild-type germline human genome are included in the vertebrate or cell of the invention. For example, the human VH gene segment repertoire of the genome in the invention does not comprise VH1-69, but does comprise one, two, three or four human VH gene segments selected from VH2-10, VH3-72, VH3-73 and VH3-74. These are gene segments that are immediately upstream of VH1-69 in a wild-type human germline heavy chain locus (see IMGT). For example, additionally or alternatively the human VH gene segment repertoire of the genome in the invention does not comprise VH1-69, but does comprise one, two, three, four or more (or all of) human VH gene segments selected from VH3-66, VH3-64, VH4-61, VH4-59, VH1-58, VH3-53, VH3-49, VH3-48, VH1-46 and VH1-45. These are gene segments that are immediately downstream of VH1-69 in a wild-type human germline heavy chain locus (see IMGT). Additionally or alternatively, the human VH gene segment repertoire of the genome in the invention does not comprise VH1-69, but does comprise one, two, three, four or



more (or all of) human VH gene segments selected from VH2-5, 7-41, 4-4, 1-3, 1-2 and 6-1.

Additionally or alternatively, the human VH gene segment repertoire of the genome in the invention does not comprise VH1-69, but does comprise one, two, three, four, 5, 6, 7, 8, 9, 10 or more (or all of) human VH gene segments selected from VH2-5, 7-41, 4-4, 1-3, 1-2, 6-1, 3-7, 1-8, 3-9, 3-11 and 3-13. Additionally or alternatively, the human VH gene segment repertoire of the genome in the invention does not comprise VH1-69, but does comprise one, two, three, four, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 or more (or all of) human VH gene segments selected from VH2-5, 7-41, 4-4, 1-3, 1-2, 6-1, 3-7, 1-8, 3-9, 3-11, 3-13, 3-15, 1-18, 3-20, 3-21, 3-23, 1-24 and 2-26. Additionally or alternatively, the human VH gene segment repertoire of the genome in the invention does not comprise VH1-69, but does comprise VH6-1 (which is commonly used in human immune responses, VH6-1 being the most proximal to the constant region in a wild-type human germline heavy chain locus) and/or VH3-23 (which is commonly used in human immune responses). In embodiment (eg, for generating VH, heavy chains or antibodies for treating and/or preventing an infectious disease, eg, HIV infection, in a human), VH1-2 is omitted in the genome or locus. In this case one, two, three or all human VH gene segments immediately 5' and 3' of VH1-2 in a wild-type germline human IgH locus (eg, see IMGT) are included in the genome, such as comprised by the same IgH locus upstream of human D and JH gene segments and a constant region.

24. The vertebrate or cell of clause 22 or 23, wherein the genome comprises a human JH gene segment repertoire that does not comprise JH6.

#### JH Bias

25. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin JH gene segment repertoire that is biased to human JH6.

In an example, the repertoire is biased to human JH6\*02 (IMGT nomenclature).

So, the inventors made a choice of human JH6\*02 on the basis of

- (i) Containing YYG and YYGXDX motifs that is conserved across several vertebrate species;
- (ii) Provision of one less TAC codon than other human JH6 variant (an AID hotspot that risks stop codons) and a choice instead of a codon that preserves the YYG and YYGXDX motifs;
- (iii) Avoidance of a GGCA AID hotspot in the region of the HCDR3/FW4 junction; and

(iv) Common occurrence (and thus conservation and acceptability) in humans of the JH6\*02 variant.

26. The vertebrate or cell of clause 25, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising a plurality of human JH6 gene segments; optionally wherein the genome is homozygous for said locus.

In an example, the plurality comprises or consists of a plurality of JH6\*02 gene segments.

27. The vertebrate or cell of clause 26, wherein the heavy chain locus comprises (in 5' to 3' order) human VH, D and JH gene segments and said JH6 gene segments are spaced from the D gene segment(s) by no more than two other JH gene segments.

28. The vertebrate or cell of clause 25, 26 or 27, wherein are no other JH gene segments in the locus between said human JH6 gene segments.

29. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin JH gene segment repertoire that consists of one or more human JH6 gene segments.

In an example, all of the gene segments are JH6\*02 gene segments.

30. The vertebrate or cell of any one of clauses 25 to 29, wherein all of said gene segments are human germline gene segments.

31. The vertebrate or cell of any one of clauses 25 to 30, comprising different variant JH6 gene segments.

In an example, the variants are all naturally-occurring (eg, appearing in the IMGT or 1000 Genome databases). In an other example, one or more variant is synthetic.

32. The vertebrate or cell of any one of clauses 25 to 31, wherein said gene segments are provided by homozygous immunoglobulin heavy chain loci.

In one embodiment, the biased JH gene segment(s) are selected from the IMGT database of variants or the 1000 Genomes database.

33. A monoclonal or polyclonal antibody composition or a population of antibody-producing cells for producing such composition, wherein the composition or population is prepared by immunising at least one vertebrate according to any preceding clause with an antigen, wherein the antibody or antibodies have human heavy chain variable regions comprising non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences; wherein the composition comprises at least one antigen-specific antibody having a HCDR3 length of at least 20 amino acids (according to IMGT).

As will be readily apparent to the skilled person, AID and TdT mutations can be determined using bioinformatics analysis to find the closest matching human germline gene segment(s) that correspond to a given variable domain sequence, aligning the sequences and determining the differences. AID has known hotspots for mutation (eg, see *Annu. Rev. Biochem.* 2007. 76:1–22; Javier M. Di Noia and Michael S. Neuberger, “Molecular Mechanisms of Antibody Somatic Hypermutation” (in particular figure 1 and associated discussion on AID hotspots in mouse); and *Curr Opin Immunol.* 1995 Apr;7(2):248-54, “Somatic hypermutation”, Neuberger MS and Milstein C (in particular, discussion on hotspots in mouse), the disclosures of which are incorporated herein by reference). By carrying out the standard bioinformatics analysis, TdT mutations (eg, to provide junctional mutations and diversity) can be determined, as will be familiar to the skilled person.

Corresponding human germline V, D and J sequences can be according to the IMGT database or 1000 Genomes database, for example.

For example, the HCDR3 length is at least 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 amino acids.

For example, the HCDR3 length is from 20 to 23 or 24 to 30, eg, from 28 to 30 amino acids.

For example, the cells are B cells (eg, immortalised B cells) or hybridomas.

Optionally the antibodies of any aspect of the invention comprise human light chain variable regions. For example, the human light chain variable regions have non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences.

34. An isolated antibody that specifically binds an antigen, the antibody comprising human heavy chain variable regions and non-human constant regions, wherein the variable regions are derived from the recombination in a non-human vertebrate of (i) a human VH gene segment selected from the group recited in clause 8 with (ii) a human D gene segment selected from the group recited in clause 1 and with a human JH gene segment (optionally JH6); wherein the antibody has a HCDR3 length of at least 20 amino acids (according to IMGT); and non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences.

In examples, the VH is selected from the group VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01 and/or

the D is selected from the group

D2-2\*02, D3-9\*01, D3-10\*01 and D3-22\*01, or

D2-2\*02, D3-9\*01 and D3-10\*01, or

D3-9\*01 and D3-10\*01, or

D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or

D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01, or

D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01,

D4-17\*01, D6-13\*01 and D6-19\*01, or  
D1-26, D2-2, D3-10 and D6-19, or  
D2-2, D3-9 and D3-10.

35. The antibody of clause 34, wherein the antibody is obtained or obtainable from a vertebrate according to any one of clauses 1 to 32.

In an embodiment, the antibody is obtained from said vertebrate, or is a copy of such an antibody.

36. A method of isolating an antibody that binds a predetermined antigen, the method comprising
- (a) providing a vertebrate (optionally a mouse or rat) according to any one of clauses 1 to 32;
  - (b) immunising said vertebrate with said antigen;
  - (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes expressing antibodies that bind to the antigen;
  - (d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by producing hybridomas therefrom; and
  - (e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes.

37. The method of clause 36, wherein in step (e) wherein the antibody has a HCDR3 length of at least 20 amino acids (according to IMGT).

The length can be 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 amino acids (according to IMGT), eg, from 20 to 23 amino acids (as produced in the examples).

38. The method of clause 36 or 37, comprising the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally exchanging the heavy chain constant region nucleotide sequence of the antibody with a nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host.

39. The method of clause 36, 37 or 38, further comprising making a copy, mutant or derivative (eg, humanised version) of the antibody produced by the method.

Humanisation can entail making the constant regions human.

40. The antibody composition, cell population, antibody or method of any one of clauses 33 to 39, wherein the antigen is an antigen of an infectious disease pathogen; optionally wherein the pathogen is a virus or bacterium.
41. The antibody composition, cell population, antibody or method of clause 40, wherein pathogen is selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus.
42. The antibody composition, cell population, antibody or method of any one of clauses 33 to 41, wherein the antigen is a HIV gp120 antigen or a HIV gp41 antigen.
43. The antibody composition, cell population, antibody or method of any one of clauses 33 to 40, wherein the antigen comprises an active site or cleft, wherein the antibody having a HCDR3 length of at least 20 amino acids specifically binds to the active site or cleft of the antigen.
44. A pharmaceutical composition comprising an antibody or antibody composition according to any one of clauses 33 to 35 and 40 to 43, or an antibody produced by the method of any one of clauses 36 to 38, for treating and/or preventing an infectious disease in a human (eg, wherein the infectious disease is caused by a pathogen selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus).
45. A repertoire of antibody heavy chains (eg, provided by antibodies) comprising one or more heavy chains whose variable domain HCDR3 has a length of at least 20 amino acids (according to IMGT) and derived from the recombination of a human VH, D and JH, wherein

the VH is selected from the group

VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01 and

the D is selected from the group

D2-2\*02, D3-9\*01, D3-10\*01 and D3-22\*01, or

D2-2\*02, D3-9\*01 and D3-10\*01, or

D3-9\*01 and D3-10\*01, or

D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or

D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01, or

D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01,

D4-17\*01, D6-13\*01 and D6-19\*01, or

D1-26, D2-2, D3-10 and D6-19, or

D2-2, D3-9 and D3-10;

and optionally the JH is JH6 (eg, JH6\*02);

Wherein

(a) the heavy chain variable domain has been produced *in vivo* in a non-human vertebrate (eg, a mouse or a rat); and/or

(b) the heavy chain variable domain comprises non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences.

In an example, the heavy chain (or all heavy chains in the repertoire) comprise non-human vertebrate constant regions (eg, mouse or rat constant regions). For example, the constant regions are gamma-type constant regions (eg, gamma-1, gamma-2 or gamma-4 type).

In an example, the repertoire is a naïve repertoire. This is shown in the examples section herein.

In an example, the repertoire is an immunised repertoire. This is shown in the examples section herein.

In an example, the repertoire is an antigen-specific repertoire (eg, provided by a plurality of hybridomas). This is shown in the examples section herein.

The repertoire can be provided by B cells (eg, immortalised B cells).

The repertoire can be provided by hybridomas.

In an example, the vectors are harboured by host cells (eg, CHO or HEK293 cells or yeast cells).

The HCDR3 length can be 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 amino acids (according to IMGT), eg, from 20 to 23 amino acids (as produced in the examples).

In an example, in (a) the vertebrate is a vertebrate according to the invention.

46. A nucleic acid collection encoding the heavy chain repertoire of clause 45.

In an example, the nucleic acids are provided in respective vectors (eg, expression vectors, eg, *E coli* or CHO or HEK293 vectors).

47. A method of obtaining an antigen-specific heavy chain (eg, provided by an antibody), the method comprising exposing the repertoire of clause 45 to a predetermined antigen and selecting one or more heavy chains that specifically bind to the antigen, wherein one or more heavy chains is isolated that has a HCDR3 length of at least 20 amino acids.

Optionally, when the heavy chain has a non-human constant region, this is swapped for a human constant region, as is conventional in the art. Thus, the invention provides a human antibody heavy chain so produced (eg, provided in combination with a human light chain to produce a human antibody which is useful for human therapeutic and/or prophylactic use, eg, to treat and/or prevent an infectious disease in a human patient).

In an example of the vertebrate or cell of any aspect of the invention, the genome comprises an immunoglobulin light chain locus comprising one or more human V gene segments and one or more



human J gene segments upstream of a constant region (eg, a human or a mouse lambda or kappa constant region).

For rearrangement and expression of heavy chains, the locus comprises control elements, such as an  $E\mu$  and  $S\mu$  between the J gene segment(s) and the constant region as is known by the skilled person. In one example, a mouse  $E\mu$  and  $S\mu$  is included in the heavy chain locus between the JH repertoire and the constant region (ie, in 5' to 3' order the locus comprises the JH gene segment(s),  $E\mu$  and  $S\mu$  and constant region). In an example, the  $E\mu$  and  $S\mu$  are  $E\mu$  and  $S\mu$  of a mouse 129-derived genome (eg, a 129Sv-derived genome, eg, 129Sv/EV (such as 129S7Sv/Ev (such as from AB2.1 or AB2.2 cells obtainable from Baylor College of Medicine, Texas, USA) or 129S6Sv/Ev)); in another example, the  $E\mu$  and  $S\mu$  are  $E\mu$  and  $S\mu$  of a mouse C57BL/6 -derived genome. In this respect, the locus can be constructed in the IgH locus of the genome of a cell selected from AB2.1, AB2.2, VGF1, CJ7 and FH14. VGF1 cells were established and described in Auerbach W, Dunmore JH, Fairchild-Huntress V, *et al*; Establishment and chimera analysis of 129/SvEv- and C57BL/6-derived mouse embryonic stem cell lines. *Biotechniques* 2000; 29:1024–8, 30, 32, incorporated herein by reference.

Additionally or alternatively, the constant region (or at least a  $C\mu$ ; or  $C\mu$  and gamma constant regions thereof) is a constant region (or  $C\mu$ ; or  $C\mu$  and gamma constant regions thereof) is of a genome described in the paragraph immediately above.

A suitable source of human DNA sequences or gene segments will be readily apparent to the skilled person. For example, it is possible to collect a DNA sample from a consenting human donor (eg, a cheek swab sample as per the Example herein) from which can be obtained suitable DNA sequences for use in constructing a locus of the invention. Other sources of human DNA are commercially available, as will be known to the skilled person. Alternatively, the skilled person is able to construct gene segment sequence by referring to one or more databases of human Ig gene segment sequences disclosed herein.

In an example, the genome comprises all or some of the following human VH gene segments

IGHV6-1

IGHV3-7

IGHV1-8

IGHV3-9

IGHV3-11

IGHV3-13

IGHV1-18

IGHV3-30

IGHV4-31

IGHV4-39

IGHV4-59

Optionally also (i) and/or (ii)

(i)

IGHV1-2

IGHV2-5 and

IGHV3-21

(ii)

IGHV1-2

IGHV2-5

IGHV3-21

IGHV1-24

For example, the genome comprises all or some of the following human VH gene segment variants

IGHV6-1\*01

IGHV3-7\*01

IGHV1-8\*01

IGHV3-9\*01

IGHV3-11\*01

IGHV3-13\*01

IGHV1-18\*01

IGHV3-30\*18

IGHV4-31\*03

IGHV4-39\*01 and

IGHV4-59\*01;

Optionally also (iii) or (iv)

(ii)

IGHV1-2\*04

IGHV2-5\*10 and

IGHV3-21\*03

(iv)

IGHV1-2\*02

IGHV2-5\*01

IGHV3-21\*01 and

IGHV1-24\*01

For example, the genome comprises all or some of the following human JH gene segment variants

IGHJ2\*01

IGHJ3\*02

IGHJ4\*02

IGHJ5\*02 and

IGHJ6\*02

For example, the genome comprises all or some of the following human D gene segments

IGHD1-1

IGHD2-2

IGHD3-9

IGHD3-10

IGHD5-12

IGHD6-13

IGHD1-14

IGHD2-15

IGHD3-16

IGHD4-17

IGHD6-19

IGHD2-21

IGHD5-24

IGHD1-26 and

IGHD7-27

and optionally also (v) or (vi)

(v)

IGHD3-3

(vi)

IGHD3-3

IGHD4-4

IGHD5-5

IGHD6-6

IGHD1-7

IGHD2-8 and

IGHD2-8

It will be understood that particular embodiments described herein are shown by way of illustration and not as limitations of the invention. The principal features of this invention can be employed in various embodiments without departing from the scope of the invention. Those skilled in the art will recognize, or be able to ascertain using no more than routine study, numerous equivalents to the specific procedures described herein. Such equivalents are considered to be within the scope of this invention and are covered by the claims. All publications and patent applications mentioned in the specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference. The use of the word "a" or "an" when used in conjunction with the term "comprising" in the claims and/or the specification may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one." The use of the term "or" in the claims is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and "and/or." Throughout this application, the term "about" is used to indicate that a value includes the inherent variation of error for the device, the method being employed to determine the value, or the variation that exists among the study subjects.

As used in this specification and claim(s), the words "comprising" (and any form of comprising, such as "comprise" and "comprises"), "having" (and any form of having, such as "have" and "has"), "including" (and any form of including, such as "includes" and "include") or "containing" (and any form of containing, such as "contains" and "contain") are inclusive or open-ended and do not exclude additional, unrecited elements or method steps

The term "or combinations thereof" as used herein refers to all permutations and combinations of the listed items preceding the term. For example, "A, B, C, or combinations thereof" is intended to include at least one of: A, B, C, AB, AC, BC, or ABC, and if order is important in a particular context, also BA, CA, CB, CBA, BCA, ACB, BAC, or CAB. Continuing with this example, expressly included are combinations that contain repeats of one or more item or term, such as BB, AAA, MB, BBC,

AAABCCCC, CBBAAA, CABABB, and so forth. The skilled artisan will understand that typically there is no limit on the number of items or terms in any combination, unless otherwise apparent from the context.

Any part of this disclosure may be read in combination with any other part of the disclosure, unless otherwise apparent from the context.

All of the compositions and/or methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and/or methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

The present invention is described in more detail in the following non limiting Examples (Examples 1-3 being prophetic). Example 4 is a worked example.

## **EXAMPLES**

### **EXAMPLE 1**

**“Recombineered BAC Vectors to add Polymorphic V-regions to the Mouse Genome”**

Figure 1 through 3 depict recombineering methods (see references above) that can be used to introduce polymorphic V-gene regions into genomic DNA. In one embodiment, a genomic fragment

from the human heavy chain region is inserted into a bacterial artificial chromosome (BAC) vector by standard techniques. Preferably, such a BAC, which can range in size from 20-kb to 200-kb or more, can be isolated from libraries of BACs by standard techniques including sequence searches of commercially available libraries or by hybridization to bacterial colonies containing BACs to identify those with a BAC of interest.

A BAC is chosen that has several VH gene segments; in Figure 1, these are generically identified as VH[a] through VH[z] for example. One skilled in the art will readily identify appropriate genomic fragments, for example, an approximately 120-kb fragment from human VH5-78 through VH1-68 which includes 5 endogenous active VH gene segments and 7 VH pseudogenes. Using recombinering techniques, the endogenous VH gene segments can be replaced by polymorphic VH or VL gene segments. In this example, two steps are required. The first step replaces the V-region coding exon of an endogenous VH gene segment with a positive-negative selection operon, in this example, an operon encoding an ampicillin resistance gene (*Amp*) and a streptomycin-sensitizing ribosomal protein (*rpsL*). Certain strains of bacteria can be selected for the absence of the *rpsL* gene by resistance to streptomycin. Short stretches of DNA homologous to sequences flanking the endogenous VH gene exon are placed 5' and 3' of the *rpsL*-*Amp* operon. In the presence of appropriate recombination factors per standard recombinering techniques (see references above) recombination between the operon fragment and the BAC will result in replacement of the endogenous VH gene exon with the operon (Figure 1a) which are selected by resistance to ampicillin. The second step uses the same homologous sequences in order to replace the inserted operon with a desired polymorphic VH gene segment. In this example, a human VH1-69 gene is inserted (Figure 1b and 1c). In particular the \*02 allele of VH1-69 is used [ref IMGT and Figure 5]. Successful integrations of the polymorphic VH gene segment are selected in bacteria that become resistant to streptomycin due to the loss of the operon, specifically the *rpsL* portion.

In this example, the two step process as described can be repeated for each of the endogenous VH gene segments or for as many endogenous gene segments that one wishes to replace with polymorphic V gene segments (Figure 1d).



As is apparent, any polymorphic V gene segment can be inserted in this manner and any endogenous V gene segment can act as a target, including pseudogenes. V gene segments in each of the heavy chain and two light chain loci can be replaced using this technique with appropriate genomic fragments available as BAC inserts.

Figure 2 depicts another method for creating a genomic fragment encoding polymorphic V gene segments. In this example, polymorphic V gene segments are inserted into a region of genomic DNA devoid of other genes, control elements or other functions. Such 'desert' regions can be selected based on sequence analysis and corresponding DNA fragments cloned into BACs or identified in existing BAC libraries. Starting with such a genomic fragment, recombineering techniques can be used to insert polymorphic V gene segments at intervals of, for example, 10-kb. In this example, a 150-kb genomic fragment might accommodate insertion of up to 15 polymorphic V gene segments. Insertion of the segments is a two-step process. The first recombineering step inserts the *rpsL-Amp* operon at a specific site. Sequences homologous to a specific site are used to flank the operon. These are used by the recombineering system to insert the element specifically into the BAC genomic fragment and positive events are selected by resistance to ampicillin (Figure 2a). The second step replaces the operon in the genomic fragment with a polymorphic V gene segment by a similar recombineering step using the same sequence homology (Figure 2b). In this example, both exons and promoter element of a polymorphic VH gene segment are inserted, resulting in replacement of the *rpsL-Amp* operon and therefore resistance to streptomycin (Figure 2c).

The two step technique for inserting polymorphic V gene segments into a specific site on the genomic fragment can be repeated multiple times resulting in a BAC genomic fragment with several polymorphic gene segments, including their promoter elements. It is apparent that the examples shown in Figures 1 and 2 can be combined wherein the technique for insertion can be used to add *extra* polymorphic V gene segments to a BAC genomic fragment as depicted in Figure 1. One might choose to add these extra segments to an IG genomic fragment since such a fragment would be more amenable to proper IG gene expression once inserted into a non-human mammal's genome. It is known that a genomic fragment can have elements such as enhancers or elements that contribute to certain chromatin conformations, both important in wild-type gene expression.

Figure 3 depicts an additional method to create genomic fragments with polymorphic V gene segments. This method depends upon the efficiency with which short (around 50 to 150 bases, preferably 100 bases) single stranded DNA fragments recombine with a homologous sequence using recombineering (Nat Rev Genet. 2001 Oct;2(10):769-79; Recombineering: a powerful new tool for mouse functional genomics; Copeland NG, Jenkins NA, Court DL). The recombinases used in recombineering preferentially bind and use such short single-stranded fragments of DNA as a substrate for initiating homologous recombination. The efficiency can be as high as 10<sup>-2</sup>, that is, a positive event can be found in approximately 100 randomly picked (not selected) clones resulting from recombineering. A positive event in this example occurring when one or more single nucleotide changes introduced into the single-stranded fragment get transferred to the BAC insert containing V gene segments and surrounding genomic DNA, said nucleotide change or changes occurring at a homologous sequence on the BAC.

Polymorphic V gene segments can differ from endogenous V gene segments by only 1 or 2, or up to 10 or 15 nucleotide changes, for example. An example of such nucleotide polymorphisms are depicted in Figure 5. Short single stranded regions that encompass the polymorphic nucleotide changes can be chemically synthesized using standard techniques. The resulting single stranded DNA fragments are introduced into bacteria and *via* recombineering techniques approximately 1 in 100 BAC fragments will have incorporated the polymorphic nucleotides *via* homologous incorporation of the single stranded fragment (Figure 3a). BACs with the desired nucleotide change can be identified by screening for example several hundred individual clones by polymerase chain reaction (PCR) amplification and sequencing, both by standard techniques. In the example, two nucleotide changes will convert a VH1-69\*01 gene segment into a VH1-69\*02 gene segment (Figure 3b).

It is clear that this process can be repeated for multiple endogenous V gene segments contained on a single BAC genomic fragment. In addition, the techniques depicted in Figure 2 can be used to add additional polymorphic V gene segments by insertion into regions between existing V gene segments. As would be evident to one skilled in the art, a combination of these techniques can be used to create numerous variations of both polymorphic and endogenous human V gene segments. And it would be evident that several different genomic fragments with engineered polymorphic V gene segments and endogenous human V gene segments can be combined to create even more variations.

**EXAMPLE 2****“Adding Polymorphic V-regions to the Genome using SRMCE of Modified BACs”**

Modified BACs with polymorphic V gene segments created using the methods described in Example 1 can be used to alter the genome of non-human mammals. These alterations can result in an intact IG locus in which normal immunoglobulin region recombination results in VDJ or VJ combinations which includes the human V gene segments. An example of how such an animal can be created is by altering the genome of, for example, mouse embryonic stem (ES) cells using the strategy outlined in Figure 4.

One technique to integrate modified BACs with polymorphic V gene segments into a genome is sequential recombinase mediated cassette exchange (SRMCE). The technique is described in WO2011004192 (Genome Research Limited), which is incorporated here in its entirety by reference.

SRMCE provides for a locus modified with a ‘landing pad’ inserted at a specific location. This insertion can either be *de novo* via homologous recombination or as a consequence of a previous BAC insertion. In this example, the landing pad is inserted in the mouse IGH locus between the most 3’ J gene segment and the C $\mu$  gene segment and a previous BAC insertion *via* SRMCE techniques have resulted in the addition of 5 human V gene segments and 2 V region pseudogenes. The landing pad has elements as shown in Figure 4 that will allow the selection of correct insertion of a second targeting BAC fragment. The specificity of this insertion is provided by cre recombinase-mediated exchange between permissive *lox* sites. A *lox* site is permissive for recombination only with a compatible *lox* site. In this example, the *loxP* site will only recombine with *loxP* and *lox2272* will only recombine with *lox2272*. This provides directionality to the insertion of the BAC fragment as depicted in Figure 4b and 4c.

ES cell clones with correct insertions are selected from a pool of clones without insertions or with non-productive insertions by resistance to puromycin. Resistance to puromycin results from the

juxtaposition of an active promoter element, PGK, with the *puroTK* coding region. Correct insertions are verified by standard techniques including PCR of junctions, PCR of internal elements, Southern blotting, comparative genomic hybridization (CGH), sequencing and *etc.* In the example, correct *lox2272-lox2272* and *loxP-loxP* recombination also results in two intact sets of *piggyBac* elements that did not exist prior to insertion. An intact *piggyBac* element is comprised of a set of inverted repeats which are depicted in the figure by "PB5'" and "PB3'". An appropriated oriented set of *piggyBac* elements are the substrate of *piggyBac* transposase which can catalyse recombination between the elements, resulting in deletion of intervening sequences as well as both elements. The DNA remaining after a *piggyBac* transposition is left intact and is lacking any remnant of the *piggyBac* element. In the example, ES cell clones with successful *piggyBac* transposition are selected by loss of the active *puroTK* element which renders the cells resistant to the drug FIAU (Figure 4c and 4d).

The final product of the SRMCE method in this example is a IGH locus with several polymorphic V gene segments inserted along with a set of endogenous unmodified VH gene segments between sequences of the mouse genome on the 5' side and the mouse IGH constant region gene segments on the 3' side. The polymorphic V gene segments are positioned such that they can participate in the recombination events associated with B cell maturation yielding VDJ gene segments. These gene segments can then be transcribed and spliced to the mouse constant region. Translation of these transcripts will result in the production of an antibody heavy chain encoded by the polymorphic V gene segment, a human DH gene segment, a human JH gene segment and a mouse constant heavy chain gene segment.

As is well known to those skilled in the art, an ES cell clone can be used to create a line of genetically modified mice *via* injection of said cells into a mouse blastocyst embryo, transferring the injected embryo to a suitable recipient and breeding the chimeric offspring that result. The modified gene locus can be propagated through breeding and made either heterozygous or homozygous depending on the genetic cross.

It is evident from the structure of the IGH locus provided in this example and by knowledge of the mechanisms involved in B cell receptor (BCR) and antibody gene rearrangements that a large set of

different combinations of polymorphic V gene segments with various DH and JH gene segments will result and these can contribute to a large repertoire of functional antibody genes in a population of B cells in genetically modified animals. In this example, several different human VH1-69 polymorphs are incorporated to provide superhuman VH diversity. This particular VH gene segment is known to be prevalent in antibodies that bind infectious disease pathogens (such as influenza virus) and therefore the antibody repertoire of a mouse with the genetic modification of this example would be expected to produce antibodies with a bias in favour of those that bind infectious disease pathogens. The repertoire, in other words, would have a larger subset of antibodies with superior affinities for pathogen antigens. Examples of such pathogens include influenza virus, hepatitis C virus (HCV) and human immunodeficiency virus-1 (HIV-1) (see also table above).

### EXAMPLE 3

#### “Alignment of 13 VH1-69 Alleles”

Building a more diverse antibody repertoire by incorporating additional V gene segment polymorphs requires availability of polymorphic alleles of V gene segments. One source of such alleles include sequence databases. In this example, 13 distinct alleles of the VH1-69 gene segment are provided. These allele sequences and comparisons are drawn from the “IMmunoGeneTics” IMGT Information System ([www.imgt.com](http://www.imgt.com)) database. Figure 5 is a diagram of the alignment of alleles \*02 through \*13 with the \*01 allele. The VH1-69\*01 nucleotide and amino acid sequence is provided at the top of the figure. Where the remaining alleles are identical to the \*01 allele sequence a dash is inserted below the sequence. Nucleotide differences are noted alongside the appropriate allele and if the sequence change results in a protein coding change, the amino acid change is indicated above the triplet.

Figure 5 depicts between 1 and 4 amino acid changes for each allele in comparison to the \*01 allele. All of the amino acid changes occur in the part of the heavy chain protein encoding the complementarity determining regions (CDRs). These regions are responsible for antigen specificity and the affinity of the antibody for the antigen. It is evident that providing additional polymorphic CDRs in a repertoire of antibodies will increase the likelihood of there being an antibody with superior binding characteristics for various antigens. In several reports, it has been observed that the VH1-69-encoded variable region of the heavy chain is often found in antibodies that bind influenza

virus, HCV and HIV-1 antigens (see table above). Therefore incorporating the polymorphic V gene segments of this example into a transgenic animal model using the methods of Examples 1 and 2 would likely result in an antibody repertoire in said transgenic animal with more antibodies that bind to antigens associated with these and other pathogens. And as is known in the art, a larger repertoire increases the probability of finding monoclonal antibodies using, for example, hybridoma technology, that bind with high affinity and specificity to a desired antigen.

This disclosure therefore describes in these examples a transgenic mouse model which can be immunized with pathogen or other antigens. Plasma B cells from such an immunized mouse can be used to make a hybridoma library that can be screened for production of antibodies that bind the pathogen antigens. This library will be superior to libraries from traditional transgenic mice for finding such antibodies given the addition of polymorphic VH1-69 gene segments to the IGH locus in said transgenic mouse.

These examples are not limiting to the human polymorphic V gene segments that can be chosen or to the methods used to introduce them into an animal model. The method can be used to construct a transgenic locus with immunoglobulin D and/or J segments. The V, D, J segments can be from a plurality of human sources (optionally more than one human ethnic population).

#### **Example 4:**

“Transgenic Mice, B-cells, Hybridomas, Antibodies & Heavy Chains Based on Human JH6\*02”

A functional human gene segment repertoire (from V<sub>H</sub>2-26 to J<sub>H</sub>6, see the IMGT database for the structure of the human IgH locus;

<http://www.imgt.org/IMGTrepertoire/index.php?section=LocusGenes&repertoire=locus&species=human&group=IGK> ) was sectored by the inventors to produce two different transgenic heavy chain

alleles (denoted S2F and S3F) and corresponding mice. The transgenic alleles were expressed in the mice and the heavy chain repertoires were assessed at the RNA transcript level. Deep sequence analysis was carried out using Bioinformatics methods to assess V, D and JH gene usage, including in variable domain sequences having a HCDR3 length of at least 20 amino acids. Endogenous, mouse variable region gene segments were inactivated by inversion (as per the method described in WO/2011/001192 this disclosure being incorporated herein by reference).

### Sequencing of Human Donor DNA Samples: Identification of Conserved JH6\*02 Variant

DNA samples from 9 anonymised consenting human donors were obtained by taking cheek swabs.

The samples were processed and the DNA Samples were extracted follow the protocol of QIAamp DNA Mini Kit (Cat.No.51304, Qiagen).

PCR reactions were set up to amplify the JH6 region and PCR products were sequenced (PCR Oligos sequence: Fwd. 5'-AGGCCAGCAGAGGGTTCCATG-3' (SEQ ID NO: 444), Rev. 5'-GGCTCCCAGATCCTCAAGGCAC-3' (SEQ ID NO: 445)).

Sequence analysis was carried out by comparing to the JH6 reference sequence from IMGT annotated database (<http://www.imgt.org/>), and this identified that all 9 donor genomes contained the human JH6\*02 variant, with this variant being in the homozygous state in 7 out of the 9 donors. The inventors also consulted the genomic sequences publicly available for Jim Watson and Craig Venter at Ensembl human genome database [<http://www.ensembl.org/>]. These too contained the human JH6\*02 variant. This confirmed to the inventors that human JH6\*02 is a common, conserved variant in humans, and thus a good candidate for construction of a transgenic IgH locus as per the invention

### Identification of Suitable Human DNA Sequence BACs

A series of human bacterial artificial chromosome (BAC) clones were identified from Ensemble (<http://www.ensembl.org/index.html>) or UCSC (<http://genome.ucsc.edu/>) human database searches based on gene name (IGH) or location (chromosome 14: 106026574-107346185). Seven human RP11 BAC clones were selected, RP11-1065N8 BAC carrying human JH6\*02. In total, the following BACs were identified as sources of human IgH locus DNA: RP11-1065N8, RP11-659B19, RP11-14117, RP-112H5, RP11-101G24, RP11-12F16 and RP11-47P23.

With a similar approach, different BAC clones (eg, different RP11 clone IDs or different sources from RP11) or genetically engineered BACs can be selected for insertion into the mouse IGH locus to provide different sets of human repertoires in the transgenic mouse.

#### Construction of Transgenic IgH Loci

Insertion of human heavy gene segments from a 1st IGH BAC (RP11-1065N8) into the IGH locus of mouse AB2.1 ES cells (Baylor College of Medicine) was performed to create a heavy chain allele denoted the S1 allele. The inserted human sequence corresponds to the sequence of human chromosome 14 from position 106494908 to position 106328951 and comprises functional heavy gene segments V<sub>H</sub>2-5, V<sub>H</sub>7-4-1, V<sub>H</sub>4-4, V<sub>H</sub>1-3, V<sub>H</sub>1-2, V<sub>H</sub>6-1, D1-1, D2-2, D3-9, D3-10, D4-11, D5-12, D6-13, D1-14, D2-15, D3-16, D4-17, D5-18, D6-19, D1-20, D2-21, D3-22, D4-23, D5-24, D6-25, D1-26, D7-27, J<sub>H</sub>1, J<sub>H</sub>2, J<sub>H</sub>3, J<sub>H</sub>4, J<sub>H</sub>5 and J<sub>H</sub>6 (in 5' to 3' order), wherein the J<sub>H</sub>6 was chosen to be the human J<sub>H</sub>6\*02 variant. The insertion was made between positions 114666435 and 114666436 on mouse chromosome 12, which is upstream of the mouse C<sub>μ</sub> region. The mouse V<sub>H</sub>, D and J<sub>H</sub> gene segments were retained in the locus, immediately upstream of (5' of) the inserted human heavy chain DNA.

A second allele, S2 was constructed in which more human functional V<sub>H</sub> gene segments were inserted upstream (5') of the 5'-most V<sub>H</sub> inserted in the S1 allele by the sequential insertion of human DNA from a second BAC (BAC2). The inserted human sequence from BAC2 corresponds to the sequence of human chromosome 14 from position 106601551 to position 106494909 and comprises functional heavy chain gene segments V<sub>H</sub>3-13, V<sub>H</sub>3-11, V<sub>H</sub>3-9, V<sub>H</sub>1-8, V<sub>H</sub>3-7. The mouse V<sub>H</sub>, D and J<sub>H</sub> gene segments were retained in the locus, immediately upstream of (5' of) the inserted human heavy chain DNA. In a subsequent step, these were inverted to inactivate them, thereby producing S2F mice in which only the human heavy chain variable region gene segments are active.

A third allele, S3 was constructed in which more human functional V<sub>H</sub> gene segments were inserted upstream (5') of the 5'-most V<sub>H</sub> inserted in the S2 allele by the sequential insertion of human DNA from a third BAC (BAC3). The inserted sequence corresponds to the sequence of human chromosome 14 from position 106759988 to position 106609301, and comprises functional heavy chain gene segments, V<sub>H</sub>2-26, V<sub>H</sub>1-24, V<sub>H</sub>3-23, V<sub>H</sub>3-21, V<sub>H</sub>3-20, V<sub>H</sub>1-18, and V<sub>H</sub>3-15. The mouse V<sub>H</sub>, D and J<sub>H</sub> gene segments were retained in the locus, immediately upstream of (5' of) the inserted



human heavy chain DNA. In a subsequent step, these were inverted to inactivate them, thereby producing S3F mice in which only the human heavy chain variable region gene segments are active.

Mice bearing either the S2F or S3F insertion into an endogenous heavy chain locus were generated from the ES cells using standard procedures. The other endogenous heavy chain locus was inactivated in the mice by insertion of an inactivating sequence comprising neo<sup>R</sup> into the mouse J<sub>H</sub>-C<sub>μ</sub> intron (to produce the "HA" allele).

Specifically, the following alleles were included:-

VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01

D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01

#### Immunisation procedure

Transgenic mice of the S2F or S3F genotype were primed with 20-40ug recombinant proteins obtained commercially or produced in house with Antigen 1 (OVA (Sigma A7641); Antigen 2 (a human infectious disease pathogen antigen) and Antigen 3 (a human antigen) via the ip route in complete Freund's adjuvant (Sigma F 5881) and 10ug/animal CpG (CpG oligo; Invivogen, San Diego, California, USA) and then boosted twice in about two weekly intervals with about half the amount of antigen in incomplete Freund's adjuvant (Sigma F 5506) and 10ug/animal CpG. Final boosts were administered two weeks later iv without any adjuvant and contained 5-10 ug protein in PBS.

#### Hybridoma fusion procedure

Spleens were taken 3 days after the final boost and spleenocytes were treated with CpG (25 μm final concentration) for and left until the following day. Cells were then fused with SP0/2 Ag14 myeloma cells (HPA Cultures Cat No 85072401) using a BTX ECM2001 electrofusion instrument. Fused cells were left to recover for 20 minutes then seeded in a T75 flask until next morning. Then the cells were spun down and plated out by dilution series on 96-well culture plates and left for about 10 days before screening. Media was changed 1-3 times during this period.

### Screening

Culture supernatants of the hybridoma wells above were screened using homogenous time resolved fluorescence assay (htfrf) using Europium cryptate labelled anti-mouse IgG (Cisbio anti-mouse Ig Europium Cryptate) and a biotin tagged target antigen with a commercially available streptavidin conjugated donor (Cisbio; streptavidin conjugated D2) or by IgG-specific 384 well ELISA. Positive wells identified by htfrf were scaled to 24-well plates or immediately counterscreened using an IgG-specific detection ELISA method. Positives identified by primary ELISA screen were immediately expanded to 24-well plates. Once cultures were expanded to 24-well stage and reached confluency, supernatants were re-tested using htfrf or IgG-specific ELISA to confirm binding to target antigen. Supernatant of such confirmed cultures were then also analysed by surface plasmon resonance using a BioRad ProteOn XPR36 instrument. For this, antibody expressed in the hybridoma cultures was captured on a biosensor GLM chip (BioRad 176-512) which had an anti-mouse IgG (GE Healthcare BR-1008-38) covalently coupled the biosensor chip surface. The antigen was then used as the analyte and passed over the captured hybridoma antibody surface. For Antigen 2 and Antigen 3, concentrations of 256nM, 64nM, 16nM, 4nM and 1nM were typically used, for Antigen 1, concentrations of 1028nM, 256nM, 64nM, 16nM and 4nM were typically used, binding curves were double referenced using a 0nM injection (i.e. buffer alone). Kinetics and overall affinities were determined using the 1:1 model inherent to the BioRad ProteOn XPR36 analysis software.

Any clones with confirmed binding activity were used for preparing total RNA and followed by PCR to recover the heavy chain variable region sequences. Standard 5'-RACE was carried out to analyse RNA transcripts from the transgenic heavy chain loci in the S2F and S3F mice. Additionally, deep sequence analysis of almost 2000 sequences produced by the mice was carried out.

### Bioinformatics Analysis

Sequences for analysis were obtained from two different methods:

- The first is from RNA extracted from the spleen: first cDNA strand was synthesized using an oligo based on the Cmu region of the mouse IGH locus as a PCR template. PCR was performed

using this oligo with an oligo dT-anchor primer. Then PCR product was cloned into pDrive vector (Qiagen) and then sequenced.

- The second is from hybridomas generated through electro-fusion: total RNA was extracted from hybridoma lines of interest using standard Trizol methods and frozen at  $-80^{\circ}\text{C}$  for long term storage. cDNA was generated from 100ng total RNA using standard Superscript III reverse transcriptase and a gene-specific reverse primer binding to all mouse IgG isotypes for heavy chain and a mouse kappa constant region primer for the light chain amplification. 2-3 ul of cDNA were then used as template in a PCR reaction using Pfu DNA polymerase and a panel of degenerate forward primers annealing to the leader sequence of the human immunoglobulin variable domain as well as one mouse pan-IgG reverse primer. PCR products were run out of a 1% agarose gel and bands of approximately 350-450 basepairs extracted and purified. DNA was then sequenced.

The sequences from the first method can either be from IgM from Naïve mice or IgG from immunised mice. The samples from the second method are all from IgG from immunised mice, and specific to the immunising antigen. Almost 2000 sequences were analysed.

The sequences were obtained as a pair of forward and reverse reads. These were first trimmed to remove low-quality base calls from the ends of the reads (trimmed from both ends until a 19 nucleotide window had an average quality score of 25 or more). The reads were combined together by taking the reverse complement of the reverse read, and aligning it against the forward read. The alignment scoring was 5 for a match, -4 for a mismatch, a gap open penalty of 10 and a gap extension penalty of 1. A consensus sequence was then produced by stepping through the alignment and comparing bases. When there was a disagreement the base with the highest quality value from sequencing was used.

The BLAST+ (Basic Local Alignment Search Tool) (Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., & Madden T.L. (2008) "BLAST+: architecture and applications." *BMC Bioinformatics* 10:421 <http://www.ncbi.nlm.nih.gov/pubmed/20003500>) program 'blastn' was then

used to find the germline J and V segments used in each sequence. A wordsize of 30 was used for V matching, and 15 for J matching. The database searched against was constructed from the NGS sequencing of the BACs which were used to generate the Kymouse.

If a sequence matched both a V and a J segment, the sequence between the two was then compared to a database of germline D segments in the mouse using 'blastn' with a wordsize of 4 and the options 'blastn-short' and 'ungapped'. This was used to assign a D segment, if possible. The CDR3 was identified by searching for the conserved "TATTACTGT" sequence in the V segment, and the "CTGGGG" in the J segment. If these motifs were not found, then up to 4 mismatches were allowed. The IMGT definition of CDR3 was used, so the CDR3 length is calculated from after the "TGT" in the V to before the "TGG" in the J. Sequences with an out of frame junction (those which do not have a CDR3 nucleotide length divisible by 3) or which contained a stop codon ("TAA", "TAG" or "TGA") were excluded.

The identity of the matching V, J and D segments as well as the CDR3 length from this assignment were then saved as a table for downstream analysis. The ratio of IGHJ6\*02 used increased from the naïve to immunised mice, as well as being enriched in the sub-population of sequences with a long HCDR3 (defined as consisting of 20 or more amino acids):

	All		HCDR3>20		% HCDR3>20
	JH6*02%	Total Count	JH6*02%	Total Count	
Naïve	22.31%	1340	91.11%	45	3.36%
Immunised	37.50%	256	66.67%	9	3.52%
Hybridoma	36.13%	119	63.64%	11	9.24%

This shows that the JH6\*02 gene segment is selected for by immunisation, as the proportion of JH6\*02 usage increases after immunisation. JH6\*02 is also used in the majority of antibodies with a long HCDR3 length, which is desirable for targets which are specifically bound by long HCDR3 length antibodies.

Additionally, the analysis revealed that certain VH and D gene segments frequently yielded HCDR3s of long length (in all of naïve, immunised and antigen-specific repertoires of heavy chains). See Table 2.

**Table 2A: Long HCDR3s from Naïve Repertoires**

<b>V</b>	<b>Average CDR3Length</b>	<b>Count</b>
IGHV1-2*02	21	3
IGHV1-18*01	21	5
IGHV3-7*01	22	3
IGHV6-1*01	21	5
IGHV3-9*01	20	2
IGHV2-5*10	20	1
IGHV7-4-1*01	21	3
IGHV1-3*01	21	5
IGHV4-4*02	20	3
IGHV3-13*01	22	1
IGHV3-23*04	20	1
IGHV1-8*01	21	10
IGHV3-21*03	23	3

<b>D</b>	<b>Average CDR3Length</b>	<b>Count</b>
IGHD2-2*02	20	1
IGHD3-9*01	21	13
IGHD3-10*01	21	26
IGHD6-13*01	20	1
IGHD4-17*01	22	2
IGHD6-19*01	23	1
IGHD3-22*01	20	1

<b>CDR3Length (All Naïve)</b>	<b>Count</b>
20	23
21	10
22	7
23	3
24	1
26	1

**Table 2B: Long HCDR3s from Immunised Repertoires**

<b>V</b>	<b>Average CDR3Length</b>	<b>Count</b>
IGHV4-4*02	20	1
IGHV3-11*01	23	2
IGHV3-7*01	21	6

<b>D</b>	<b>Average CDR3Length</b>	<b>Count</b>
IGHD2-2*02	22	2
IGHD3-10*01	22	5
IGHD6-19*01	20	1
IGHD1-26*01	20	1

<b>CDR3Length (All Immunised)</b>	<b>Count</b>
20	4
21	1
22	2
24	1
25	1

**Table 2C: Long HCDR3s from Antigen-Specific Repertoires**

<b>V</b>	<b>Average CDR3Length</b>	<b>Count</b>
IGHV4-4*02	20	2
IGHV1-3*01	21	3
IGHV3-11*01	21	1
IGHV3-7*01	22	1
IGHV1-8*01	22	2
IGHV3-20*d01	22	1
IGHV3-9*01	20	1

<b>D</b>	<b>Average CDR3Length</b>	<b>Count</b>
IGHD2-2*02	22	1
IGHD3-9*01	21	1
IGHD3-10*01	21	9

<b>CDR3Length (All Antigen-Specific)</b>	<b>Count</b>
20	4
22	2
21	4
24	1

**CLAIMS:**

1. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin D gene segment repertoire that is biased to the human D2 and/or D3 family or biased to one, more or all human D gene segments selected from the group D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.
2. The vertebrate or cell of claim 1, wherein the D gene segment repertoire consists of or substantially consists of one, two or three human gene segments selected from the group D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.
3. The vertebrate or cell of claim 1 or 2, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising (in 5' to 3' order) human VH, D and JH gene segments and said human D gene segments recited in claim 1 are spaced from the VH gene segment(s) by no more than four other D gene segments (eg, by no D gene segments).
4. The vertebrate or cell of any preceding claim, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising said human D gene segments and there are no other D gene segments in the locus between said human D gene segments.
5. The vertebrate or cell of any preceding claim, wherein the genome comprises three or more copies of a human D gene segment selected from D1-26, D2-2, D2-15, D3-3, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.
6. The vertebrate or cell of claim 5, wherein the genome comprises first and second human D gene segments selected from D1-26, D2-2, D2-15, D3-3, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19 when the first D gene segment is present as three or more copies and wherein the second D gene segment is present as three or more copies.
7. The vertebrate or cell of any preceding claim, wherein the D gene segments are selected from D2-2\*02, D3-9\*01, D3-10\*01 and D3-22\*01, or selected from D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01.
8. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell), optionally according to any preceding claim, whose genome comprises a human immunoglobulin VH gene segment repertoire that is biased to one, more or all of gene segments selected from the group VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.
9. The vertebrate or cell of claim 8, wherein the VH gene segment repertoire substantially consists of or substantially consists of one, two or three human gene segments selected from VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.



10. The vertebrate or cell of claim 8 or 9, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising (in 5' to 3' order) human VH, D and JH gene segments and said human VH gene segments are spaced from the D gene segment(s) by no more than four other VH gene segments (eg, by no VH gene segments).
11. The vertebrate or cell of any one of claims 8 to 10, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising said human VH gene segments and there are no other VH gene segments in the locus between said human VH gene segments.
12. The vertebrate or cell of any one of claims 8 to 11, wherein the genome comprises three or more copies of a human VH gene segment selected from the group consisting of VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.
13. The vertebrate or cell of claim 12, wherein the genome comprises first and second human VH gene segments selected from the group consisting of VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1 when the first VH gene segment is present as three or more copies and wherein the second VH gene segment is present as three or more copies.
14. The vertebrate or cell of any one of claims 8 to 13, wherein the VH gene segments are selected from the group consisting of VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01.
15. The vertebrate or cell of any one of claims 8 to 16, comprising an immunoglobulin heavy chain locus comprising two or more copies of a VH gene segment selected from said group.
16. The vertebrate or cell of any claim, wherein the genome comprises a human JH gene segment repertoire consisting of one or more human JH6 gene segments.
17. The vertebrate or cell of any preceding claim, wherein the genome comprises a human immunoglobulin JH gene segment repertoire that is biased to JH6, optionally JH6\*02.
18. The vertebrate or cell of claim 17, wherein the JH gene segment repertoire consists or substantially consists of three or more human JH6 gene segments.
19. The vertebrate or cell of any preceding claim, wherein the sequence of each of said human gene segments is a human germline gene segment sequence.
20. The vertebrate or cell of any preceding claim, wherein one, more or all of the selected gene segments are present in the genome as two or more copies, the copies being variants of each other.

21. The vertebrate or cell of any preceding claim, wherein said human gene segments are provided by homozygous immunoglobulin heavy chain loci.
22. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin VH gene segment repertoire, one or more human D gene segments and one or more human JH gene segments, wherein the VH repertoire does not comprise one, more or all VH gene segments selected from the group VH1-2, VH1-3, VH1-8, VH1-18, VH5-51, VH1-69, VH2-5, VH3-7, VH3-9, VH3-11, VH3-13, VH3-20, VH3-21, VH3-23, VH4-4, VH6-1 and VH7-4-1.
23. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin D gene segment repertoire, one or more human VH gene segments and one or more human JH gene segments, wherein the D repertoire does not comprise one, more or all D gene segments selected from the group D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19.
24. The vertebrate or cell of claim 22 or 23, wherein the genome comprises a human JH gene segment repertoire that does not comprise JH6.
25. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin JH gene segment repertoire that is biased to human JH6.
26. The vertebrate or cell of claim 25, wherein the genome comprises an unrearranged immunoglobulin heavy chain locus comprising a plurality of human JH6 gene segments; optionally wherein the genome is homozygous for said locus.
27. The vertebrate or cell of claim 26, wherein the heavy chain locus comprises (in 5' to 3' order) human VH, D and JH gene segments and said JH6 gene segments are spaced from the D gene segment(s) by no more than two other JH gene segments.
28. The vertebrate or cell of claim 25, 26 or 27, wherein are no other JH gene segments in the locus between said human JH6 gene segments.
29. A non-human vertebrate (eg, a mouse or a rat) or a non-human vertebrate cell (eg, a mouse cell or a rat cell) whose genome comprises a human immunoglobulin JH gene segment repertoire that consists of one or more human JH6 gene segments.
30. The vertebrate or cell of any one of claims 25 to 29, wherein all of said gene segments are human germline gene segments.

31. The vertebrate or cell of any one of claims 25 to 30, comprising different variant JH6 gene segments.
32. The vertebrate or cell of any one of claims 25 to 31, wherein said gene segments are provided by homozygous immunoglobulin heavy chain loci.
33. A monoclonal or polyclonal antibody composition or a population of antibody-producing cells for producing such composition, wherein the composition or population is prepared by immunising at least one vertebrate according to any preceding claim with an antigen, wherein the antibody or antibodies have human heavy chain variable regions comprising non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences; wherein the composition comprises at least one antigen-specific antibody having a HCDR3 length of at least 20 amino acids (according to IMGT).
34. An isolated antibody that specifically binds an antigen, the antibody comprising human heavy chain variable regions and non-human constant regions, wherein the variable regions are derived from the recombination in a non-human vertebrate of (i) a human VH gene segment selected from the group recited in claim 8 with (ii) a human D gene segment selected from the group recited in claim 1 and with a human JH gene segment (optionally JH6); wherein the antibody has a HCDR3 length of at least 20 amino acids (according to IMGT); and non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences.
35. The antibody of claim 34, wherein the antibody is obtained or obtainable from a vertebrate according to any one of claims 1 to 32.
36. A method of isolating an antibody that binds a predetermined antigen, the method comprising
  - (a) providing a vertebrate (optionally a mouse or rat) according to any one of claims 1 to 32;
  - (b) immunising said vertebrate with said antigen;
  - (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes expressing antibodies that bind to the antigen;
  - (d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by producing hybridomas therefrom; and
  - (e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes.
37. The method of claim 36, wherein in step (e) wherein the antibody has a HCDR3 length of at least 20 amino acids (according to IMGT).
38. The method of claim 36 or 37, comprising the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally exchanging the heavy chain

constant region nucleotide sequence of the antibody with a nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host.

39. The method of claim 36, 37 or 38, further comprising making a copy, mutant or derivative (eg, humanised version) of the antibody produced by the method.
40. The antibody composition, cell population, antibody or method of any one of claims 33 to 39, wherein the antigen is an antigen of an infectious disease pathogen; optionally wherein the pathogen is a virus or bacterium.
41. The antibody composition, cell population, antibody or method of claim 40, wherein pathogen is selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus.
42. The antibody composition, cell population, antibody or method of any one of claims 33 to 41, wherein the antigen is a HIV gp120 antigen or a HIV gp41 antigen.
43. The antibody composition, cell population, antibody or method of any one of claims 33 to 40, wherein the antigen comprises an active site or cleft, wherein the antibody having a HCDR3 length of at least 20 amino acids specifically binds to the active site or cleft of the antigen.
44. A pharmaceutical composition comprising an antibody or antibody composition according to any one of claims 33 to 35 and 40 to 43, or an antibody produced by the method of any one of claims 36 to 38, for treating and/or preventing an infectious disease in a human (eg, wherein the infectious disease is caused by a pathogen selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus).
45. A repertoire of antibody heavy chains (eg, provided by antibodies) comprising one or more heavy chains whose variable domain HCDR3 has a length of at least 20 amino acids (according to IMGT) and derived from the recombination of a human VH, D and JH, wherein

the VH is selected from the group

VH1-2\*02, VH1-3\*01, VH1-8\*01, VH1-18\*01, VH2-5\*10, VH3-7\*01, VH3-9\*01, VH3-11\*01, VH3-13\*01, VH3-21\*03, VH3-23\*04, VH4-4\*02, VH6-1\*01 and VH7-4-1\*01 and

the D is selected from the group

D2-2\*02, D3-9\*01, D3-10\*01 and D3-22\*01, or

D2-2\*02, D3-9\*01 and D3-10\*01, or

D3-9\*01 and D3-10\*01, or

D1-26, D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or

D1-26\*01, D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01, D4-17\*01, D6-13\*01 and D6-19\*01, or

D2-2, D3-9, D3-10, D3-22, D4-17, D6-13 and D6-19, or D2-2\*02, D3-9\*01, D3-10\*01, D3-22\*01,

D4-17\*01, D6-13\*01 and D6-19\*01, or  
D1-26, D2-2, D3-10 and D6-19, or  
D2-2, D3-9 and D3-10;

and optionally the JH is JH6 (eg, JH6\*02);

Wherein

(a) the heavy chain variable domain has been produced *in vivo* in a non-human vertebrate (eg, a mouse or a rat); and/or

(b) the heavy chain variable domain comprises non-human vertebrate AID-pattern somatic hypermutations, (eg, mouse or rat AID-pattern mutations) when compared to corresponding human germline V, D and J sequences and/or non-human (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations when compared to corresponding human germline V, D and J sequences.

46. A nucleic acid collection encoding the heavy chain repertoire of claim 45.
47. A method of obtaining an antigen-specific heavy chain (eg, provided by an antibody), the method comprising exposing the repertoire of claim 45 to a predetermined antigen and selecting one or more heavy chains that specifically bind to the antigen, wherein one or more heavy chains is isolated that has a HCDR3 length of at least 20 amino acids.
48. A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:
- (c) An immunoglobulin heavy chain locus comprising one or more human V gene segments, one or more human D gene segments and one or more human J gene segments upstream of a constant region; optionally wherein the heavy chain locus is according to (a) of claim 81; and
  - (d) An immunoglobulin light chain locus comprising either
    - (i) one or more human VH gene segments and one or more human J gene segments upstream of a constant region (optionally a rearranged  $V_HJ_LC_L$  or  $V_HJ_\lambda C_L$ , wherein the  $C_L$  is  $C_\lambda$  or  $C_k$ ); or
    - (ii) one or more human VL gene segments, one or more human D gene segments and one or more human J<sub>H</sub> gene segments upstream of a constant region (optionally a rearranged  $V_LDJ_HC_L$  or  $V_\lambda DJ_H C_L$ , wherein the  $C_L$  is  $C_\lambda$  or  $C_k$ ); or
    - (iii) one or more human VL gene segments selected from the group consisting of: a  $V_{\lambda II}$  gene family member,  $V_{\lambda VII 4A}$ ,  $V_{\lambda II 2.1}$ ,  $V_{\lambda VII 4A}$ , a  $V_{\lambda 1}$  gene family member, a  $V_{\lambda 3}$  gene family member, IGLV1S2,  $V_{\lambda 3-cML70}$ , IalH2, IalVI, IalH3, Kv325, a  $V_{kI}$  gene

family member,  $\kappa$ I-15A (KL012),  $V_{\kappa}$ II family member, a  $V_{\kappa}$ III family member, a  $V_{\kappa}$ I gene family member,  $\kappa$ I-15A (KL012),  $V_{\kappa}$ II A2 (optionally the A2a allele),  $V_{\kappa}$  A27 (Humkv325) and a gene segment at least 80% identical thereto, and one or more human  $J_L$  gene segments upstream of a constant region;

Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains produced by recombination of the light chain locus.

49. The vertebrate or cell of claim 48, wherein in (b)(i) the J gene segments of the light chain locus are  $J_{\lambda}$  gene segments and optionally the constant region of the light chain locus is a lambda constant region; or in (b)(ii) the VL is a  $V_{\lambda}$  and optionally the constant region of the light chain locus is a lambda constant region.

50. The vertebrate or cell of claim 48 or 49, wherein

in (b)(i) the V gene segment repertoire of the light chain locus comprises or consists of one or more VH gene segments selected from the group consisting of: a  $V_{HIII}$  gene family member (optionally, a  $V_{HIIIa}$  or  $V_{HIIIb}$  family member), a  $V_{HIV}$  gene family member,  $V_{HIII}$  9.1 (VH3-15),  $V_{HIII}$  VH26 (VH3-23),  $V_{H3}$ -21, LSG6.1, LSG12.1, DP77 (V3-21),  $V_H$  H11, VH1GRR, ha3h2,  $V_{HI}$ -ha1c1,  $V_{HIII}$ -VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto; or

in (b)(iii) the light chain locus V gene segment repertoire consists of one VL gene segment type (optionally and one or mutants thereof), wherein the VL gene segment is selected from said group of VL gene segments.

51. The vertebrate or cell of any one of claims 48 to 50, wherein in (a) said constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region and/or in (b) said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

52. The vertebrate or cell of any one of claims 48 to 51, wherein endogenous heavy and light chain expression has been inactivated.
53. A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:
- (f) An immunoglobulin heavy chain locus comprising one or more human V gene segments, one or more human D gene segments and one or more human J gene segments upstream of a constant region; and
  - (g)
    - (i) An unrearranged immunoglobulin light chain locus comprising one or more human VH gene segments and one or more human J gene segments upstream of a constant region, wherein each human VH gene segment is a human gene segment identical to (or mutant of) a human VH gene segment used to produce a rearranged VDJ encoding a heavy chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism); or
    - (ii) An immunoglobulin light chain locus comprising a rearranged VJ region or VDJ region upstream of a constant region, wherein the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VH gene segment that is identical to (or mutant of) the human VH gene segment used to produce a rearranged VDJ encoding a heavy chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);

- (h) Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments or VJ or VDJ in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains derived from the light chain locus;
- (i) Optionally when (b)(i) applies, each said VH gene segment in the light chain locus is selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto;
- (j) Optionally when (b)(ii) applies, the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto.
54. The vertebrate or cell of claim 53, wherein the J gene segments of the light chain locus are  $J_{\lambda}$  gene segments and optionally the constant region of the light chain locus is a lambda constant region.
55. The vertebrate or cell of claim 53 or 54, wherein the V gene segment repertoire of the light chain locus comprises or consists of human VH1-69 gene segment.
56. The vertebrate or cell of claim 53, 54 or 55, wherein in (a) said constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.



57. The vertebrate or cell of claim 53, 54, 55, or 56, wherein in (b) said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.
58. The vertebrate or cell of any one of claims 53 to 57, wherein the genome of said vertebrate or cell is homozygous for light chain locus (b)(i) or (ii); optionally wherein:
- the V gene segment repertoire of the light chain loci consists of one or more human VH gene segments selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto; or
  - the recombined VJ or VDJ repertoire of the light chain loci consists of sequences identical to one or more nucleotide sequences produced by the recombination of a human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto, with a human J gene segment and optionally a human D gene segment.
59. The vertebrate or cell of any one of claims 53 to 58, wherein endogenous heavy and light chain expression has been inactivated, and wherein light chain loci according to any one of claims 53 to 58 are the only functional light chain loci in the genome of the vertebrate or cell.
60. The vertebrate or cell of any one of claims 53 to 59, wherein each immunoglobulin light chain locus of said vertebrate or cell is according to (b)(i) recited in claim 53 and comprises only a single human VH gene segment selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical thereto, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain so that all light chain loci comprise the

same, single human VH gene segment.

61. The vertebrate or cell of any one of claims 53 to 59, wherein each immunoglobulin light chain locus of said vertebrate or cell is according to (b)(ii) recited in claim 53 and comprises only a single rearranged VJ or VDJ region, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain so that all light chain loci comprise the same, single rearranged VJ or VDJ region.
62. The vertebrate or cell according to claim 60 or 61, wherein each immunoglobulin light chain locus further comprises a VH gene segment or rearranged region that is a mutant respectively of said selected human VH gene segment or rearranged region, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain mutant VH gene segment or rearranged region.
63. The vertebrate or cell according to claim 60, wherein each immunoglobulin light chain locus comprises only two human VH gene segments selected from said group, optionally wherein the genome of the vertebrate or cell is homozygous for said two light chain human VH gene segments.
64. The vertebrate or cell according to claim 61, wherein each immunoglobulin light chain locus comprises only two of said rearranged VJ or VDJ regions, optionally wherein the genome of the vertebrate or cell is homozygous for said two light chain rearranged VJ or VDJ regions.
65. A monoclonal or polyclonal antibody composition prepared by immunisation of at least one vertebrate (eg, mouse or rat) according to any one of claims 48 to 64 with an antigen, optionally wherein the antigen is an antigen of an infectious disease pathogen, optionally wherein the same antigen is used to immunise all the vertebrates; optionally wherein the antibody or antibodies are IgG-type.
66. A method of isolating an antibody that binds a predetermined antigen, the method comprising
  - (a) providing a vertebrate (optionally a mouse or rat) according to any one of claims 48 to 64;
  - (b) immunising said vertebrate with said antigen (optionally wherein the antigen is an antigen of an infectious disease pathogen);
  - (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes

expressing antibodies that bind to the antigen;

(d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by producing hybridomas therefrom; and

(e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes.

67. The method of claim 66, comprising the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally exchanging the heavy chain constant region nucleotide sequence of the antibody with a nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host.

68. The method of claim 66 or 67, further comprising making a mutant or derivative of the antibody produced by the method of claim 66 or 67.

69. A method of producing a polyclonal antibody mixture, the method comprising carrying out the method of any one of claims 66 to 68 by separately immunising first and second vertebrates (optionally first and second mice or first and second rats) with antigen and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):

(i) the vertebrates are immunised with the same antigen or different antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));

(ii) prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene) and optionally the identical J repertoire; optionally the light chain loci of the mammals are identical prior to immunisation;

(iii) prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.

70. A method of producing a polyclonal antibody mixture, the method comprising carrying out the method of any one of claims 66 to 68 by immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):
- (i) the antigens are expressed by the same pathogenic organism (or a family member thereof);
  - (ii) prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene) and optionally the identical J repertoire; optionally the light chain loci of the mammals are identical prior to immunisation;
  - (iii) prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.
71. A method of producing host cells capable of expressing a polyclonal antibody mixture, the method comprising, in a method according to claim 67:-
- (a) immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));
  - (k) isolating nucleic acid encoding first and second anti-antigen antibodies from B lymphocytes from said vertebrates;
  - (l) determining the nucleotide sequences of the heavy and light chain variable regions of the first antibody;
  - (m) determining the nucleotide sequence of the heavy variable region and optionally the light chain variable region of the second antibody;
  - (n) inserting the heavy chain variable region coding sequence of each antibody into a heavy chain expression vector; optionally wherein the constant region coding sequence of each heavy chain is exchanged for a nucleotide sequence that encodes a human or humanised constant region;

- (o) inserting the light chain variable region coding sequence of the first antibody into a light chain expression vector; optionally wherein the constant region coding sequence of the light chain of the first antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (p) optionally inserting the light chain variable region coding sequence of the second antibody into a light chain expression vector; optionally wherein the constant region coding sequence of the light chain of the second antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region; and
- (q) introducing each expression vector into a host cell and co-expressing antibody chains in a mixture of said host cells to produce antibodies, each antibody comprising one or both of said heavy chain variable regions and a light chain; optionally wherein the expression vectors are introduced together into the same host cell (eg, a CHO or HEK293 cell) so that the cell is capable of expressing antibody light chains and heavy chains, such that the cell or a plurality of the host cells express antibodies, each comprising one or both of said heavy chain variable regions and a light chain;

(r) optionally:

prior to immunisation the light chain loci of the vertebrates contain the identical VH gene repertoire (optionally a single VH gene segment) and optionally the identical J repertoire (optionally a single J gene segment); optionally the light chain loci of the vertebrates are identical prior to immunisation; or

prior to immunisation the light chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the light chain loci of the vertebrates are identical prior to immunisation.

72. A method of producing a monoclonal or polyclonal antibody mixture, the method comprising carrying out the method of claim 71 and expressing a monoclonal antibody or polyclonal mixture of said antibodies; optionally followed by isolating an antibody comprising the heavy chain variable region of the first and/or second antibodies.

73. The method of any one of claims 66 to 72, wherein each vertebrate used for immunisation is provided by

- (a) isolating from a human blood or tissue (eg, B lymphocytes, PBMCs, bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
- (b) determining which human germline VH gene segment was recombined in the human to produce the nucleotide sequence of said B lymphocyte that encodes the heavy chain variable region of the antibody;
- (c) constructing a transgenic vertebrate wherein said human germline VH gene segment is provided in a light chain locus thereof according to any one of claims 48 to 63; and
- (d) providing said transgenic vertebrate for immunisation in the method of any one of claims 66 to 72.

74. The method of any one of claims 66 to 72, wherein each vertebrate used for immunisation is provided by

- (a) isolating from a human blood or tissue (eg, B lymphocytes, PBMCs, bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
- (b) determining a nucleotide sequence of said B lymphocyte that encodes a rearranged VDJ or VJ region of the antibody;
- (c) constructing a transgenic vertebrate wherein said rearranged VDJ or VJ region is provided in a light chain locus thereof according to any one of claims 48 to 64; and
- (d) providing said transgenic vertebrate for immunisation in the method of any one of claims 66 to 72.

75. An isolated antibody (eg, IgG-type antibody) obtainable or obtained by the method of claim 72, or claim 73 or 74 when dependent from claim 72, or a mutant or derivative antibody thereof wherein (i) the isolated antibody comprises two copies of the heavy chain variable region of said first antibody paired with two copies of the light chain variable region of said first antibody; or (ii) the isolated antibody comprises two copies of the heavy chain variable region of said second

antibody paired with two copies of the light chain variable region of said first antibody; or (iii) the isolated antibody is a bispecific antibody comprising one copy of the heavy chain variable region of said first antibody paired with a copy of the light chain variable region of the first antibody, and one copy of the heavy chain variable region of said the antibody paired with a copy of the light chain variable region of the first antibody, optionally wherein the bispecific antibody binds to said first and second antigens recited in claim 71; optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease.

76. A monoclonal or polyclonal antibody mixture (eg, IgG-type antibody or antibodies), wherein the monoclonal antibody or mixture comprises or consists of antibodies produced by the method of any one of claims 65 to 71 and 72 to 74, or a mutant or derivative antibody thereof optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease, wherein optionally wherein each antibody binds an antigen of an infectious disease pathogen, preferably the same antigen.
77. The use of an isolated, monoclonal or polyclonal antibody according to claim 65, 75 or 76, or a mutant or derivative antibody thereof in the manufacture of a medicament for the treatment and/or prevention of an infectious disease, optionally wherein the infectious disease is a disease caused by a bacterial or viral pathogen.
78. The antibody, mixture or use according to claim 75, 76 or 77, wherein the infectious disease is selected from the group consisting of a disease caused by a pathogen selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus.
79. A nucleotide sequence encoding an antibody of claim 65, 75 or 76, optionally wherein the nucleotide sequence is part of a vector.
80. A pharmaceutical composition comprising the antibody or antibodies of claim 65, 75 or 76 and a diluent, excipient or carrier.
81. A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

- (c) An immunoglobulin heavy chain locus comprising either:-
- (i) one or more human VL gene segments, one or more human D gene segments and one or more human J gene segments upstream of a constant region (optionally a rearranged  $V_L D J_H C_H$  or  $V_\lambda D J_H C_H$ ); or
  - (ii) one or more human VH gene segments selected from the group consisting of: a VHIII gene family member (optionally, a VHIIIa or VHIIIb family member), a VHIV gene family member, VHIII 9.1 (VH3-15), VHIII VH26 (VH3-23), VH3-21, LSG6.1, LSG12.1, DP77 (V3-21), VH H11, VH1GRR, ha3h2, VHI-ha1c1, VHIII-VH2-1, VH4.18, ha4h3, Hv1051, 71-2, Hv1f10, VH4.11, 71-4, VH251, VH1-69 and a gene segment at least 80% identical; one or more human D gene segments and one or more human  $J_H$  gene segments upstream of a constant region; and
- (d) An immunoglobulin light chain locus comprising one or more human V gene segments and one or more human J gene segments upstream of a constant region, optionally wherein the light chain locus is according to (b)(i) or (b)(ii) of claim 48.

Wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof, and the gene segments in the light chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising heavy chains produced by recombination of the heavy chain locus and light chains produced by recombination of the light chain locus.

82. The vertebrate or cell of claim 81, wherein the V gene segment repertoire of the light chain locus comprises or consists of one or more VL gene segments selected from the group consisting of a  $V_\lambda$ II gene family member,  $V_\lambda$ VII 4A,  $V_\lambda$ II 2.1,  $V_\lambda$ VII 4A, a  $V_\lambda$ 1 gene family member, a  $V_\lambda$ 3 gene family member, IGLV1S2,  $V_\lambda$ 3-cML70, IalH2, IalVl, IalH3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto; and/or wherein or in (a)(ii) the heavy chain locus V gene segment repertoire consists of one VH gene segment type (optionally and one or mutants thereof), wherein the VH gene segment is selected from said group of VH gene segments.

83. The vertebrate or cell of any claim 81 or 82, wherein in (a) said constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region and/or in (b)



said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

84. The vertebrate or cell of any one of claims 81 to 83, wherein endogenous heavy and light chain expression has been inactivated.

85. A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises:

- (a) (i) An unrearranged immunoglobulin heavy chain locus comprising one or more human VL gene segments, one or more human D gene segments and one or more J<sub>H</sub> gene segments upstream of a constant region, wherein each human VL gene segment is a human gene segment identical to (or mutant of) a human VL gene segment used to produce a rearranged VJ encoding a light chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism); or
- (ii) An immunoglobulin heavy chain locus comprising a rearranged VJ region or VDJ region upstream of a constant region, wherein the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VL gene segment that is identical to (or mutant of) the human VL gene segment used to produce a rearranged VJ encoding a light chain variable region of a human antibody from an antibody-expressing cell wherein said antibody binds to an antigen of an infectious disease pathogen (optionally the variable regions of said antibody being identical to an antibody from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);

- (b) An immunoglobulin light chain locus comprising one or more human V gene segments and one or more human J gene segments upstream of a constant region; and
- (c) Wherein the gene segments in the light chain locus are operably linked to the constant region thereof, and the gene segments or VJ or VDJ in the heavy chain locus are operably linked to the constant region thereof, so that upon immunisation the mouse is capable of producing an antibody comprising light chains produced by recombination of the light chain locus and heavy chains derived from the heavy chain locus;
- (d) Optionally when (a)(i) applies, each said VL gene segment in the heavy chain locus is selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ 1 gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, Ialh2, Ialvl, Ia3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto;
- (e) Optionally when (a)(ii) applies, the nucleotide sequence of the recombined region is identical to a nucleotide sequence produced by the recombination of a human J gene segment and optionally a human D gene segment with a human VL gene segment selected from the group consisting of a VL gene segment selected from the group consisting of a V $\lambda$ II gene family member, V $\lambda$ VII 4A, V $\lambda$ II 2.1, V $\lambda$ VII 4A, a V $\lambda$ 1 gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, Ialh2, Ialvl, Ia3h3, Kv325, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ I gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto.

86. The vertebrate or cell of claim 85, wherein the VL gene segments of the heavy chain locus are V $\lambda$  gene segments.

87. The vertebrate or cell of claim 85 or 86, wherein in (a) said constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.
88. The vertebrate or cell of claim 85, 86 or 87, wherein in (b) said constant region is a light chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.
89. The vertebrate or cell of any one of claims 85 to 88, wherein the genome of said vertebrate or cell is homozygous for heavy chain locus (a)(i) or (ii); optionally wherein:
- the V gene segment repertoire of the heavy chain loci consists of one or more human VL gene segments selected from the group consisting of a VL gene segment selected from the group consisting of a  $V_{\lambda II}$  gene family member,  $V_{\lambda VII}$  4A,  $V_{\lambda II}$  2.1,  $V_{\lambda VII}$  4A, a  $V_{\lambda 1}$  gene family member, a  $V_{\lambda 3}$  gene family member, IGLV1S2,  $V_{\lambda 3}$ -cML70, Iah2, Ialvl, Ia3h3, Kv325, a  $V_{\kappa I}$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa II}$  family member, a  $V_{\kappa III}$  family member, a  $V_{\kappa I}$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa II}$  A2 (optionally the A2a allele),  $V_{\kappa}$  A27 (Humkv325) and a gene segment at least 80% identical thereto; or
  - the recombined VJ or VDJ repertoire of the heavy chain loci consists of sequences identical to one or more nucleotide sequences produced by the recombination of a human VL gene segment selected from the group consisting of a VL gene segment selected from the group consisting of a  $V_{\lambda II}$  gene family member,  $V_{\lambda VII}$  4A,  $V_{\lambda II}$  2.1,  $V_{\lambda VII}$  4A, a  $V_{\lambda 1}$  gene family member, a  $V_{\lambda 3}$  gene family member, IGLV1S2,  $V_{\lambda 3}$ -cML70, Iah2, Ialvl, Ia3h3, Kv325, a  $V_{\kappa I}$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa II}$  family member, a  $V_{\kappa III}$  family member, a  $V_{\kappa I}$  gene family member,  $\kappa I$ -15A (KL012),  $V_{\kappa II}$  A2 (optionally the A2a allele),  $V_{\kappa}$  A27 (Humkv325) and a gene segment at least 80% identical thereto with a human J gene segment and optionally a human D gene segment.
90. The vertebrate or cell of any one of claims 85 to 89, wherein endogenous heavy and light chain expression has been inactivated, and wherein heavy chain loci according to any one of claims 85 to 89 are the only functional heavy chain loci in the genome of the vertebrate or cell.
91. The vertebrate or cell of any one of claims 85 to 90, wherein each immunoglobulin heavy chain locus of said vertebrate or cell is according to (a)(i) recited in claim 85 and comprises only a single human VL gene segment selected from the group consisting of a VL gene segment selected from the group consisting of a  $V_{\lambda II}$  gene family member,  $V_{\lambda VII}$  4A,  $V_{\lambda II}$  2.1,  $V_{\lambda VII}$  4A, a

V $\lambda$ 1 gene family member, a V $\lambda$ 3 gene family member, IGLV1S2, V $\lambda$ 3-cML70, I $\alpha$ h2, I $\alpha$ vl, I $\alpha$ 3h3, K $\nu$ 325, a V $\kappa$ l gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II family member, a V $\kappa$ III family member, a V $\kappa$ l gene family member,  $\kappa$ I-15A (KL012), V $\kappa$ II A2 (optionally the A2a allele), V $\kappa$  A27 (Humkv325) and a gene segment at least 80% identical thereto, optionally wherein the genome of the vertebrate or cell is homozygous for said heavy chain so that all heavy chain loci comprise the same, single human VL gene segment.

92. The vertebrate or cell of any one of claims 85 to 90, wherein each immunoglobulin heavy chain locus of said vertebrate or cell is according to (a)(ii) recited in claim 85 and comprises only a single rearranged VJ or VDJ region, optionally wherein the genome of the vertebrate or cell is homozygous for said heavy chain so that all heavy chain loci comprise the same, single rearranged VJ or VDJ region.
93. The vertebrate or cell according to claim 91 or 92, wherein each immunoglobulin heavy chain locus further comprises a VL gene segment or rearranged region that is a mutant respectively of said selected human VL gene segment or rearranged region, optionally wherein the genome of the vertebrate or cell is homozygous for said light chain mutant VL gene segment or rearranged region.
94. The vertebrate or cell according to claim 91, wherein each immunoglobulin heavy chain locus comprises only two human VL gene segments selected from said group, optionally wherein the genome of the vertebrate or cell is homozygous for said two heavy chain human VL gene segments.
95. The vertebrate or cell according to claim 92, wherein each immunoglobulin heavy chain locus comprises only two of said rearranged VJ or VDJ regions, optionally wherein the genome of the vertebrate or cell is homozygous for said two heavy chain rearranged VJ or VDJ regions.
96. A monoclonal or polyclonal antibody composition prepared by immunisation of at least one vertebrate (eg, mouse or rat) according to any one of claims 81 to 95 with an antigen, optionally wherein the antigen is an antigen of an infectious disease pathogen, optionally wherein the same antigen is used to immunise all the vertebrates; optionally wherein the antibody or antibodies are IgG-type.

97. A method of isolating an antibody that binds a predetermined antigen, the method comprising
- (a) providing a vertebrate (optionally a mouse or rat) according to any one of claims 81 to 95;
  - (b) immunising said vertebrate with said antigen (optionally wherein the antigen is an antigen of an infectious disease pathogen);
  - (c) removing B lymphocytes from the vertebrate and selecting one or more B lymphocytes expressing antibodies that bind to the antigen;
  - (d) optionally immortalising said selected B lymphocytes or progeny thereof, optionally by producing hybridomas therefrom; and
  - (e) isolating an antibody (eg, and IgG-type antibody) expressed by the B lymphocytes.
98. The method of claim 97, comprising the step of isolating from said B lymphocytes nucleic acid encoding said antibody that binds said antigen; optionally exchanging the heavy chain constant region nucleotide sequence of the antibody with a nucleotide sequence encoding a human or humanised heavy chain constant region and optionally affinity maturing the variable region of said antibody; and optionally inserting said nucleic acid into an expression vector and optionally a host.
99. The method of claim 97 or 98, further comprising making a mutant or derivative of the antibody produced by the method of claim 97 or 98.
100. A method of producing a polyclonal antibody mixture, the method comprising carrying out the method of any one of claims 97 to 99 by separately immunising first and second vertebrates (optionally first and second mice or first and second rats) with antigen and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):
- (i) the vertebrates are immunised with the same antigen or different antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));
  - (ii) prior to immunisation the heavy chain loci of the vertebrates contain the identical VL gene

repertoire (optionally a single VL gene) and optionally the identical D and/or J repertoire;  
optionally the heavy chain loci of the mammals are identical prior to immunisation;

(iii) prior to immunisation the heavy chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the heavy chain loci of the vertebrates are identical prior to immunisation.

101. A method of producing a polyclonal antibody mixture, the method comprising carrying out the method of any one of claims 97 to 99 by immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens and combining the anti-antigen antibodies isolated from each vertebrate (or mutants or derivatives of said antibodies) to produce a polyclonal antibody mixture; optionally wherein the following apply separately or in combination ((i) and (ii); or (i) and (iii)):

(i) the antigens are expressed by the same pathogenic organism (or a family member thereof));

(ii) prior to immunisation the heavy chain loci of the vertebrates contain the identical VL gene repertoire (optionally a single VL gene) and optionally the identical D and/or J repertoire;  
optionally the heavy chain loci of the mammals are identical prior to immunisation;

(iii) prior to immunisation the heavy chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the heavy chain loci of the vertebrates are identical prior to immunisation.

102. A method of producing host cells capable of expressing a polyclonal antibody mixture, the method comprising, in a method according to claim 98:-

(a) immunising one or a plurality of vertebrates (optionally mice or rats) with first and second antigens (optionally wherein the different antigens are expressed by the same pathogenic organism (or a family member thereof));

(b) isolating nucleic acid encoding first and second anti-antigen antibodies from B lymphocytes from said vertebrates;

(c) determining the nucleotide sequences of the heavy and light chain variable regions of the first antibody;

- (d) determining the nucleotide sequence of the light variable region and optionally the heavy chain variable region of the second antibody;
- (e) inserting the light chain variable region coding sequence of each antibody into a light chain expression vector; optionally wherein the constant region coding sequence of each light chain is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (f) inserting the heavy chain variable region coding sequence of the first antibody into a heavy chain expression vector; optionally wherein the constant region coding sequence of the heavy chain of the first antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region;
- (g) optionally inserting the heavy chain variable region coding sequence of the second antibody into a heavy chain expression vector; optionally wherein the constant region coding sequence of the heavy chain of the second antibody is exchanged for a nucleotide sequence that encodes a human or humanised constant region; and
- (h) introducing each expression vector into a host cell and co-expressing antibody chains in a mixture of said host cells to produce antibodies, each antibody comprising one or both of said light chain variable regions and a heavy chain; optionally wherein the expression vectors are introduced together into the same host cell (eg, a CHO or HEK293 cell) so that the cell is capable of expressing antibody light chains and heavy chains, such that the cell or a plurality of the host cells express antibodies, each comprising one or both of said light chain variable regions and a heavy chain;
- (i) optionally:

prior to immunisation the heavy chain loci of the vertebrates contain the identical VL gene repertoire (optionally a single VL gene segment) and optionally the identical D and/or J repertoire (optionally a single D and J gene segment); optionally the heavy chain loci of the vertebrates are identical prior to immunisation; or

prior to immunisation the heavy chain loci of the vertebrates contain the identical rearranged VJ or VDJ repertoire (optionally a single VJ or VDJ); optionally the heavy chain loci of the vertebrates are identical prior to immunisation.

103. A method of producing a monoclonal antibody or polyclonal antibody mixture, the method comprising carrying out the method of claim 102 and expressing a monoclonal antibody or

polyclonal mixture of said antibodies; optionally followed by isolating an antibody comprising the light chain variable region of the first and/or second antibodies.

104. The method of any one of claims 97 to 103, wherein each vertebrate used for immunisation is provided by
- (a) isolating from a human blood or tissue (eg, B lymphocytes, PBMCs, bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
  - (b) determining which human germline VL gene segment was recombined in the human to produce the nucleotide sequence of said B lymphocyte that encodes the light chain variable region of the antibody;
  - (c) constructing a transgenic vertebrate wherein said human germline VL gene segment is provided in a heavy chain locus thereof according to any one of claims 81 to 95; and
  - (d) providing said transgenic vertebrate for immunisation in the method of any one of claims 97 to 103.
105. The method of any one of claims 97 to 103, wherein each vertebrate used for immunisation is provided by
- (a) isolating from a human blood or tissue (eg, B lymphocytes, PBMCs, bone marrow, spleen, tonsil or lymph node) sample a B lymphocyte that expresses an antibody that binds a predetermined antigen (eg, an antigen expressed by an infectious disease pathogen; optionally wherein said serum or tissue was from a human individual suffering, susceptible to, or recovered from, a disease or condition caused or mediated by an organism harbouring or secreting said antigen; or from a human individual harbouring said organism);
  - (b) determining a nucleotide sequence of said B lymphocyte that encodes a rearranged VDJ or VJ region of the antibody;
  - (c) constructing a transgenic vertebrate wherein said rearranged VDJ or VJ region is provided in a heavy chain locus thereof according to any one of claims 81 to 95; and
  - (d) providing said transgenic vertebrate for immunisation in the method of any one of claims 97 to 103.



106. An isolated antibody (eg, IgG-type antibody) obtainable or obtained by the method of claim 103, or claim 104 or 105 when dependent from claim 103, or a mutant or derivative antibody thereof wherein (i) the isolated antibody comprises two copies of the heavy chain variable region of said first antibody paired with two copies of the light chain variable region of said first antibody; or (ii) the isolated antibody comprises two copies of the heavy chain variable region of said second antibody paired with two copies of the light chain variable region of said first antibody; or (iii) the isolated antibody is a bispecific antibody comprising one copy of the heavy chain variable region of said first antibody paired with a copy of the light chain variable region of the first antibody, and one copy of the heavy chain variable region of said the antibody paired with a copy of the light chain variable region of the first antibody, optionally wherein the bispecific antibody binds to said first and second antigens recited in claim 102; optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease.
107. A monoclonal or polyclonal antibody mixture (eg, IgG-type antibody or antibodies), wherein the monoclonal antibody or mixture comprises or consists of antibodies produced by the method of any one of claims 97 to 101 and 103 to 105, or a mutant or derivative antibody thereof optionally for use in medicine, optionally for the treatment and/or prevention of an infectious disease, wherein optionally wherein each antibody binds an antigen of an infectious disease pathogen, preferably the same antigen.
108. An isolated chimaeric antibody for treating and/or preventing an infectious disease or condition, the antibody comprising a non-human vertebrate (optionally a mouse or rat) heavy chain constant regions and human variable regions that bind an antigen of an infectious disease pathogen, wherein the antibody is obtainable or obtained in a method comprising immunisation of a vertebrate of any one of claims 48 to 64 and 81 to 95 with said antigen.
109. An isolated human antibody for treating and/or preventing an infectious disease or condition, the antibody comprising human heavy chain constant regions and human variable regions that bind an antigen of an infectious disease pathogen, wherein the antibody is obtainable or obtained in a method comprising affinity maturation of antibody variable regions *in vivo* in a transgenic non-human vertebrate (eg, mouse or rat) when said variable regions are operably linked to heavy chain constant regions of said vertebrate (eg, mouse or rat heavy chain constant regions) by (a) immunisation of a vertebrate of any one of claims 48 to 64 and 81 to 95 with said antigen, (b) isolation of nucleic acid encoding a chimaeric antibody according to claim

108, (c) replacing the nucleotide sequences of the nucleic acid that encode the non-human vertebrate heavy chain constant regions with nucleotide sequence encoding human heavy chain constant regions to produce nucleic acid encoding a human antibody; (d) expressing the human antibody *in vitro* (optionally from CHO or HEK293 cells harbouring the human nucleic acid) and (e) isolating the human antibody (optionally with further affinity maturation of the antibody and/or producing a derivative thereof).

110. An antibody comprising human variable domains that bind a predetermined antigen (eg, an antigen expressed by a bacterial or viral pathogen), wherein the variable domain sequences are encoded by rearranged VDJ and VJ regions, each of the VDJ and/or VJ being a hybrid region produced by the *in vivo* rearrangement of human heavy and light chain variable region gene segments (V and J and optionally D segments); optionally wherein the antibody comprises human constant regions.
111. A mixture of first and second human antibodies, each antibody being according to claim 109 or 110 and capable of binding to an antigen of an infectious disease pathogen (optionally wherein the first antibody binds a first antigen and the second antibody binds a second antigen, said antigens being from the same pathogen; or wherein the antigens are the same).
112. The mixture of claim 111, wherein the light chain amino acid sequence of the first antibody is identical to the light chain amino acid sequence of the second antibody, or has up to 15 amino acid changes therefrom.
113. The use of an isolated, monoclonal or polyclonal antibody or mixture according to any one of claims 96 and 106 to 112, in the manufacture of a medicament for the treatment and/or prevention of an infectious disease, optionally wherein the infectious disease is a disease caused by a bacterial or viral pathogen.
114. The antibody, mixture or use according to any one of claims 96 and 106 to 112, wherein the infectious disease is selected from the group consisting of a disease caused by a pathogen selected from the group consisting of Haemophilus influenza, E coli, Neisseria meningitidis, a herpes family virus, cytomegalovirus (CMV), HIV and influenza virus.

115. First and second nucleotide sequences encoding the heavy and light chains of an antibody recited in any one of claims 96 and 106 to 112 or at least the variable regions thereof, optionally wherein each nucleotide sequence is part of a vector.
116. A host cell comprising one or more expression vectors encoding the heavy chains of the first and second antibodies recited in claim 110 or 112, and the light chain of the first antibody recited in claim 110 or 112 (and optionally also the light chain of the second antibody recited in claim 110 or 112).
117. A pharmaceutical composition comprising the antibody or antibodies recited in any one of claims 96 and 106 to 112 and a diluent, excipient or carrier; optionally wherein the composition is provided in a container connected to an IV needle or syringe or in an IV bag.
118. A synthetic immunoglobulin locus comprising one or more variable and J gene segments (and optionally one or more D gene segments) operably linked 5' of a constant region, wherein the locus comprises a 5' to 3' V(D)J arrangement selected from the group consisting of immunoglobulin locus can be constructed with one or more of the following arrangements (5' to 3'):-
- (a) [V (heavy, lambda or kappa)] – [two-turn RSS] --- [one-turn RSS] – [D] – [JH], wherein said RSSs are in an opposite orientation;
  - (b) [VH] – [D] – [two-turn RSS] --- [one-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
  - (c) [VH] – [D] – [one-turn RSS] --- [two-turn RSS] – [J kappa], wherein said RSSs are in an opposite orientation;
  - (d) [VH or V kappa] - [two-turn RSS] --- [one-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
  - (e) [V kappa] - [one-turn RSS] --- [two-turn RSS] – [JH or J lambda], wherein said RSSs are in an opposite orientation;
  - (f) [V (heavy, lambda or kappa)] – [one-turn RSS] --- [two-turn RSS] – [D] – [JH], wherein said RSSs are in an opposite orientation;
  - (g) [VH] – [D] – [one-turn RSS] --- [two-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;

- (h) [VH] – [D] – [two-turn RSS] --- [one-turn RSS] – [J kappa], wherein said RSSs are in an opposite orientation;
- (i) [VH or V kappa] - [one-turn RSS] --- [two-turn RSS] – [J lambda], wherein said RSSs are in an opposite orientation;
- (j) [V kappa] - [two-turn RSS] --- [one-turn RSS] – [JH or J lambda], wherein said RSSs are in an opposite orientation.

119. A non-human vertebrate (eg, a mouse or rat) or cell comprising the locus of claim 118.

# Recombineered BAC Vectors to add Polymorphic V-regions to the Mouse Genome

## 1. Cassette Exchange --- Replace existing V-regions

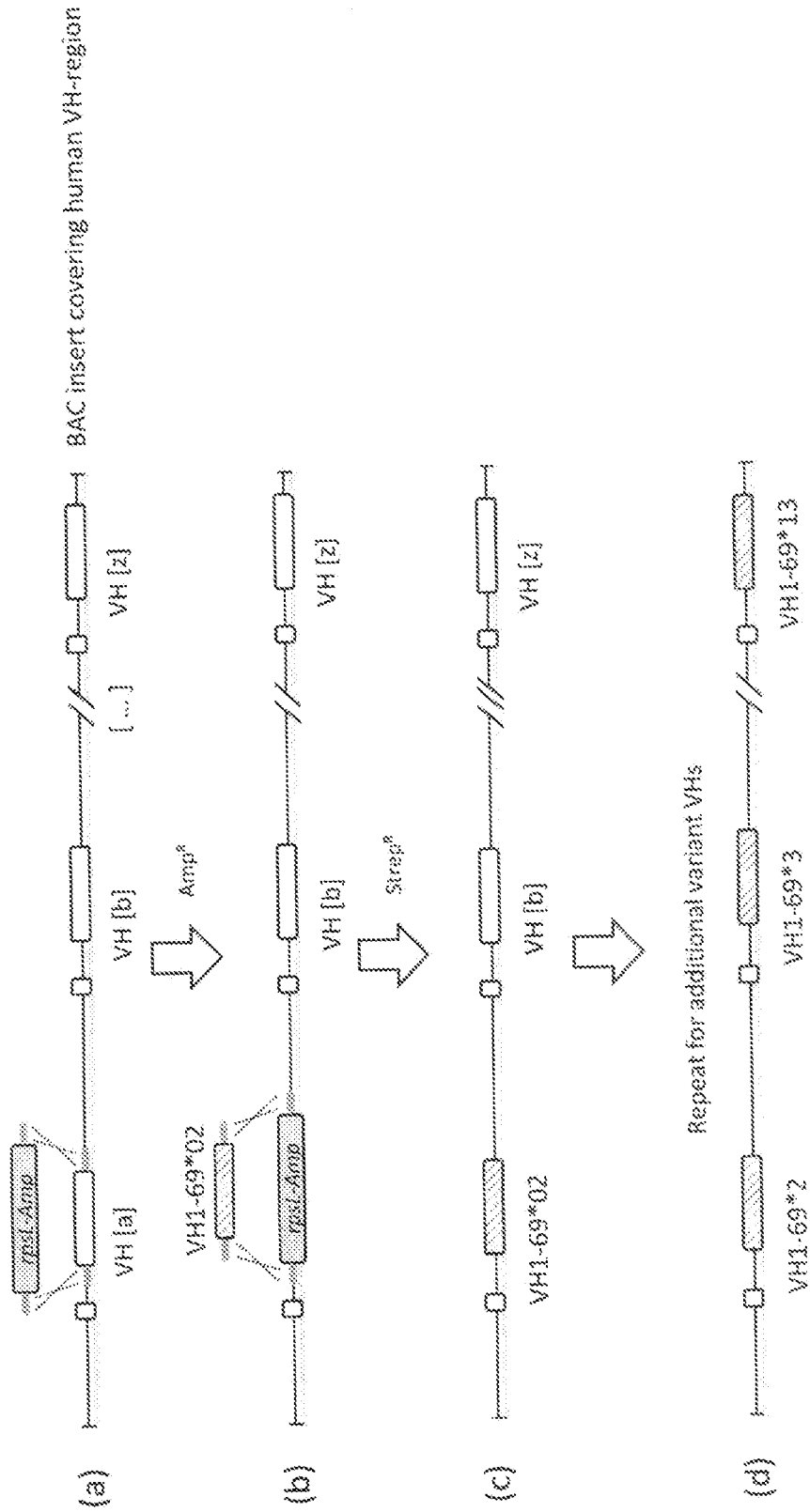


Figure 1

# Recombined BAC Vectors to add Heterogeneous V-regions to the Mouse Genome

## 2. Insertion of V-regions into Heterologous Genomic DNA

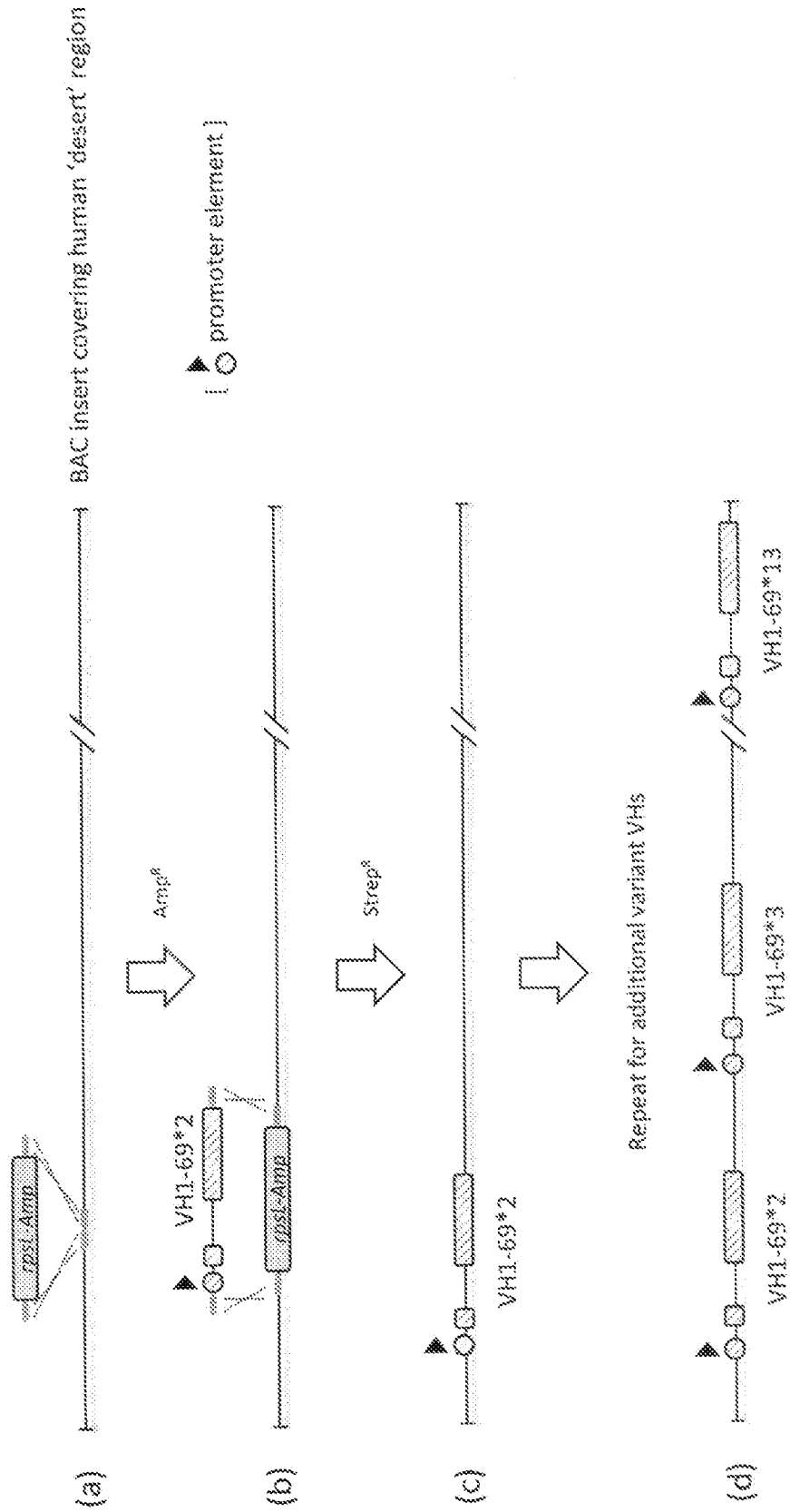


Figure 2

# Recombineered BAC Vectors to add Heterogeneous V-regions to the Mouse Genome

## 3. Single Strand Nucleotide Changes

[ Single Stranded targeting fragment with 2 nucleotide changes ]

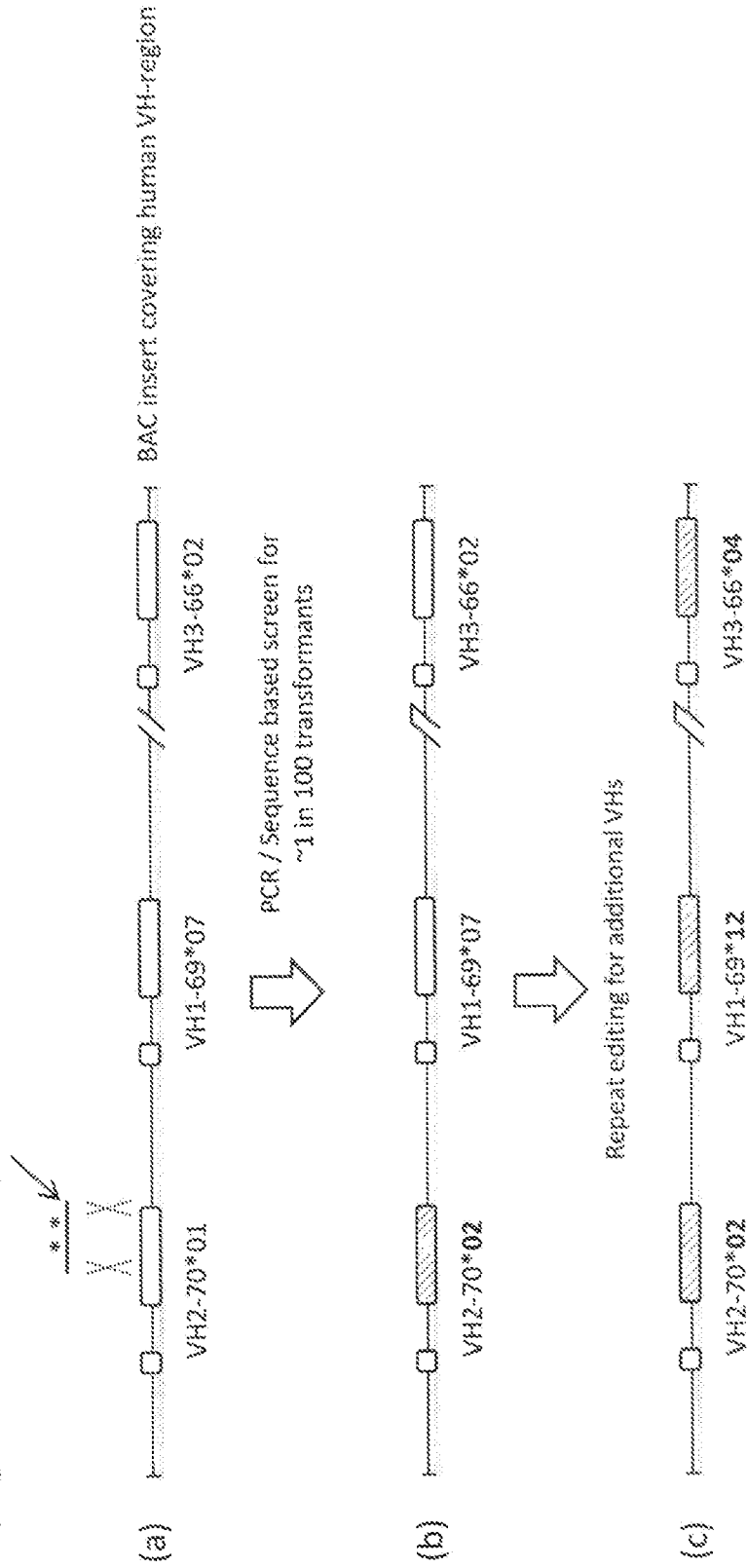


Figure 3

# Adding Polymorphic V-regions to the Genome using SRMCE of Modified BACs

1. Insertion at a 'Landing Pad' site during SRMCE

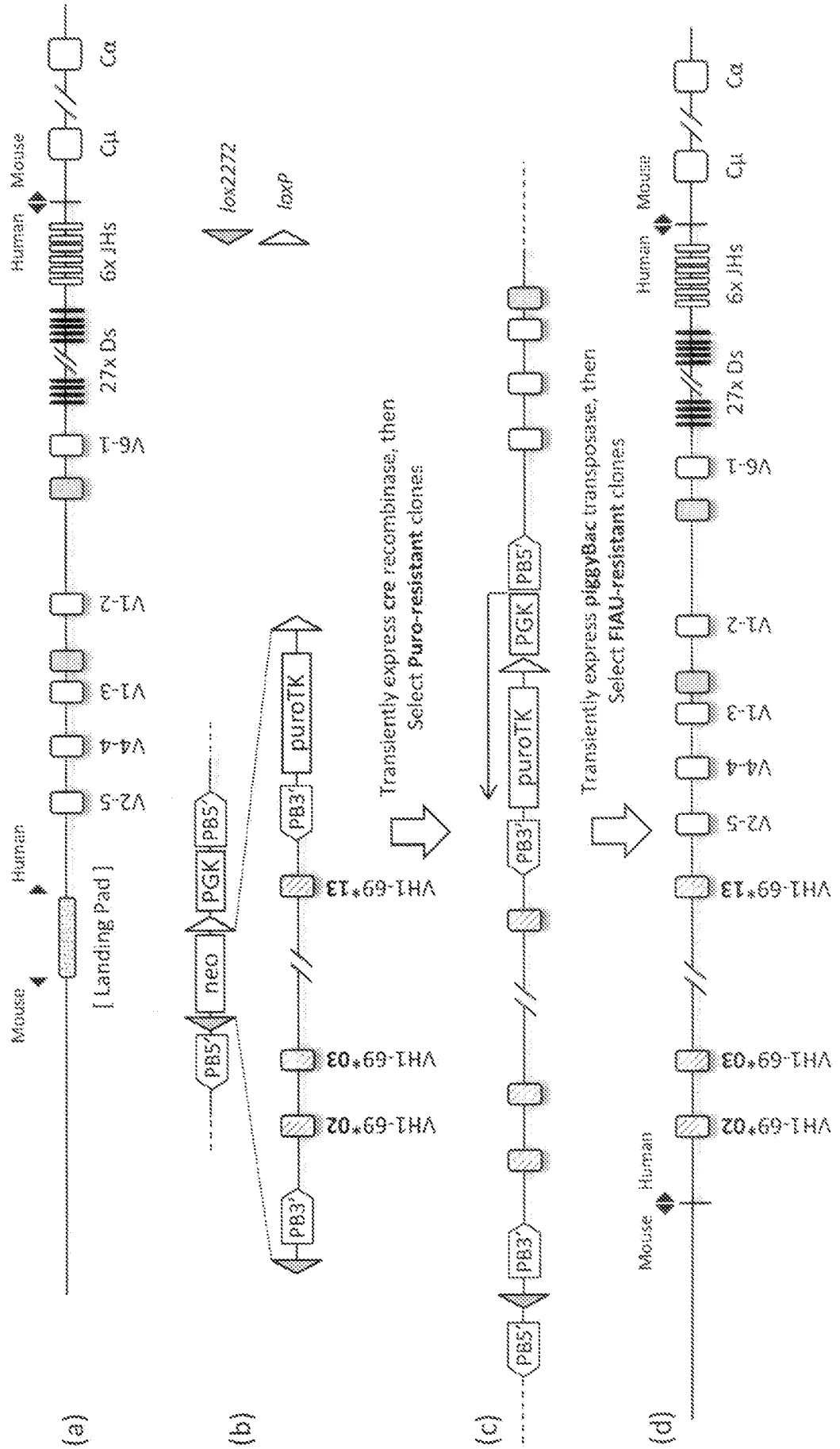


Figure 4



```

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Q V Q L V Q S G A E V K K F G S S V K V S C K A S
CAG GTG CAG CTG GTG CAG TCT GCG GCT ... GAG GTC AAG AAG CCT GGG TCC TCG GTG ARG GTC TCC TGC AAG GCT TCT
Z17506 ,IGHV1-69*02, y15H6(YAC7)---A .....
X92340 ,IGHV1-69*03, 576TA6 .....
MR3132 ,IGHV1-69*04, h91L63 .....
X67905 ,IGHV1-69*05, FR.VH1.2 .....
L22563 ,IGHV1-69*06, h91051K .....
Z29978 ,IGHV1-69*07, DA-2 .....
Z14309 ,IGHV1-69*08 .....
Z14307 ,IGHV1-69*09 .....
Z14300 ,IGHV1-69*10 .....
Z14296 ,IGHV1-69*11 .....
Z14301 ,IGHV1-69*12 .....
Z14210 ,IGHV1-69*13 .....

```

Alignment of 13 IGHV1-69 alleles showing the variable (V) coding region only. Nucleotides that differ from VH1-69 allele \*01 are indicated at the appropriate position whereas identical nucleotides are marked with a dash. Where nucleotide changes result in amino acid differences, the encoded amino acid is shown above the corresponding triplet. Boxed regions correspond to CDR1, CDR2 and CDR3 as indicated.

Figure 5 (part 1 of 4)

	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			
	G	G	T	F	S	S	Y	A					I	S	W	V	E	Q	A	P	S	Q	S	L	S	R	M			
	GGA	GGC	ACC	ATC	AGC	ACC	TAT	TCT					ATC	ASC	TGG	GTG	CGA	CAG	GCC	CCG	GGA	CAA	GGG	GTT	GAG	GGG	ATG			
122582																														
227506								A																						
X92340																														
M03152																														
X67905																														
L22583																														
Z09978																														
Z14309																														
Z14307																														
Z14300																														
Z14296																														
Z14301																														
Z14214																														

Alignment of 13 IGHV1-69 alleles showing the variable (V) coding region only. Nucleotides that differ from VH1-69 allele \*01 are indicated at the appropriate position whereas identical nucleotides are marked with a dash. Where nucleotide changes result in amino acid differences, the encoded amino acid is shown above the corresponding triplet. Boxed regions correspond to CDR1, CDR2 and CDR3 as indicated.

Figure 5 (part 2 of 4)

	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	
122582	IGHV1-69*01, hv1051	GGA	GGG	ATC	ATC	CTT	ATT	GGT	ACA	GCA	...	AAC	TAC	GCA	CAG	CAG	...	GGC	AGA	GTC	ACC	ATT	ACC	GGG	...	...	...	...
227506	IGHV1-69*02, Y1GH6(YAC7	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
X92340	IGHV1-69*03, 570E248	K	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
883132	IGHV1-69*04, hv1263	K	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
X67905	IGHV1-69*05, FR.VH1.2	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
122583	IGHV1-69*06, hv1051K	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Z59970	IGHV1-69*07, DA-2	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
214309	IGHV1-69*08	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
214307	IGHV1-69*09	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
214300	IGHV1-69*10	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
214296	IGHV1-69*11	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
214301	IGHV1-69*12	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
214214	IGHV1-69*13	R	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Alignment of 13 IGHV1-69 alleles showing the variable (V) coding region only. Nucleotides that differ from VH1-69 allele \*01 are indicated at the appropriate position whereas identical nucleotides are marked with a dash. Where nucleotide changes result in amino acid differences, the encoded amino acid is shown above the corresponding triplet. Boxed regions correspond to CDR1, CDR2 and CDR3 as indicated.

Figure 5 (part 3 of 4)

CDR3...

81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	
D	S	S	T	S	T	A	Y	M	E	L	S	S	L	R	G	E	D	T	A	Y	Y	Y	C	A	R	
GAC	GAA	TTC	ACG	ASC	ACA	GCC	TAC	AIG	SAG	CTG	AGC	AGC	CTG	AGA	TCT	SAG	SAC	ACG	GCC	GTG	TAT	TAC	TGT	GCG	AGA	GA
	K																									
227506																										
	A																									
292340																										
	K																									
263132																										
	A																									
267905																										
	K																									
222583																										
	A																									
229976																										
	K																									
214309																										
	A																									
214307																										
	K																									
214300																										
	A																									
214296																										
	K																									
214301																										
	A																									
214214																										
	K																									

Alignment of 13 IGHV1-69 alleles showing the variable (V) coding region only. Nucleotides that differ from VH1-69 allele \*01 are indicated at the appropriate position whereas identical nucleotides are marked with a dash. Where nucleotide changes result in amino acid differences, the encoded amino acid is shown above the corresponding triplet. Boxed regions correspond to CDR1, CDR2 and CDR3 as indicated.

Figure 5 (part 4 of 4)

Figure 6

