ANTIBODIES, VARIABLE DOMAINS & CHAINS TAILORED FOR HUMAN USE

FIELD OF THE INVENTION

The present invention relates to the provision of antibody therapeutics and prophylactics that are tailored specifically for human use.

The present invention provides libraries, 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, heavy chains, polypeptides, nucleotide sequences, pharmaceutical compositions and uses are also provided by the invention.

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 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 of the human V(D) J repertoire. Existing transgenic immunoglobulin loci are 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.

The inventors considered that it would be desirable to tailor the genomes of these transgenic non-human vertebrates (and thus antibody and antibody chain products of these) to address the variability – and commonality – in the natural antibody gene usage of humans. The inventors wanted to do this in order to better address human use of antibody-based therapeutic and prophylactic drugs.

It would be desirable also 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.

SUMMARY OF THE INVENTION

The present invention has been developed from extensive bioinformatics analysis of natural antibody gene segment distributions across a myriad of different human populations and across more than two thousand samples from human individuals. The inventors have undertaken this huge task to more thoroughly understand and design non-human vertebrate systems and resultant antibodies to better address human medical therapeutics as a whole, as well as to enable rational design to address specific ethnic populations of humans. Using such rational design, the inventors have constructed transgenic non-human vertebrates and isolated antibodies, antibody chains and cells expressing these in a way that yields products that utilise gene segments that have been purposely included on the basis of the human bioinformatics analysis. The examples illustrate worked experiments where the inventors isolated many cells and antibodies to this effect.

The invention also relates to synthetically-extended & ethnically-diverse superhuman immunoglobulin gene repertoires. The present invention thus provides for novel and potentially expanded synthetic immunoglobulin diversities, thus providing a pool of diversity from which human antibody therapeutic leads can be selected. This expanded pool is useful when seeking to find antibodies with desirable characteristics, such as relatively high affinity to target antigen without the need for further affinity maturation (eg, using laborious *in vitro* techniques such as phage or ribosome display), or improved biophysical characteristics, or to address targets and new epitopes that have previously been difficult to address with antibodies are not reached by prior antibody binding sites.

The invention also provides for diversity that is potentially biased towards variable gene usage common to members of a specific human population, which is useful for generating antibodies for treating and/or preventing diseases or conditions within such population. This ability to bias the

antibody repertoire allows one to tailor antibody therapeutics with the aim of more effectively treating and/or preventing disease or medical conditions in specific human populations.

The present inventors realised the possibility of providing immunoglobulin gene segments from disparate sources in transgenic loci, in order to provide for novel and potentially-expanded antibody diversities from which antibody therapeutics (and antibody tool reagents) could be generated. This opens up the potential of transgenic human-mouse/rat technologies to the possibility of interrogating different and possibly larger antibody sequence-spaces than has hitherto been possible.

In rationally designing transgenic antibody loci, as well as antibodies and antibody chains, the inventors also realised that a relatively long HCDR3 length (at least 20 amino acids) is often desirable to address epitopes. For example, naturally-occurring antibodies have been isolated from humans infected with infectious disease pathogens, such antibodies having a long HCDR3 length. Neutralising antibodies have been found in this respect. A long HCDR3 length would be desirable to address other antigens (eg, receptor clefts or enzyme active sites), not just limited to infectious disease pathogens, and thus the inventors realised the general desirability of the possibility of engineering transgenic loci to be able to produce long HCDR3 antibodies and heavy chains. The inventors, through laborious execution of bioinformatics on in excess of 2000 human DNA samples *via* the 1000 Genomes project together with rational sequence choices, identified that the inclusion of the specific human gene segment variant JH6*02 is desirable for producing long HCDR3 antibodies and chains.

Additional rational design and bioinformatics has led the inventors to realise that specific human constant region variants are conserved across many diverse human populations. The inventors realised that this opens up the possibility of making a choice to humanise antibodies, chains and variable domains by using such specific constant regions in products, rather than arbitrarily choosing the human constant region (or a synthetic version of a human constant region). This aspect of the invention also enables one to tailor antibody-based drugs to specific human ethnic populations, thereby more closely matching drug to patient (and thus disease setting) than has hitherto been performed. It can be a problem in the state of the art that antibodies are humanised with an arbitrary choice of human constant region (presumably derived from one (often unknown) ethnic

population or non-naturally occurring) that does not function as well in patients of a different human ethnic population. This is important, since the constant region has the major role in providing antibody effector functions, eg, for antibody recycling, cellular and complement recruitment and for cell killing.

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

First Configuration

A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) comprising a genome having a superhuman immunoglobulin heavy chain human VH and/or D and/or J gene repertoire.

A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) comprising a genome having a superhuman immunoglobulin light chain human VL gene repertoire; optionally wherein the vertebrate or cell is according to the first configuration.

A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises a transgenic immunoglobulin locus (eg, a heavy chain locus or a light chain locus), said locus comprising immunoglobulin gene segments according to the first and second human immunoglobulin gene segments (optionally V segments) as mentioned below operably connected upstream of an immunoglobulin constant region; optionally wherein the genome is homozygous for said transgenic immunoglobulin locus;
optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or
optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or
optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

A transgenic non-human vertebrate (eg, a mouse or rat) or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises a transgenic immunoglobulin locus comprising a plurality of human immunoglobulin gene segments operably connected upstream of a non-human vertebrate constant region for the production of a repertoire of chimaeric antibodies, or chimaeric light or heavy chains, having a non-human vertebrate constant region and a human variable region; wherein the transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments, a first (optionally a V segment) of said gene segments and a second (optionally a V segment) of said gene segments being different and derived from the genomes of first and second human individuals respectively, wherein the individuals are different; and optionally not related;
optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or
optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or
optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

A transgenic non-human vertebrate (eg, a mouse or rat) or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises first and second transgenic immunoglobulin loci, each locus comprising a plurality of human immunoglobulin gene segments operably connected upstream of a non-human vertebrate constant region for the production of a repertoire of chimaeric antibodies, or chimaeric light or heavy chains, having a non-human vertebrate constant region and a human variable region;

wherein (i) the first transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments, (ii) the second transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments; and (iii) wherein a first (optionally a V) gene segment of said first locus and a second (optionally a V) gene segment of said second gene locus are different and derived from the genomes of first and second human individuals respectively, wherein the individuals are different; and optionally not related;

optionally wherein the first and second loci are on different chromosomes (optionally chromosomes with the same chromosome number) in said genome;

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

A method of constructing a cell (eg, an ES cell) according to the invention, the method comprising

(a) identifying functional V and J (and optionally D) gene segments of the genome sequence of a (or said) first human individual;

(b) identifying one or more functional V and/or D and/or J gene segments of the genome sequence of a (or said) second human individual, wherein these additional gene segments are not found in the genome sequence of the first individual;

(c) and constructing a transgenic immunoglobulin locus in the cell, wherein the gene segments of (a) and (b) are provided in the locus operably connected upstream of a constant region.

In one embodiment, the gene segment(s) in step (b) are identified from an immunoglobulin gene database selected from the 1000 Genomes, Ensembl, Genbank and IMGT databases.

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, 16th March 2012; the Ensembl database is assembly GRCh37.p8 (10/04/2012); the IMGT database is available at www.imgt.org .

In one embodiment, the first and second human individuals are members of first and second ethnic populations respectively, wherein the populations are different, optionally wherein the human

immunoglobulin gene segment derived from the genome sequence of the second individual is low-frequency (optionally rare) within the second ethnic population.

This configuration of the invention also provides a method of making a transgenic non-human vertebrate (eg, a mouse or rat), the method comprising

(a) constructing an ES cell (eg, a mouse C57BL/6N, C57BL/6J, 129S5 or 129Sv strain ES cell) by carrying out the method above;

(b) injecting the ES cell into a donor non-human vertebrate blastocyst (eg, a mouse C57BL/6N, C57BL/6J, 129S5 or 129Sv strain blastocyst);

(c) implanting the blastocyst into a foster non-human vertebrate mother (eg, a C57BL/6N, C57BL/6J, 129S5 or 129Sv strain mouse); and

(d) obtaining a child from said mother, wherein the child genome comprises a transgenic immunoglobulin locus.

In one embodiment, the invention provides a method of isolating an antibody that binds a predetermined antigen (eg, a bacterial or viral pathogen antigen), the method comprising immunising a non-human vertebrate according to the invention.

Second Configuration

A library of antibody-producing transgenic cells whose genomes collectively encode a repertoire of antibodies, wherein

(a) a first transgenic cell expresses a first antibody having a chain encoded by a first immunoglobulin gene, the gene comprising a first variable domain nucleotide sequence produced following recombination of a first human unrearranged immunoglobulin gene segment;

(b) a second transgenic cell expresses a second antibody having a chain encoded by a second immunoglobulin gene, the second gene comprising a second variable domain nucleotide sequence produced following recombination of a second human unrearranged immunoglobulin gene segment, the first and second antibodies being non-identical;

(c) the first and second gene segments are different and derived from the genome sequences of first and second human individuals respectively, wherein the individuals are different; and optionally not related;

(d) wherein the cells are non-human vertebrate (eg, mouse or rat) cells.

In one embodiment, the first and second human individuals are members of first and second ethnic populations respectively, wherein the populations are different; optionally wherein the ethnic populations are selected from those identified in the 1000 Genomes database.

In another embodiment, the second human immunoglobulin gene segment is a polymorphic variant of the first human immunoglobulin gene segment; optionally wherein the second gene segment is selected from the group consisting of a gene segment in any of Tables 1 to 7 and 9 to 14 below (eg, selected from Table 13 or Table 14), eg, the second gene segment is a polymorphic variant of VH1-69.

Third Configuration

An isolated antibody having

(a) a heavy chain encoded by a nucleotide sequence produced following recombination in a transgenic non-human vertebrate cell of an unrearranged human immunoglobulin V gene segment with a human D and human J segment, optionally with affinity maturation in said cell, wherein one of the gene segments is derived from the genome of an individual from a first human ethnic population; and the other two gene segments are derived from the genome of an individual from a second, different, human ethnic population, and wherein the antibody comprises heavy chain constant regions of said non-human vertebrate (eg, rodent, mouse or rat heavy chain constant regions); and/or

(b) a light chain encoded by a nucleotide sequence produced following recombination in a transgenic non-human vertebrate cell of an unrearranged human immunoglobulin V gene segment with a human J segment, optionally with affinity maturation in said cell, wherein one of the gene segments is derived from the genome of an individual from a first human ethnic population (optionally the same as the first population in (a)); and the other gene segment is derived from the genome of an individual from a second, different, human ethnic population (optionally the same as the second population in (a)), and wherein the antibody comprises light chain constant regions of said non-human vertebrate (eg, rodent, mouse or rat heavy light constant regions);

(c) Optionally wherein each variable domain of the antibody is a human variable domain.

(d) Optionally wherein the heavy chain constant regions are gamma-type constant regions.

The invention also provides an isolated nucleotide sequence encoding the antibody, optionally wherein the sequence is provided in an antibody expression vector, optionally in a host cell.

The invention also provides a method of producing a human antibody, the method comprising replacing the non-human vertebrate constant regions of the antibody of the third configuration with human antibody constant regions.

The invention also provides a pharmaceutical composition comprising an antibody according to the third configuration, or an antibody produced according to the method above 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 invention also provides an antibody-producing cell that expresses the second antibody recited in any one of the configurations.

In an alternative configuration, the invention contemplates the combination of nucleotide sequences of first and second immunoglobulin gene segments (eg, two or more polymorphic variants of a particular human germline VH or VL gene segment) to provide a synthetic gene segment. Such synthetic gene segment is used, in one embodiment, to build a transgenic immunoglobulin locus, wherein the synthetic gene segment is provided in combination with one or more human variable and J regions (and optionally one or more human D regions) operably connected upstream of a constant region. When provided in the genome of a non-human vertebrate or cell (eg, mouse or rat cell, eg, ES cell), the invention provides for superhuman gene segment diversity. The sequences to be combined can be selected from gene segments that have been observed to be commonly used in human antibodies raised against a particular antigen (eg, a flu antigen, such as haemaglutinin). By combining the sequences, the synthetic gene segment may

recombine *in vivo* to produce an antibody that is well suited to the treatment and/or prevention of a disease or condition (eg, influenza) mediated by said antigen.

Fourth Configuration

A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises an immunoglobulin heavy chain locus comprising human gene segment JH6*02, one or more VH gene segments and one or more D gene segments upstream of a constant region; wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof so that the mouse is capable of producing an antibody heavy chain produced by recombination of the human JH6*02 with a D segment and a VH segment.

A non-human vertebrate cell (optionally a mouse cell or a rat cell) whose genome comprises an immunoglobulin heavy chain locus comprising human gene segment JH6*02, one or more VH gene segments and one or more D gene segments upstream of a constant region; wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof for producing (eg, in a subsequent progeny cell) an antibody heavy chain produced by recombination of the human JH6*02 with a D segment and a VH segment.

A heavy chain (eg, comprised by an antibody) isolated from a vertebrate of the invention wherein the heavy chain comprises a HCDR3 of at least 20 amino acids.

A method for producing a heavy chain, VH domain or an antibody specific to a target antigen, the method comprising immunizing a non-human vertebrate according to the invention with the antigen and isolating the heavy chain, VH domain or an antibody specific to a target antigen or a cell producing the heavy chain, VH domain or an antibody, wherein the heavy chain, VH domain or an antibody comprises a HCDR3 that is derived from the recombination of human JH6*02 with a VH gene segment and a D gene segment.

A heavy chain, VH domain or an antibody produced by the method.

A B-cell or hybridoma expressing a heavy chain VH domain that is identical to the VH domain of the heavy chain.

A nucleic acid encoding the VH domain of the heavy chain of claim 22, 23 or 28, or encoding the heavy chain.

A vector (eg, a CHO cell or HEK293 cell vector) comprising the nucleic acid; optionally wherein the vector is in a host cell (eg, a CHO cell or HEK293 cell).

A pharmaceutical composition comprising the antibody, heavy chain or VH domain (eg, comprised by an antibody), together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, heavy chain or antibody).

The antibody, heavy chain or VH domain (eg, comprised by an antibody) as above for use in medicine.

The use of an antibody, heavy chain or VH domain (eg, comprised by an antibody) as above in the manufacture of a medicament for treating and/or preventing a medical condition in a human.

Fifth Configuration

A method of producing an antibody heavy chain, the method comprising
(a) providing an antigen-specific heavy chain variable domain; and
(b) combining the variable domain with a human heavy chain constant region to produce an antibody heavy chain comprising (in N- to C-terminal direction) the variable domain and the constant region;
wherein

the human heavy chain constant region is an IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region.

An antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region . Optionally, the variable domain comprises mouse-pattern AID somatic mutations.

A polypeptide comprising (in N- to C- terminal direction) a leader sequence, a human variable domain that is specific for an antigen and a human constant region that is an IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region wherein (i) the leader sequence is not the native human variable domain leader sequence; and/or (ii) the variable domain comprises mouse AID-pattern somatic mutations and/or mouse Terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations.

A nucleotide sequence encoding (in 5' to 3' direction) a leader sequence and a human antibody heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region; and the leader sequence being operable for expression of the heavy chain and wherein the leader sequence is not the native human variable domain leader sequence.

A nucleotide sequence encoding (in 5' to 3' direction) a promoter and a human antibody heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region ; and the promoter being operable for expression of the heavy chain and wherein the promoter is not the native human promoter.

A vector (eg, a CHO cell or HEK293 cell vector) comprising a IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region nucleotide sequence that is 3' of a cloning site for the insertion of a human antibody heavy chain variable domain nucleotide sequence, such

that upon insertion of such a variable domain sequence the vector comprises (in 5' to 3' direction) a promoter, a leader sequence, the variable domain sequence and the constant region sequence so that the vector is capable of expressing a human antibody heavy chain when present in a host cell.

Sixth Configuration

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 3 human variable region gene segments of the same type (eg, at least 3 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 3 human Vκ1-39 gene segments, at least 3 human D2-2 gene segments or at least 3 human Jκ1 gene segments), wherein at least two of the human gene segments are variants that are not identical to each other.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) cis at the same Ig locus.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *trans* at the same Ig locus; and optionally a third human gene segment of the same type, wherein the third gene segment is *cis* with one of said 2 different gene segments.

A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human variable region gene segments, wherein the plurality comprises at least 2 human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), a first of said different gene segments is provided in the genome of a

first vertebrate of the population, and a second of said different gene segments being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second gene segment.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 3 human variable region gene segments of the same type (eg, at least 3 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 3 human Vκ1-39 gene segments, at least 3 human D2-2 gene segments or at least 3 human Jκ1 gene segments), wherein at least two of the human gene segments are variants that are not identical to each other.

A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *cis* at the same Ig locus.

A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *trans* at the same Ig locus; and

optionally a third human gene segment of the same type, wherein the third gene segment is *cis* with one of said 2 different gene segments.

A method of providing an enhanced human immunoglobulin variable region gene segment repertoire, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human variable region gene segments, wherein the method comprises providing at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein a first of said different gene segments is provided in the genome of a first vertebrate of the population, and a second of said different gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second gene segment.

A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same Ig locus in said genome.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 and 9 to 14 (eg, selected from Table 13 or Table 14) (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence.

A population of non-human vertebrates (eg, mice or rats) comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 and 9 to 14 (eg, selected from Table 13 or Table 14) (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence, wherein the first gene segment is provided in the genome of a first vertebrate of the population, and the second gene segment is provided in the genome of a second vertebrate of the population.

A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 and 9 to 14 (eg, selected from Table 13 or Table 14) (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence.

In one aspect of this configuration, the invention relates to human D gene segment variants as described further below.

In one aspect of this configuration, the invention relates to human V gene segment variants as described further below.

In one aspect of this configuration, the invention relates to human J gene segment variants as described further below.

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); and

Figure 5 (in 4 parts): Alignment of 13 IGHV1-69 variants showing the variable (V) coding region only. Nucleotides that differ from VH1-69 variant *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 6 is a schematic illustrating gene segment diversity and the effect of including variant variants in *cis* according to the invention:-

(a) Situation in a normal person: Recombination on the same chromosome limits combinations of variants, for instance the antibody gene V4-4 can only be recombined within variant 1 to form for instance for instance V4-4-D-J6 or V4-4-D-$J2^A$. Similarly the variant $V4\text{-}4^A$ can't be recombined with either J6 or $J2^A$ from variant 1 and can only be joined with J-genes from variant 2 to form $V4\text{-}4^A$-D-$J6^A$ and $V4\text{-}4^A$-D-J2. V4-4-J2/J6 complexity = 4.

(b) Situation in a transgenic mouse: Only one variant is provided so the genome is limited. V4-4-J6/J2 complexity = 2.

(c) Supra mouse of the invention: The variants are added in *cis* and thus can be recombined in every combination, expanding the repertoire. For instance V4-4 can be combined with J6A, J6, J2A or J2 and similarly V4-4A can be recombined with these same J-genes. The V4-4-J6/J2 complexity = 8, which in this simple example is double that of a person and 4X that of a mouse with a single variant.

Figure 7: Alignment of human JH6*02 variants. Nucleotides that differ from JH6 *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. Accession numbers (eg, J00256) are shown to the left of the IMGT variant name.

Figure 8: Alignment of JH sequences from various species.

Figure 9: Codon Table

Figure 10: BAC database extract

BRIEF DESCRIPTION OF THE TABLES

Table 1:	Human IgH V Polymorphic Variants
Table 2:	Human IgH D Polymorphic Variants
Table 3:	Human IgH J Polymorphic Variants
Table 4:	Human Ig Vk Polymorphic Variants
Table 5:	Human Ig Vλ Polymorphic Variants
Table 6:	Human IgH Jk Polymorphic Variants
Table 7:	Human IgH Jλ Polymorphic Variants
Table 8:	1000 Genomes Project Human Populations
Table 9:	Immunoglobulin Gene Usage in Human Antibody Responses to Infectious Disease Pathogens
Table 10A:	Human IgH JH5 Variant Occurrences
Table 10B:	Non-Synonymous Human IgH JH5 Variants
Table 11A:	Human IgH JH6 Variant Occurrences
Table 11B:	Non-Synonymous Human IgH JH6 Variants
Table 12A:	Human IgH JH2 Variant Occurrences
Table 12B:	Non-Synonymous Human IgH JH2 Variants
Table 13:	Variant Frequency Analyses & Human Population Distributions

Table 14: Frequent Human Variant Distributions

Table 15: Human Gene Segment Usage: Heavy Chain Repertoires From Naïve Non-Human Vertebrates

Table 16: Human Gene Segment Usage: Heavy Chain Repertoires From Immunised Non-Human Vertebrates

Table 17: Human Gene Segment Usage: Heavy Chain Repertoires From Antigen-Specific Hybridomas

Table 18: Sequence Correlation Table

Table 19: Summary Of Function Correlated With Human Gamma Constant Region Sub-Type

Table 20: Gene Segments Prevalent In Few Human Populations

Table 21: Genomic and sequence information

DETAILED DESCRIPTION OF THE INVENTION

A suitable source of JH6*02 and other human DNA sequences for use in the invention 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.

An example source for human V, D and J gene segments according to the invention are Bacterial Artificial Chromosomes (RPCI-11 BACs) obtained from Roswell Park Cancer Institute (RPCI)/Invitrogen. See 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...

"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";
- 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.

Superhuman Immunoglobulin Gene Repertoires

The invention relates to synthetically-extended & ethnically-diverse superhuman immunoglobulin gene repertoires. The human immunoglobulin repertoires are beyond those found in nature (ie, "Superhuman"), for example, they are more diverse than a natural human repertoire or they comprise combinations of human immunoglobulin gene segments from disparate sources in a way that is non-natural. Thus, the repertoires of the invention are "superhuman" immunoglobulin repertoires, and the invention relates to the application of these in transgenic cells and non-human vertebrates for utility in producing chimaeric antibodies (with the possibility of converting these into fully-human, isolated antibodies using recombinant DNA technology). The present invention thus provides for novel and potentially expanded synthetic immunoglobulin diversities, which provides for a pool of diversity from which antibody therapeutic leads (antibody therapeutics and antibody tool reagents) can be selected. This opens up the potential of transgenic human-mouse/rat technologies to the possibility of interrogating different and possibly larger antibody sequence-spaces than has hitherto been possible. To this end, in one embodiment, the invention provides a SUPERHUMAN MOUSE™ (aka SUPRA-MOUSE™) and a SUPERHUMAN RAT™ (aka SUPRA-RAT™)

In developing this thinking, the present inventors have realised the possibility of mining the huge genetics resources now available to the skilled person thanks to efforts such as the HapMap Project, 1000 Genomes Project and sundry other immunoglobulin gene databases (see below for more details). Thus, in some embodiments, the inventors realised the application of these genome sequencing developments in the present invention to generate synthetically-produced and ethnically-diverse artificial immunoglobulin gene repertoires. In one aspect, the inventors realised that such repertoires are useful 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.

The present invention provides libraries, 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.

Variation Analysis

The present inventors have realized methods and antibody loci designs that harness the power of genetic variation analysis. The reference human genome provides a foundation for experimental work and genetic analysis of human samples. The reference human is a compilation of the genomes from a small number of individuals and for any one segment of the genome a high quality single reference genome for one of the two chromosomes is available. Because the reference genome was assembled from a series of very large insert clones, the identity of these clones is known. Accordingly, experimental work with human genomic DNA is usually conducted on the clones from which the reference sequence was derived.

Individual humans differ in their sequence and recently several individuals have had their genomes sequenced, for instance James Watson and Craig Venter. Comparison of the genome sequence of these individuals has revealed differences between their sequences and the reference genome in both coding and non-coding parts of the genome, approximately 1 in 1000 bases are different. Some variants will be significant and contribute to differences between individuals. In extreme cases

these will result in genetic disease. Variation can be implicated in differing responses to drugs administered to human patients, eg, yielding an undesirable lowering of patient response to treatment.

The 1000-Genomes Project has the objective of identifying the most frequent variations in the human genome. This public domain project involved sequencing the genomes of more than 1000 individuals from diverse ethnic groups, comparing these sequences to the reference and assembling a catalogue of variants. This has enabled the annotation of variants in coding regions, but because this sequence wasn't derived from large clones of DNA, the analysis of the sequence from diploid individuals can't discriminate the distribution of the variation between the maternal and paternally inherited chromosomes. Where more than one variant is identified in a protein coding gene, it is not possible to illuminate the distribution of the pattern of variants in each version of the protein. For example, if two variants are detected in different positions of the same protein in an individual, this could have resulted from one copy with two variants and none in the other or each copy could have just one variant. To illuminate the sequence of real proteins, the 1000-Genome Project has sequenced mother-father-child trios. This allows one to "phase" the sequence variants, in other words identify blocks of sequence that are inherited from one or other parent and deconvolute the variants.

To further understand the variation within the 1000-genome set a tool has been developed that can identify the significant variants (defined as non-synonymous amino acid changes) from a region of DNA from the phased data in the 1000-genome data set. This tool has been made available online http://www.1000genomes.org/variation-pattern-finder. This tool allows an investigator to download non-synonymous variation delimited between specific coordinates. The downloaded files are configured as individual genotypes, but the data is phased so the haplotype information and the frequencies of specific halotypes in different populations can be extracted.

The inventors' analysis of the 1000-genome data for the individual human coding segments of the C, V D and J genes from the heavy and light chains reveals that there is significant variation in these segments. Individuals will usually have two different heavy chain alleles and also different light chain alleles at both kappa and lambda loci. The repertoire of antibodies that can be generated

from each allele will be different. This variation will contribute to a better or differing immune response to certain antigens.

Humanized mice that have hitherto been generated with immunoglobulin heavy and light chain loci contain just one type of immunoglobulin locus. Even if these mice contain a full human heavy chain locus, the variation will be less than contained in a typical human because only one set of C, V, D and J genes are available, while a typical human would have two sets.

The inventors have devised ways to improve on this limitation when constructing transgenic non-human vertebrates and cells for human antibody and variable region production *in vivo*.

Mice can be generated with two different loci, each engineered to have a different repertoire of V, D and J segments. This could be in a single mouse or two or more separate mouse strains and would be analogous to or beyond the repertoire found in a normal human. The engineering of such a mouse would go beyond the repertoire described in humanized mice to date which only have one set of alleles.

However, the inventors also realized that this also has limitations, because the different loci would not normally interact to shuffle V, D and J variants between loci. This same limitation is also inherent in a human, thus this system does not utilize the advantage of recombining variants in all combinations.

To go beyond the normal repertoire in humans and take advantage of combinations of C, V, D and J variants the inventors decided, in one embodiment, to provide these on the same chromosome in *cis*. See figure 6. These loci would be characterized by having more than the normal number of J, D or V genes. For example n=6 for the J genes, but including one J6 variant and one J2 variant would increase this to n=8. This could be combined with additional variants for the D and V genes, for example. By detailed analysis of the 1000- Genomes database, the inventors have devised a collection of candidate polymorphic human variant gene segments, eg, JH gene segments (eg, see the examples), that can be built into the design of transgenic heavy and light chain loci in mice for

expressing increasingly diverse and new, synthetic repertoires of human variable regions. Moreover, by utilizing naturally-occurring human variant gene segments, as per embodiments of the invention, this addresses compatibility with human patients since the inventor's analysis has drawn out candidate variants that are naturally conserved and sometimes very prevalent amongst human ethnic populations. Additionally this enables one to tailor the configurations of the invention to provide for antibody-based drugs that better address specific human ethnic populations.

In an example according to any configuration of the invention, loci (and cells and vertebrates comprising these) are provided in which gene segments from different human populations are used. This is desirable to increase antibody gene diversity to better address more diverse human patients. In an example, the gene segments are from first and second different human populations respectively, and thus the second gene segment is found in the second human population, but not so (or rarely) in the first human population. Rarely means, for example, that the gene segment is found in 5, 4, 3, 2, or 1 or zero individuals in the first population in the 1000 Genomes database. For example, the first gene segment may be shown as present in a first population by reference to Table 13 or 14 herein, the second gene segment may be shown as present in the second population by reference to Table 13 and not in the first population. Optionally, the first gene segment may also be shown as being present in the second population by reference to Table 13 or 14.

In any configuration or aspect of the invention, where a V gene segment is used, this may be used optionally with the native leader sequence. For example, use of genomic DNA (eg, from BACs as in the examples) will mean that the native leader will be used for each V gene segment incorporated into the locus and genomes of the invention. In an alternative, the skilled person may wish to inert a non-native leader sequence together with one or more of the V gene segments. Similarly, in any configuration or aspect of the invention, where a V gene segment is used, this may be used optionally with the native 5' UTR sequence. For example, use of genomic DNA (eg, from BACs as in the examples) will mean that the native 5' UTR sequence will be used for each V gene segment incorporated into the locus and genomes of the invention. In an alternative, the skilled person may wish to exclude the native 5' UTR sequence.

The present invention provides, in a first configuration

(a) Superhuman heavy chain gene repertoires

A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) comprising a genome having a superhuman immunoglobulin heavy chain human VH and/or D and/or J gene repertoire.

In one aspect the cell of the invention is an embryonic stem cell. For example, 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 blastocystys 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.

The natural human immunoglobulin gene segment repertoire consists of (see eg, www.imgt.org):-

VH: total-125 ; functional-41

DH: total-27; functional-23

JH: total-8; functional-6

Vk: total-77; functional-38

Jk: total-5; functional-5

V lambda: total-75; functional-31

J lambda: total-7; functional-5

In one embodiment, the vertebrate or cell genome comprises a transgenic immunoglobulin heavy chain locus comprising a plurality of human immunoglobulin VH gene segments, one or more human D gene segments and one or more human J gene segments, wherein the plurality of VH gene segments consists of more than the natural human repertoire of functional VH gene segments; optionally wherein the genome is homozygous for said transgenic heavy chain locus.

In one embodiment of the vertebrate or cell, the VH gene repertoire consists of a plurality of VH gene segments derived from the genome sequence of a first human individual, supplemented with one or more different VH gene segments derived from the genome sequence of a second, different human individual. Optionally the D and J segments are derived from the genome sequence of the first human individual. Optionally the VH gene segments from the genome sequence of the second individual are selected from the VH gene segments listed in Table 1, 13 or 14. In this way, the locus provides a superhuman repertoire of D gene segments.

Optionally the individuals are not related. Individuals are "not related" in the context of any configuration or aspect of the invention, for example, if one of the individuals does not appear in a family tree of the other individual in the same generation or going back one, two, three or four generations. Alternatively, are not related, for example, if they do not share a common ancestor in the present generation or going back one, two, three or four generations.

In one embodiment of the vertebrate or cell, the transgenic locus comprises more than 41 functional human VH gene segment species, and thus more than the natural human functional repertoire. Optionally the locus comprises at least 42, 43, 44, 45, 46, 47, 48, 49 or 50 functional human VH gene segment species (eg, wherein the locus comprises the full functional VH repertoire of said first individual supplemented with one or more VH gene segments derived from the genome sequence of the second human individual and optionally with one or more VH gene segments derived from the genome sequence of a third human individual). In this way, the locus provides a superhuman repertoire of VH gene segments that is useful for generating a novel gene and antibody diversity for

use in therapeutic and tool antibody selection.

In one embodiment of the vertebrate or cell, the transgenic locus comprises a first VH gene segment derived from the genome sequence of the first individual and a second VH gene segment derived from the genome sequence of the second individual, wherein the second VH gene segment is a polymorphic variant of the first VH gene segment. For example, the VH gene segments are polymorphic variants of VH1-69 as illustrated in the examples below. Optionally the locus comprises a further polymorphic variant of the first VH gene segment (eg, a variant derived from the genome sequence of a third human individual). In this way, the locus provides a superhuman repertoire of VH gene segments.

In one embodiment of the vertebrate or cell, the genome (alternatively or additionally to the superhuman VH diversity) comprises a transgenic immunoglobulin heavy chain locus comprising a plurality of human immunoglobulin VH gene segments, a plurality of human D gene segments and one or more human J gene segments, wherein the plurality of D gene segments consists of more than the natural human repertoire of functional D gene segments. Optionally the genome is homozygous for said transgenic heavy chain locus.

In one embodiment of the vertebrate or cell, the D gene repertoire consists of a plurality of D gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more different D gene segments derived from the genome sequence of a (or said) second, different human individual. Optionally the individuals are not related. Optionally the J segments are derived from the genome sequence of the first human individual. Optionally the D gene segments from the genome sequence of the second individual are selected from the D gene segments listed in Table 2, 13 or 14. In this way, the locus provides a superhuman repertoire of D gene segments.

In one embodiment of the vertebrate or cell, the transgenic locus comprises more than 23 functional human D gene segment species; optionally wherein the locus comprises at least 24, 25, 26, 27, 28, 29, 30 or 31 functional human D gene segment species (eg, wherein the locus comprises the full functional D repertoire of said first individual supplemented with one or more D gene segments derived from the genome sequence of the second human individual and optionally with one or more

D gene segments derived from the genome sequence of a third human individual). In this way, the locus provides a superhuman repertoire of D gene segments.

In one embodiment of the vertebrate or cell, the transgenic locus comprises a first D gene segment derived from the genome sequence of the first individual and a second D gene segment derived from the genome sequence of the second individual, wherein the second D gene segment is a polymorphic variant of the first D gene segment. Optionally the locus comprises a further polymorphic variant of the first D gene segment (eg, a variant derived from the genome sequence of a third human individual). In this way, the locus provides a superhuman repertoire of D gene segments.

In one embodiment of the vertebrate or cell (alternatively or additionally to the superhuman VH and/or JH diversity), the genome comprises a (or said) transgenic immunoglobulin heavy chain locus comprising a plurality of human immunoglobulin VH gene segments, one or more human D gene segments and a plurality of human JH gene segments, wherein the plurality of J gene segments consists of more than the natural human repertoire of functional J gene segments; optionally wherein the genome is homozygous for said transgenic heavy chain locus.

In one embodiment of the vertebrate or cell, the JH gene repertoire consists of a plurality of J gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more different J gene segments derived from the genome sequence of a (or said) second, different human individual. Optionally the individuals are not related. Optionally D segments are derived from the genome sequence of the first human individual. Optionally the J gene segments from the genome sequence of the second individual are selected from the J gene segments listed in Table 3 13 or 14. In this way, the locus provides a superhuman repertoire of JH gene segments.

In one embodiment of the vertebrate or cell, the transgenic locus comprises more than 6 functional human JH gene segment segments. Optionally the locus comprises at least 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 functional human JH gene segments (eg, wherein the locus comprises the full functional JH repertoire of said first individual supplemented with one or more JH gene segments derived from

the genome sequence of the second human individual and optionally with one or more JH gene segments derived from the genome sequence of a third human individual). In this way, the locus provides a superhuman repertoire of JH gene segments.

In one embodiment of the vertebrate or cell, the transgenic locus comprises a first JH gene segment derived from the genome sequence of the first individual and a second JH gene segment derived from the genome sequence of the second individual, wherein the second JH gene segment is a polymorphic variant of the first JH gene segment. Optionally the locus comprises a further polymorphic variant of the first JH gene segment (eg, a variant derived from the genome sequence of a third human individual). In this way, the locus provides a superhuman repertoire of JH gene segments.

(b) Superhuman light chain gene repertoires

The first configuration of the invention also provides:-

A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) comprising a genome having a superhuman immunoglobulin light chain human VL gene repertoire. Optionally the vertebrate or cell comprises a heavy chain transgene according to aspect (a) of the first configuration. Thus, superhuman diversity is provided in both the heavy and light chain immunoglobulin gene segments in the cell and vertebrate. For example, the genome of the cell or vertebrate is homozygous for the heavy and light chain transgenes and endogenous antibody expression is inactivated. Such a vertebrate is useful for immunisation with a predetermined antigen to produce one or more selected antibodies that bind the antigen and have human variable regions resulting from recombination within the superhuman gene segment repertoire. This provides potentially for a novel antibody and gene sequence space from which to select therapeutic, prophylactic and tool antibodies.

In one embodiment of aspect (b) of the first configuration, the vertebrate or cell genome comprises

(i) a transgenic immunoglobulin kappa light chain locus comprising a plurality of human

immunoglobulin Vκ gene segments and one or more human J gene segments, wherein the plurality of Vκ gene segments consists of more than the natural human repertoire of functional Vκ gene segments; optionally wherein the genome is homozygous for said transgenic kappa light chain locus; and/or

(ii) a transgenic immunoglobulin lambda light chain locus comprising a plurality of human immunoglobulin Vλ gene segments and one or more human J gene segments, wherein the plurality of Vλ gene segments consists of more than the natural human repertoire of functional Vλ gene segments; optionally wherein the genome is homozygous for said transgenic lambda light chain locus.

In this way, the locus provides a superhuman repertoire of VL gene segments.

In one embodiment of the vertebrate or cell,

(i) the Vκ gene repertoire consists of a plurality of Vκ gene segments derived from the genome sequence of a first human individual, supplemented with one or more Vκ gene segments derived from the genome sequence of a second, different human individual; optionally wherein the individuals are not related; optionally wherein the J segments are derived from the genome sequence of the first human individual; and optionally wherein the Vκ gene segments from the genome sequence of the second individual are selected from the Vκ gene segments listed in Table 4, 13 or 14; and

(i) the Vλ gene repertoire consists of a plurality of Vλ gene segments derived from the genome sequence of a first human individual, supplemented with one or more Vλ gene segments derived from the genome sequence of a second, different human individual; optionally wherein the individuals are not related; optionally wherein the J segments are derived from the genome sequence of the first human individual; and optionally wherein the Vλ gene segments from the genome sequence of the second individual are selected from the Vλ gene segments listed in Table 5, 13 or 14.

In this way, the locus provides a superhuman repertoire of VL gene segments.

In one embodiment of the vertebrate or cell,

-the kappa light transgenic locus comprises more than 38 functional human Vκ gene segment species; optionally wherein the locus comprises at least 39, 40, 41, 42, 43, 44, 45, 46, 47 or 48 functional human Vκ gene segment species (eg, wherein the locus comprises the full functional Vκ repertoire of said first individual supplemented with one or more Vκ gene segments derived from the genome sequence of the second human individual and optionally with one or more Vκ gene segments derived from the genome sequence of a third human individual); and

-the lambda light transgenic locus comprises more than 31 functional human Vλ gene segment species; optionally wherein the locus comprises at least 32, 33, 34, 35, 36, 37, 38, 39, 40 or 41 functional human Vλ gene segment species (eg, wherein the locus comprises the full functional Vλ repertoire of said first individual supplemented with one or more Vλ gene segments derived from the genome sequence of the second human individual and optionally with one or more Vλ gene segments derived from the genome sequence of a third human individual).

In this way, the locus provides a superhuman repertoire of VL gene segments.

In one embodiment of the vertebrate or cell,

-the kappa light transgenic locus comprises a first Vκ gene segment derived from the genome sequence of the first individual and a second Vκ gene segment derived from the genome sequence of the second individual, wherein the second Vκ gene segment is a polymorphic variant of the first Vκ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Vκ gene segment (eg, a variant derived from the genome sequence of a third human individual); and

-the lambda light transgenic locus comprises a first Vλ gene segment derived from the genome sequence of the first individual and a second Vλ gene segment derived from the genome sequence of the second individual, wherein the second Vλ gene segment is a polymorphic variant of the first Vλ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Vλ gene segment (eg, a variant derived from the genome sequence of a third human individual).

In this way, the locus provides a superhuman repertoire of VL gene segments.

In one embodiment of the vertebrate or cell, the genome comprises a (or said) transgenic immunoglobulin light chain locus comprising a plurality of human immunoglobulin VL gene segments and a plurality of human JL gene segments, wherein the plurality of J gene segments consists of more than the natural human repertoire of functional J gene segments; optionally wherein the genome is homozygous for said transgenic heavy chain locus.

In one embodiment of the vertebrate or cell,
(i) the Jκ gene repertoire consists of a plurality of Jκ gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more Jκ gene segments derived from the genome sequence of a (or said) second, different human individual; optionally wherein the individuals are not related; optionally wherein the Vκ segments are derived from the genome sequence of the first human individual; optionally wherein the Jκ gene segments from the genome sequence of the second individual are selected from the Jκ gene segments listed in Table 6, 13 or 14; and
(ii) the Jκ gene repertoire consists of a plurality of Jλ gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more Jλ gene segments derived from the genome sequence of a (or said) second, different human individual; optionally wherein the individuals are not related; optionally wherein the Vλ segments are derived from the genome sequence of the first human individual; optionally wherein the Jλ gene segments from the genome sequence of the second individual are selected from the Jλ gene segments listed in Table 7, 13 or 14.

In this way, the locus provides a superhuman repertoire of JL gene segments.

In one embodiment of the vertebrate or cell,
(i) the transgenic light chain locus comprises more than 5 functional human Jκ gene segment species; optionally wherein the locus comprises at least 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 functional human Jκ gene segment species (eg, wherein the locus comprises the full functional Jκ repertoire of said first individual supplemented with one or more Jκ gene segments derived from the genome sequence of the second human individual and optionally with one or more Jκ gene segments derived

from the genome sequence of a third human individual); and/or

(i) the transgenic light chain locus comprises more than 5 functional human Jλ gene segment species; optionally wherein the locus comprises at least 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 functional human Jλ gene segment species (eg, wherein the locus comprises the full functional Jλ repertoire of said first individual supplemented with one or more Jλ gene segments derived from the genome sequence of the second human individual and optionally with one or more Jλ gene segments derived from the genome sequence of a third human individual).

In this way, the locus provides a superhuman repertoire of JL gene segments.

In one embodiment of the vertebrate or cell,

(i) the kappa light transgenic locus comprises a first Jκ gene segment derived from the genome sequence of the first individual and a second Jκ gene segment derived from the genome sequence of the second individual, wherein the second Jκ gene segment is a polymorphic variant of the first Jκ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Jκ gene segment (eg, a variant derived from the genome sequence of a third human individual); and

(ii) the lambda light transgenic locus comprises a first Jλ gene segment derived from the genome sequence of the first individual and a second Jλ gene segment derived from the genome sequence of the second individual, wherein the second Jκ gene segment is a polymorphic variant of the first Jλ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Jλ gene segment (eg, a variant derived from the genome sequence of a third human individual).

In this way, the locus provides a superhuman repertoire of JL gene segments.

Further aspects of the first configuration are described below.

The present invention provides, in a second configuration

A library of antibody-producing transgenic cells whose genomes collectively encode a repertoire of antibodies, wherein

(a) a first transgenic cell expresses a first antibody having a chain (eg, heavy chain) encoded by a first immunoglobulin gene, the gene comprising a first variable domain nucleotide sequence produced following recombination of a first human unrearranged immunoglobulin gene segment (eg, a VH);
(b) a second transgenic cell expresses a second antibody having a chain (eg, a heavy chain) encoded by a second immunoglobulin gene, the second gene comprising a second variable domain nucleotide sequence produced following recombination of a second human unrearranged immunoglobulin gene segment (eg, a VH), the first and second antibodies being non-identical;
(c) the first and second gene segments are different and derived from the genome sequences of first and second human individuals respectively, wherein the individuals are different; and optionally not related;
(d) wherein the cells are non-human vertebrate (eg, mouse or rat) cells (eg, B-cells or hybridomas).

In one embodiment, the library is provided *in vitro*. In another embodiment, the library is provided *in vivo* by one or a plurality of transgenic non-human vertebrates. For example, the or each vertebrate is according to any aspect of the first configuration of the invention.

In one embodiment, the library encodes an antibody repertoire of from 10 to 10^9 antibodies, for example, 10, 20, 30, 40, 50, 100 or 1000 to 10^8; or 10, 20, 30, 40, 50, 100 or 1000 to 10^7; or 10, 20, 30, 40, 50, 100 or 1000 to 10^6; or 10, 20, 30, 40, 50, 100 or 1000 to 10^5; or 10, 20, 30, 40, 50, 100 or 1000 to 10^4 antibodies. In an example, library encodes an antibody repertoire of at least 10^3, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, or 10^{10} antibodies.

The first variable domain nucleotide sequence is produced following recombination of the first human unrearranged immunoglobulin gene segment with one or more other immunoglobulin gene segments (for example, human immunoglobulin gene segments). For example, where the first gene segment is a VH, the first variable domain nucleotide sequence (a VH domain) is produced following recombination of the VH with a human D and JH segments *in vivo*, optionally with somatic hypermutation, in the first transgenic cell or an ancestor thereof. For example, where the first gene segment is a VL, the first variable domain nucleotide sequence (a VL domain) is produced following recombination of the VL with a human JL segment *in vivo*, optionally with somatic hypermutation, in

the first transgenic cell or an ancestor thereof.

The second variable domain nucleotide sequence is produced following recombination of the second human unrearranged immunoglobulin gene segment with one or more other immunoglobulin gene segments (for example, human immunoglobulin gene segments). For example, where the second gene segment is a VH, the second variable domain nucleotide sequence (a VH domain) is produced following recombination of the VH with a human D and JH segments *in vivo*, optionally with somatic hypermutation, in the second transgenic cell or an ancestor thereof. For example, where the second gene segment is a VL, the second variable domain nucleotide sequence (a VL domain) is produced following recombination of the VL with a human JL segment *in vivo*, optionally with somatic hypermutation, in the second transgenic cell or an ancestor thereof.

The first and second gene segments are respectively derived from genome sequences of first and second human individuals. In one example, such a gene segment is isolated or cloned from a sample cell taken from said individual using standard molecular biology techniques as know to the skilled person. The sequence of the gene segment may be mutated (eg, by the introduction of up to 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 nucleotide changes) prior to use in the present invention. In another example, a gene segment is derived by identifying a candidate human immunoglobulin gene segment in a database (see guidance below) and a nucleotide sequence encoding a gene segment for use in the present invention is made by reference (eg, to be identical or a mutant with up to 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 nucleotide changes to the reference sequence) to the database sequence. The skilled person will be aware of methods of obtaining nucleotide sequences by reference to databases or by obtaining from cellular samples.

In one embodiment of the vertebrate, cell or library of any configuration of the invention, the first and second human individuals are members of first and second ethnic populations respectively, wherein the populations are different. This, therefore, provides for superhuman gene diversity in transgenic loci, cells and vertebrates as per the invention.

Human Populations

Optionally the ethnic populations are selected from those identified in the 1000 Genomes Project of database. In this respect, see Table 8 which provides details of the ethnic populations on which the 1000 Genomes database is based.

N A Rosenberg *et al* (Science 20 December 2002: vol. 298 no. 5602 2342-2343) studied the genetic structure of human populations of differing geographical ancestry. In total, 52 populations were sampled, these being populations with:

African ancestry
(Mbuti Pygmies, Biaka Pygmies, San peoples, and speakers of Niger-Kordofanian languages (Bantu, Yoruba or Mandenka populations),

Eurasian ancestry
(European ancestry (Orcadian, Adygel, Basque, French, Russians, Italians, Sardinian, Tuscan),
Middle Eastern ancestry (Mozabite, Bedouin, Druze, Palestinians),
Central/South Asian ancestry (Balochl, Brahul, Makrani, Sindhi, Pathan, Burusho, Hazara, Uygur, Kalash)),

East Asian ancestry
(Han, Dal, Daur, Hezhen, Lahu, Miao, Oroqen, She, Tujia, Tu, Xibo, Yi, Mongola, Naxi, Cambodian, Japanese, Yakut), Oceanic ancestry (Melanesian, Papuan); or

Americas ancestry
(Karitiana, Surui, Colombian, Maya, Pima).

The International HapMap Project, Nature, 2003 Dec 18;426(6968):789-96, discloses that goal of the HapMap Project: to determine the common patterns of DNA sequence variation in the human genome by determining the genotypes of one million or more sequence variants, their frequencies and the degree of association between them in DNA samples from populations with ancestry from parts of Africa, Asia and Europe. The relevant human populations of differing geographical ancestry include Yoruba, Japanese, Chinese, Northern European and Western European populations. More specifically:-

Utah population with Northern or Western European ancestry (samples collected in 1980 by the Centre d'Etude du Polymorphisme Humain (CEPH));
population with ancestry of Yoruba people from Ibadan, Nigeria;
population with Japanese ancestry; and
population with ancestry of Han Chinese from China.

The authors, citing earlier publications, suggest that ancestral geography is a reasonable basis for sampling human populations.

A suitable sample of human populations from which the populations used in the present invention are selected is as follows:-

(a) *European ancestry*
(b) Northern European ancestry; Western European ancestry; Toscani ancestry; British ancestry, Finnish ancestry or Iberian ancestry.
(c) More specifically, population of Utah residents with Northern and/or Western European ancestry; Toscani population in Italia; British population in England and/or Scotland; Finnish population in Finland; or Iberian population in Spain.

(a) *East Asian ancestry*
(b) Japanese ancestry; Chinese ancestry or Vietnamese ancestry.
(c) More specifically, Japanese population in Toyko, Japan; Han Chinese population in Beijing, China; Chinese Dai population in Xishuangbanna; Kinh population in Ho Chi Minh City, Vietnam; or Chinese population in Denver, Colorado, USA.

(a) *West African ancestry*
(b) Yoruba ancestry; Luhya ancestry; Gambian ancestry; or Malawian ancestry.

(c) More specifically, Yoruba population in Ibadan, Nigeria; Luhya population in Webuye, Kenya; Gambian population in Western Division, The Gambia; or Malawian population in Blantyre, Malawi.

(a) Population of The Americas

(b) Native American ancestry; Afro-Caribbean ancestry; Mexican ancestry; Puerto Rican ancestry; Columbian ancestry; or Peruvian ancestry.

(c) More specifically, population of African Ancestry in Southwest US; population of African American in Jackson, MS; population of African Caribbean in Barbados; population of Mexican Ancestry in Los Angeles, CA; population of Puerto Rican in Puerto Rico; population of Colombian in Medellin, Colombia; or population of Peruvian in Lima, Peru.

(a) South Asian ancestry

(b) Ahom ancestry; Kayadtha ancestry; Reddy ancestry; Maratha; or Punjabi ancestry.

(c) More specifically, Ahom population in the State of Assam, India; Kayadtha population in Calcutta, India; Reddy population in Hyderabad, India; Maratha population in Bombay, India; or Punjabi population in Lahore, Pakistan.

In any configuration of the invention, in one embodiment, each human population is selected from a population marked "(a)" above.

In any configuration of the invention, in another embodiment, each human population is selected from a population marked "(b)" above.

In any configuration of the invention, in another embodiment, each human population is selected from a population marked "(c)" above.

In one embodiment of the library of the vertebrate, cell or library of the invention, the first and second ethnic populations are selected from the group consisting of an ethnic population with

European ancestry, an ethnic population with East Asian, an ethnic population with West African ancestry, an ethnic population with Americas ancestry and an ethnic population with South Asian ancestry.

In one embodiment of the library of the vertebrate, cell or library of the invention, the first and second ethnic populations are selected from the group consisting of an ethnic population with Northern European ancestry; or an ethnic population with Western European ancestry; or an ethnic population with Toscani ancestry; or an ethnic population with British ancestry; or an ethnic population with Icelandic ancestry; or an ethnic population with Finnish ancestry; or an ethnic population with Iberian ancestry; or an ethnic population with Japanese ancestry; or an ethnic population with Chinese ancestry; or an ethnic population Vietnamese ancestry; or an ethnic population with Yoruba ancestry; or an ethnic population with Luhya ancestry; or an ethnic population with Gambian ancestry; or an ethnic population with Malawian ancestry; or an ethnic population with Native American ancestry; or an ethnic population with Afro-Caribbean ancestry; or an ethnic population with Mexican ancestry; or an ethnic population with Puerto Rican ancestry; or an ethnic population with Columbian ancestry; or an ethnic population with Peruvian ancestry; or an ethnic population with Ahom ancestry; or an ethnic population with Kayadtha ancestry; or an ethnic population with Reddy ancestry; or an ethnic population with Maratha; or an ethnic population with Punjabi ancestry.

In one embodiment of any configuration of the vertebrate, cell or library of the invention, the human immunoglobulin gene segment derived from the genome sequence of the second individual is low-frequency (optionally rare) within the second ethnic population. Optionally human immunoglobulin gene segment has a Minor Allele Frequency (MAF) (cumulative frequency) of between 0.5% - 5%, optionally less than 0.5%, in the second human population, eg, as in the 1000 Genomes database.

In one embodiment of any configuration of the vertebrate, cell or library of the invention, the first variable region nucleotide sequence is produced by recombination of the first human immunoglobulin gene segment with a first J gene segment and optionally a first D gene segment, wherein the first human immunoglobulin gene segment is a V gene segment and the V, D and J segments are derived from the first human population, optionally from the genome of one individual

of the first human population.

In one embodiment of the library of the vertebrate, cell or library of the invention, the second variable region nucleotide sequence is produced by recombination of the second human immunoglobulin gene segment with a second J gene segment and optionally a second D gene segment, wherein the second human immunoglobulin gene segment is a V gene segment derived from the second population and the D and/or J segments are derived from the first human population, optionally the D and J gene segments being from the genome of one individual of the first human population.

In one embodiment of the library of the vertebrate, cell or library of the invention, all of the D and J segments that have been recombined with the first and second V gene segments are D and J segments derived from the first human population, optionally the D and J gene segments being from the genome of one individual of the first human population.

In one embodiment of the library, the second human immunoglobulin gene segment is a polymorphic variant of the first human immunoglobulin gene segment; optionally wherein the second gene segment is selected from the group consisting of a gene segment in any of Tables 1 to 7 and 9 to 14 (eg, selected from Table 13 or 14).

In one embodiment of the library, the first and second human immunoglobulin gene segments are both (i) V_H gene segments; (ii) D segments; (iii) J segments (optionally both J_H segments, both J_K segments or both J_λ segments); (iv) constant regions (optionally both a gamma constant region, optionally both a C gamma-1 constant region); (v) CH1 regions; (vi) CH2 regions; or (vii) CH3 regions.

The library is, for example, a naive and optionally has a library size of from 10 or 10^2 to 10^9 cells. For example, from 10, 20, 30, 40, 50, 100 or 1000 to 10^8; or 10, 20, 30, 40, 50, 100 or 1000 to 10^7; or 10, 20, 30, 40, 50, 100 or 1000 to 10^6; or 10, 20, 30, 40, 50, 100 or 1000 to 10^5; or 10, 20, 30, 40, 50, 100 or 1000 to 10^4 cells.

The library has, for example, been selected against a predetermined antigen and optionally has a library size of from 10 or 10^2 to 10^9 cells. For example, from 10, 20, 30, 40, 50, 100 or 1000 to 10^8; or 10, 20, 30, 40, 50, 100 or 1000 to 10^7; or 10, 20, 30, 40, 50, 100 or 1000 to 10^6; or 10, 20, 30, 40, 50, 100 or 1000 to 10^5; or 10, 20, 30, 40, 50, 100 or 1000 to 10^4 cells.

In one embodiment of the library of the invention, said first and second cells are progeny of first and second ancestor non-human vertebrate cells respectively, wherein the first ancestor cell comprises a genome comprising said first human immunoglobulin gene segment; and the second ancestor cell comprises a genome comprising said second human immunoglobulin gene segment.

The invention further provides a library of antibody-producing transgenic cells whose genomes collectively encode a repertoire of antibodies, wherein the library comprises the first and second ancestor cells described above.

The invention further provides a library of hybridoma cells produced by fusion of the library of the invention (eg, a B-cell library) with fusion partner cells and optionally has a library size of from 10 or 10^2 to 10^9 cells. For example, from 10, 20, 30, 40, 50, 100 or 1000 to 10^8; or 10, 20, 30, 40, 50, 100 or 1000 to 10^7; or 10, 20, 30, 40, 50, 100 or 1000 to 10^6; or 10, 20, 30, 40, 50, 100 or 1000 to 10^5; or 10, 20, 30, 40, 50, 100 or 1000 to 10^4 cells. Production of hybridomas is well known to the skilled person. Examples of fusion partners are SP2/0-g14 (obtainable from ECACC), P3X63-Ag8.653 (obtainable from LGC Standards; CRL-1580), NS1 and NS0 cells. PEG fusion or electrofusion can be carried out, as is conventional.

The invention provides, in a third configuration:-

An isolated antibody having

(a) a heavy chain encoded by a nucleotide sequence produced following recombination in a transgenic non-human vertebrate cell of an unrearranged human immunoglobulin V gene segment with a human D and human J segment, optionally with affinity maturation in said cell, wherein one of the gene segments (eg, VH) is derived from the genome of an individual from a first human ethnic population; and the other two gene segments (eg, D and JH) are derived from the genome of an individual from a second (eg, a second and third respectively), different, human ethnic population,

and wherein the antibody comprises heavy chain constant regions (eg, C gamma) of said non-human vertebrate (eg, rodent, mouse or rat heavy chain constant regions); and/or

(b) a light chain encoded by a nucleotide sequence produced following recombination in a transgenic non-human vertebrate cell of an unrearranged human immunoglobulin V gene segment with a human J segment, optionally with affinity maturation in said cell, wherein one of the gene segments (eg, VL) is derived from the genome of an individual from a first human ethnic population (optionally the same as the first population in (a)); and the other gene segment (eg, JL) is derived from the genome of an individual from a second, different, human ethnic population (optionally the same as the second population in (a)), and wherein the antibody comprises light chain constant regions of said non-human vertebrate (eg, rodent, mouse or rat heavy light constant regions);

(c) Optionally wherein each variable domain of the antibody is a human variable domain.

(d) Optionally wherein the heavy chain constant regions are mu- or gamma-type constant regions.

The invention also provides an isolated nucleotide sequence encoding the antibody of the third configuration, optionally wherein the sequence is provided in an antibody expression vector, optionally in a host cell. Suitable vectors are mammalian expression vectors (eg, CHO cell vectors or HEK293 cell vectors), yeast vectors (eg, a vector for expression in *Picchia pastoris*, or a bacterial expression vector, eg, a vector for *E. coli* expression.

The invention also provides a method of producing a human antibody, the method comprising replacing the non-human vertebrate constant regions of the antibody of the third configuration with human antibody constant regions (eg, a C variant disclosed in table 13 or 18). The skilled person will be aware of standard molecular biology techniques to do this. For example, see Harlow, E. & Lane, D. 1998, 5th 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'd edition) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY).

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

The invention also provides a pharmaceutical composition comprising an antibody according to the third configuration, or a human antibody 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 invention also provides an antibody-producing cell (eg, a mammalian cell, eg, CHO or HEK293; a yeast cell, eg, *P pastoris;* a bacterial cell, eg, *E coli*; a B-cell; or a hybridoma) that expresses the second antibody of the third configuration or the isolated antibody of the invention.

The first configuration of the invention also provides:-

A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises a transgenic immunoglobulin locus (eg, a heavy chain locus or a light chain locus), said locus comprising immunoglobulin gene segments according to the first and second human immunoglobulin gene segments (optionally V segments) described above in connection with the third configuration. The gene segments are operably connected upstream of an immunoglobulin constant region; optionally wherein the genome is homozygous for said transgenic immunoglobulin locus.

Optionally the immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or

Optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or

Optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

In this way, a superhuman immunoglobulin gene repertoire is provided in a transgenic non-human vertebrate or vertebrate cell according to the invention.

The first configuration also provides:-

A transgenic non-human vertebrate (eg, a mouse or rat) or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises a transgenic immunoglobulin locus comprising a plurality of human immunoglobulin gene segments operably connected upstream of a non-human vertebrate constant region for the production of a repertoire of chimaeric antibodies, or chimaeric light or heavy chains, having a non-human vertebrate constant region and a human variable region; wherein the transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments, a first (optionally a V segment) of said gene segments and a second (optionally a V segment) of said gene segments being different and derived from the genomes of first and second human individuals respectively, wherein the individuals are different; and optionally not related;

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

In this way, a superhuman immunoglobulin gene repertoire is provided in a transgenic non-human vertebrate or vertebrate cell according to the invention.

The first configuration also provides:-

A transgenic non-human vertebrate (eg, a mouse or rat) or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises first and second transgenic immunoglobulin loci, each locus comprising a plurality of human immunoglobulin gene segments operably connected upstream of a non-human vertebrate constant region for the production of a repertoire of chimaeric antibodies, or chimaeric light or heavy chains, having a non-human vertebrate constant region and a human variable region;

wherein (i) the first transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments, (ii) the second transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments; and (iii) wherein a first (optionally a V) gene segment of said first locus and a second (optionally a V) gene segment of said second gene locus are different and derived from the genomes of first and second human individuals respectively, wherein the individuals are different; and optionally not related;

optionally wherein the first and second loci are on different chromosomes (optionally chromosomes with the same chromosome number) in said genome;

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

In this way, a superhuman immunoglobulin gene repertoire is provided in a transgenic non-human vertebrate or vertebrate cell according to the invention.

In these embodiments of the first configuration, the immunoglobulin gene segments are optionally as described for the third configuration.

In these embodiments of the first configuration, the genome optionally comprises a third immunoglobulin gene segment (optionally a V segment), the third gene segment being derived from a human individual that is different from the individual from which the first (and optionally also the second) gene segment is derived; optionally wherein the first, second and third gene segments are polymorphic variants of a human immunoglobulin gene segment (eg, VH1-69 – see the examples for

further description).

In these embodiments of the first configuration, the genome of the vertebrate or cell is optionally homozygous for the first, second and optional third gene segment, wherein a copy of the first, second and optional third gene segments are provided together on the same chromosome operably connected upstream of a common non-human vertebrate constant region.

For example, each first, second and optional third gene segment is a V gene segment.

In one example, the library of the invention is provided by a collection of non-human vertebrates (optionally a collection of rodents, mice or rats); optionally, wherein a first member of said collection produces said first antibody but not said second antibody, and a second member of the collection produces said second antibody (but optionally not said first antibody). It is therefore contemplated to make non-human vertebrates where different human genomes have been used as a source for building the transgenic loci in the vertebrates. For example, a first vertebrate comprises a transgenic heavy chain locus having gene segments only from a first (and optionally a second) human population or individual; a second vertebrate comprises a transgenic heavy chain locus having gene segments only from a third (and optionally a fourth) human population or individual; and optionally third and more vertebrates can be built similarly based on unique or overlapping human population genomes. However, when provided as a mixed population of transgenic vertebrates, the mixed population provides a collective pool of human immunoglobulin genes that is greater than found in a natural human repertoire. This is useful to extend the antibody and gene sequence space beyond those possible with prior transgenic mice and rats bearing human immunoglobulin loci. As explained above, these have been based on a single human genome.

In one embodiment, the collection of non-human vertebrates bear human immunoglobulin genes confined to human populations that are together grouped under the same population genus “(a)” mentioned above. This provides for a gene repertoire that is biased to producing human antibody variable regions prevalent in the population genus (a) and thus useful for generating antibody therapeutics/prophylactics for members of said population. Alternatively, where gene segments from different human populations are provided in a single transgene according to the invention (not necessarily in a collection of vertebrates), the different human populations are for example together

grouped under the same population genus "(a)" mentioned above.

The invention also provides a repertoire of antibodies expressed from a library of cells according to the invention.

In the non-human vertebrate or cell of any configuration of the invention, the constant region of the transgenic locus is, in one example, an endogenous constant region of said vertebrate (eg, endogenous mouse or rat constant region, eg, from the same strain of mouse or rat as the non-human vertebrate itself).

The invention also provides a method of constructing a cell (eg, an ES cell) according to the invention, the method comprising

(a) identifying functional V and J (and optionally D) gene segments of the genome sequence of a (or said) first human individual;

(b) identifying one or more functional V and/or D and/or J gene segments of the genome sequence of a (or said) second human individual, wherein these additional gene segments are not found in the genome sequence of the first individual;

(c) and constructing a transgenic immunoglobulin locus in the cell, wherein the gene segments of (a) and (b) are provided in the locus operably connected upstream of a constant region.

Optionally the cell comprises a heavy chain locus constructed according to steps (a) to (c) and/or a light chain locus (kappa and/or lambda loci) constructed according to steps (a) to (c).

Optionally the cell is homozygous for the or each transgenic locus; optionally wherein antibody expression from loci endogenous to said cell has been inactivated. This is useful for confining the functional antibody gene repertoire, and thus antibody production, to antibodies bearing human variable regions.

Optionally the gene segment(s) in step (b) are identified from an immunoglobulin gene database selected from the 1000 Genomes, Ensembl, Genbank and IMGT databases.

Optionally the first and second human individuals are members of first and second ethnic populations respectively, wherein the populations are different, optionally wherein the human immunoglobulin gene segment derived from the genome sequence of the second individual is low-frequency (optionally rare) within the second ethnic population.

The invention also provides a method of making a transgenic non-human vertebrate (eg, a mouse or rat), the method comprising

(a) constructing an ES cell (eg, a mouse C57BL/6N, C57BL/6J, 129S5 or 129Sv strain ES cell) by carrying out the method above;

(b) injecting the ES cell into a donor non-human vertebrate blastocyst (eg, a mouse C57BL/6N, C57BL/6J, 129S5 or 129Sv strain blastocyst);

(c) implanting the blastocyst into a foster non-human vertebrate mother (eg, a C57BL/6N, C57BL/6J, 129S5 or 129Sv strain mouse); and

(d) obtaining a child from said mother, wherein the child genome comprises a transgenic immunoglobulin locus.

The invention provides a transgenic non-human vertebrate (eg, a mouse or rat) made by the method or a progeny thereof. The invention also provides a population of such non-human vertebrates.

Microinjection of ES cells into blastocysts and generation of transgenic mice therafter are conventional practices in the state of the art, and the skilled person is aware of techniques useful to effect this. C57BL/6N, C57BL/6J, 129S5 or 129Sv mouse strains and ES cells are readily and publicly available.

The invention also provides a method of isolating an antibody that binds a predetermined antigen (eg, a bacterial or viral pathogen antigen), the method comprising

(a) providing a vertebrate (optionally a mouse or rat) according to 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; and

(f) optionally producing a derivative or variant of the antibody.

This method optionally further comprises after step (e) 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.

Bioinformatics Analysis & Selection of Immunoglobulin Gene Segments

See also the discussion on variation analysis above.

The skilled person will know of sources of human antibody gene sequences, such as IMGT (www.imgt.org), GenBank (www.ncbi.nlm.nih.gov/genbank) 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)

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):-

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.

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.

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.

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.

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.

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

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-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, and hybridoma/cell-culture databases are provided.

1.9. Humanization bY Design (J. Saldanha, 2000; http://people.cryst.bbk.ac.uk/~ubcg07s). 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 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-protocolsantibody-engineering-methods-and-protocols.pdf

As a source of genomic sequence variation data, the skilled person will also be aware of the following available databases and resources (including updates thereof):-

1. HapMap (The International HapMap Consortium. 2003; http://hapmap.ncbi.nlm.nih.gov/index.html.en). The HapMap Project is an international project that aims to compare the genetic sequences of different individuals to identify chromosomal regions containing shared genetic variants. The HapMap www site provides tools to identify chromosomal regions and the variant therein, with options to drill down to population level frequency data.

2. 1000 Genomes (The 1000 Genomes Project Consortium 2010; http://www.1000genomes.org/). This resource provides complete genomic sequence for 2500 unidentified individuals from one of 25 distinct population groups, with the aim of identifying genomic variants of >1%. The site provides the ability to interrogate data utilizing online tools (e.g. 'Variation Pattern Finder') and to download variant data for individual population groups.

3. Japanese SNP Database (H.Haga et al. 2002; http://snp.ims.u-tokyo.ac.jp/index.html). Based on a study identifying 190,562 human genetic variants this site catalogues genomic variants with useful features for searching and summarizing data.

It is possible to identify variants in immunoglobulin genes classed as low-frequency or rare variants that segregate with specific human ethnic populations. For the purpose of this analysis, a low-frequency immunoglobulin gene segment is classed as one with 'Minor Allele Frequency' (MAF) (cumulative frequency) of between 0.5% - 5%, rare variants are those classed as having a MAF of less than 0.5% in a particular human population.

The following bioinformatics protocol is envisaged to indentify human immunoglobulin gene segments for use in the present invention:

(a) Identify one or more genomic regions containing gene segments of interest ('target genomic regions') and calculate the genomic coordinates, using coordinates that match the sequence assembly build used by either the 1000 Genomes project or International HapMap project (or another selected human gene database of choice).

(b) Identify genomic variants mapped to the genomic regions previously identified in (a). Retrieve variant frequencies for variants for each super population and preferably sub-population where such data is available. Tools readily available on the HapMap WWW site and the VWC tools for the 1000Genomes Project are useful for this step.

(c) Filter list of genomic variants from target genomic regions to contain only variants classed as either 'Non-synonymous' single nucleotide polymorphisms (SNPs) or genomic 'insertions or delections' (indels). Filter further to include those that are present in exonic sequences only.

(d) Correlate population frequency data for each of the identified variants for each of the super populations (for example 'European Ancestry', 'East Asian ancestry', 'West African ancestry', 'Americas', and 'South Asian ancestry') to identify those variants that segregate with less than two super-populations. Further correlate all identified variants with each of the sub-populations (for example, 'European ancestry' super-population might be subdivided into groups such as 'CEU – Utah residents with Northern or Western European ancestry', 'TSI Toscani in Italia' and 'British from England and Scotland') and produce a second score for rarity of variants in within a super-population.

(e) Collect one or more gene segments that show segregation to specific sub-populations for construction of synthetic loci according to the invention.

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

By detailed analysis of the 1000 Genomes database, the inventors have devised a collection of candidate polymorphic antibody gene segment variants, eg, human variant JH gene segments (eg, see Example 4), that can be built into the design of transgenic heavy chain loci in mice for expressing increasingly diverse and new, synthetic repertoires of human variable regions. To this end, the invention provides the following embodiments.

The present invention provides in a fourth configuration:-

Selection of Human JH6*02 Variant

Transgenic IgH Loci, Non-Human Vertebrates, Cells & Antibodies Based on Human JH6*02

As explained above, in designing transgenic Ig heavy chain loci the present inventors have considered the huge amount of data available from the 1000 Genomes project (see www.1000genomes.org) that analyses gene distributions amongst many human populations, and in particular data on Ig gene segments. The inventors were also aware of human gene segments disclosed in the IMGT database (see www.imgt.org) and in Ensembl (see www.ensembl.org). The inventors needed to make choices about which human gene segments to include amongst the large number of human gene segments presented in these databases and the other sources of human Ig gene segment information known in the art, including those other databases disclosed herein. When choosing human JH gene segments, the inventors were aware that human JH6 encodes a relatively long amino acid sequence, and thus the inventors thought it desirable to include this for increasing the chances of producing IgH chains with relatively long HCDR3 regions. Antibodies with long HCDR3 (at least 20 amino acids according to IMGT nomenclature) have been shown to neutralise a variety of pathogens effectively 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 (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 *e al*, incorporated herein by reference) (Fig. 1). The usage analysis indicates that these antibodies have the preference to use human JH6 with human D2-2, D3-3 or D2-15 (Brinley, BS *et al*, Figs. 2-5). See also 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"; Breden F *et al*, incorporated herein by reference. Around 20% of all HCDR3 of antibodies use JH6. However, in those antibodies with HCDR3 of more than 24 amino acids, 70% use JH6 (Brinley, BS *et al*, Fig.2).

There is a need in the art for genetically modified non-human vertebrates and cells that can make antibodies and heavy chains that have long human HCDR3s, as well as antibodies, chains and VH domains that can be selected from such vertebrates and cells wherein these can address target epitopes better accessed by long HCDR3s.

The inventors, therefore, chose in this configuration of the invention to include a human JH6 gene segment as a mandatory human gene segment in their IgH locus design. Several different naturally-occurring human JH6 variants are known (eg, JH6*01 to *04 as well as others; IMGT nomenclature). The inventors considered this when deciding upon which human JH6 variant should be included in the transgenic IgH locus design. An alignment of some human JH6 variants is shown in Figure 7 (from www.imgt.org; dashes indicate identical nucleotides; nucleotide changes versus the *01 variant are shown by underlined nucleotides and corresponding amino acid changes are shown by underlined amino acids; Genbank accession numbers (release 185.0) are shown prefixed by J, X, M or A). The inventors used sequencing of human genomic DNA samples, inspection of public IgH DNA databases as well as informed choices on the basis of variant sequences as means to arrive at a rational choice of which JH6 variant to use.

The 1000 Genomes database uses human JH6*03 as the reference sequence, which would be a possible choice for the skilled person wishing to construct a transgenic IgH locus. The inventors noticed (eg, Figure 7 herein) that position 6 in JH6*03 is a tyrosine (Y) encoded by a TAC codon, whereas some other naturally-occurring human variants have a glycine (G) encoded by a GGT codon (the glycine being present as a YYG motif, forming part of a larger YYGXDX motif). To understand the potential significance of this, the inventors carried out analysis of JH sequences from other vertebrate species. The inventors surprisingly noticed that YYG and YYGXDX motifs are conserved across many vertebrate species (see Figures 7 & 8). This suggested to the inventors, therefore, that preservation of this motif might be desirable, which could guide the choice of JH6 variant for use in the present invention.

Another pointer arose when the inventors considered the TAC codon versus the GGT codon encoding Y or G respectively. The inventors considered the impact of these nucleotide sequences on the action of activation-induced cytidine deaminase (AID). The inventors knew that activation-induced cytidine deaminase (AID) is believed to initiate Ig somatic hypermutation (SHM) in a multi-step mechanism and they addressed this activity when rationally designing the locus. AID catalyses the deamination of C to U in DNA, generating mutations at C bases. Cytidines located within hotspot motifs are preferentially deaminated. Certain motifs are hotspots for AID acitivity (DGYW, WRC, WRCY, WRCH, RGYW, AGY, TAC, WGCW, wherein W=A or T, Y=C or T, D=A, G or T, H=A or C or T, and R=A or G). The presence of a TAC codon encoding Y at position 6 in JH6*03 creates AID mutation hotspots (the cytidine being the substrate of AID), these hotspots being the underlined motifs in the previous sentence. The inventors considered the impact of this and in doing so they considered possible mutants created by AID activity at the cytidine. Reference is made to Figure 9. The inventors noticed that a mutation at the third base of the TAC codon would yield 3 possible outcomes: Y, stop or stop. Thus, out of the three stop codons possible in the genetic code (the other being encoded by TGA – see Figure 9), two of them would be provided by mutation of the cytidine in the TAC codon encoding position 6 in JH6*03. The inventors, therefore, considered that this might increase the chances of non-productive IgH variable region production in transgenic loci based on JH6*03. Moreover, the inventors noticed that provision of a GGT codon instead (as per the other human JH6 variants) seemed preferable since mutation of the third base would never yield a stop codon (see Figure 9), and furthermore would retain coding, and thus conservation, of glycine at position 6, which the inventors also noticed was is in the YYG and YYGXDX motifs conserved across species.

Having decided against using JH6*03, the inventors needed to make a choice from other possible human variants. The MDV motif is at the C-terminus of HCDR3 based on human JH6, the adjacent framework 4 (FW4) starting with the WGQ motif (with reference to the sequence shown encoded by JH6*01; Figure 7). In making their choices for locus design, the inventors wished to maximise conservation of this HCDR3/FW4 junction in product IgH chains and antibodies including these. The inventors believed this to be desirable for heavy chain variable domain functionality and conformation. The inventors thought that this might in some cases be desirable to minimise immunogenicity (suitable for human pharmaceutical use). Consistent with these considerations, the inventors wanted to make a choice that would minimise mutation around the HCDR3/FW4 junction as a result of SHM *in vivo* to conserve junction configuration. See Rogozin & Diaz; "Cutting Edge: DGYW/WRCH Is a Better Predictor of Mutability at G:C Bases in Ig Hypermutation Than the Widely Accepted RGYW/WRCY Motif and Probably Reflects a Two-Step Activation-Induced Cytidine Deaminase-Triggered Process"; Journal of Immunology; March 15, 2004 vol. 172 no. 6 3382-3384. An example of a DGYW motif is GGCA. The inventors had this in mind when analysing the variant sequences.

With these considerations in mind, the inventors decided specifically to use human JH6*02 as the mandatory human JH6 for their IgH locus design. JH6*01 was rejected as the mandatory JH6 gene segment since the nucleotide sequence GGG CAA (encoding G and Q) contains a GGCA motif which is an AID recognition hotspot. The inventors realised that JH6*04 also contains such a motif due to the presence of the sequence GGC AAA encoding G and K (positions 11 and 12 respectively). The inventors also realised that the *02 variant has a C instead of a G that is in the *01 variant, the C desirably being a synonymous change (ie, not changing the encoded amino acid sequence around the CDR3/FW4 junction) and also this does not provide a GGCA AID hotspot motif. The inventors, therefore, decided that the mandatory JH6 should have this C base and this too pointed them to using the human JH6*02 variant.

In one example of any configuration of the invention herein, the only JH6 species included in the locus or genome is human JH6*02.

The inventors obtained 9 anonymised DNA samples from cheek swabs of 9 consenting human adults. Sequencing was performed on IgH locus DNA to confirm natural JH6 variant usage. It was found that the genome of all 9 humans contained a JH6*02 variant gene segment. In 7 out of the 9 humans, the genome was homozygous for JH6*02 (ie, each chromosome 14 had JH6*02 as its JH6 gene segment in the IgH locus). The inventors also inspected the publicly-available sequence information from the genomes of well-known scientists Craig Venter and Jim Watson. Both of these genomes contain JH6*02 too. This indicated to the inventors that this variant is common in humans.

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

(i) Containing the YYG and YYGXDX motifs that is conserved across several vertebrate species;

(ii) Provision of one less TAC codon (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.

This rationale was tested by the inventors in laboratory examples, in order to see if human JH6*02 could desirably participate in antibody gene segment recombination and heavy chain production in a foreign (non-human vertebrate) setting, and moreover to assess if long HCDR3s based on human JH6*02 could be produced *in vivo* (in naïve and immunised settings) in such non-human systems. It was noted that in some non-human settings, such as a mouse, the YYG and YYGXDX motifs are not conserved, and thus the inventors decided that it was important to test whether or not JH6*02 (having the YYG and YYGXDX motifs) could function properly in such a foreign setting to participate in VDJ recombination and selection against antigen.

Thus, as explained further in the examples, the inventors constructed transgenic JH6*02-containing IgH loci in ES cells, generated transgenic non-human vertebrates from the ES cells (both naïve and immunised with a range of different target antigen types), isolated antibodies and heavy chain sequences based on JH6*02 as well as B-cells expressing these and made hybridomas expressing antigen-specific antibodies that are based on the chosen JH6*02 variant. The inventors found that the JH6*02 variant was extensively used and could contribute to the production of HCDR3 of at least

20 amino acids in many different heavy chains (including antigen-specific heavy chains). The chosen variant was preferably used over other JH gene segments in all settings (naïve, immunised and antigen-specific) for the production of HCDR3 of at least 20 amino acids.

Thus, the present invention provides an IgH locus including human JH6*02 (IMGT nomenclature) as a mandatory JH gene segment. In one embodiment, the locus comprises non-human vertebrate (eg, mouse or rat) constant region gene segments downstream (ie, 3' of) the human JH6*02; and one or more VH gene segments (eg, a plurality of human VH gene segments) and one or more D gene segments (eg, a plurality of human D gene segments) upstream of (ie, 5' of) the human JH6*02. For example, the locus is comprised by a vector (eg, a DNA vector, eg, a yeast artificial chromosome (YAC), BAC or PAC). Such a vector (eg, YAC) can be introduced into a non-human vertebrate (eg, mouse or rat) cell using standard techniques (eg, pronuclear injection) so that the locus is integrated into the cell genome for expression of IgH chains comprising at least one chain whose variable domain is a product of the recombination of human JH6*02 with a VH and a D gene segment.

In another example, the locus (eg, with a completely human, rat or mouse constant region, or a human/mouse chimaeric constant region) can be provided in the genome of a non-human vertebrate (eg, mouse or rat) cell. For example, the cell is an ES cell or an antibody-producing cell (eg, an isolated B-cell, an iPS cell or a hybridoma).

In another example, the invention provides a non-human vertebrate (eg, a mouse or a rat) comprising an IgH locus of the invention which comprises a human JH6*02 gene segment, wherein the locus can express an IgH chain whose variable domain is a product of the recombination of human JH6*02 with a VH and a D gene segment. As shown in the examples, the inventors have successfully produced such mice which produce such IgH chains with VH domains based on human JH6*02. The inventors isolated and sequenced IgH chains from the mice before (naïve) and after (immunised) exposure to a range of target antigens and confirmed by comparison to IMGT IgH gene segment sequences that the isolated chains (and antibodies containing these) were produced based on JH6*02. Such chains were found in naïve mice, as well as in antigen-specific antibodies from immunised mice. B-cells were isolated from immunised mice, wherein the B-cells express antibodies based on JH6*02 and hybridomas were generated from the B-cells, the hybridomas expressing antigen-specific antibodies based on JH6*02. The inventors, therefore, provided the locus,

vertebrate, cell and hybridoma of the invention based on the use of human JH6*02 and showed that antibodies based on JH6*02 and B-cells expressing these can be successfully produced and isolated following immunisation of the vertebrates, corresponding hybridomas being a good source of antibodies whose VH domains are based on JH6*02, eg for administration to a patient, eg, for human medicine. Furthermore, it was found possible to produce and isolated antigen-specific antibodies whose VH domains are based on JH6*02 and which had a relatively long HCDR3 (eg, 20 amino acids).

Thus, the present invention provides embodiments as in the following clauses:-

1. A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises an immunoglobulin heavy chain locus comprising human gene segment JH6*02, one or more VH gene segments and one or more D gene segments upstream of a constant region; wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof so that the mouse is capable of producing an antibody heavy chain produced by recombination of the human JH6*02 with a D segment and a VH segment.

   In another example, the invention provides
   A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises an immunoglobulin heavy chain locus comprising one, more or all of human IGHV gene segments selected from V3-21, V3-13, V3-7, V6-1, V1-8, V1-2, V7-4-1, V1-3, V1-18, V4-4, V3-9, V3-23, V3-11 and V3-20 (eg, one, more or all of V3-21*03, V3-13*01, V3-7*01, V6-1*01, V1-8*01, V1-2*02, V7-4-1*01, V1-3*01, V1-18*01, V4-4*01, V3-9*01 and V3-23*04). These segments were found in naive repertoires to be productive to produce HCDR3s of at least 20 amino acids in length. In an embodiment, the locus comprises a human JH6, eg, JH6*02.

   The invention also provides a HCDR3, VH domain, antibody heavy chain or antibody having a HCDR3 size of at least 20 amino acids. Optionally, the HCDR3 or VH domain (or VH domain of the heavy chain or antibody) comprises mouse AID-pattern somatic hypermutations and/or mouse dTd-pattern mutations. This 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.

These segments were found in naive repertoires to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

In an example, the vertebrate is naïve. In another embodiment, the vertebrate instead is immunised with a target antigen.

In an example, the vertebrate or cell mentioned below is capable of so producing an antibody heavy chain upon immunisation with a target antigen. In an example, the vertebrate is an immunised vertebrate that produces antibody heavy chains specific for a target antigen and wherein the variable domains of the heavy chains are the product of recombination between a VH, D and JH6*02. For example, the D is selected from human D3-3, D2-15, D3-9; D4-17; D3-10; D2-2; D5-24; D6-19; D3-22; D6-13; D5-12; D1-26; D1-20; D5-18; D3-16; D2-21; D1-14; D7-27; D1-1; D6-25; D2-14 and D4-23 (eg, selected from D3-9*01; D4-17*01; D3-10*01; D2-2*02; D5-24*01; D6-19*01; D3-22*01; D6-13*01; D5-12*01; D1-26*01; D1-20*01; D5-18*01; D3-16*02; D2-21*02; D1-14*01; D7-27*02; D1-1*01; D6-25*01; D2-15*01; and D4-23*01). For example, the D is human D3-9 or D3-10. In an example, the HCDR3 length is at least 20 amino acids (eg, 20, 21, 23 or 24).

In an example of the vertebrate or cell, the genome comprises additional human JH gene segments (eg, JH2, 3, 4 and 5 gene segments).

In an example of the vertebrate or cell, 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 Eμ and Sμ between the J gene segment(s) and the constant region as is known by the skilled person. In one example, a mouse Eμ and Sμ is included in the heavy chain locus between the JH6*02 and the constant region (ie, in 5' to 3' order the locus comprises the JH6*02, Eμ and Sμ and constant region). In an example, the Eμ and Sμ are Eμ and Sμ 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 Eμ and Sμ are Eμ and Sμ 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 Cμ; or Cμ and gamma constant regions thereof) is a constant region (or Cμ; or Cμ and gamma constant regions thereof) is of a genome described in the paragraph immediately above.

A suitable source of JH6*02 and other human DNA sequences 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.

2. The vertebrate of clause 1, wherein the vertebrate has been immunised with a target antigen and wherein the variable domain of the heavy chain is the product of recombination between a VH, D and JH6*02 and wherein the HCDR3 length is at least 20 amino acids (eg, 20, 21, 23 or 24).

   Optionally, the immunised vertebrate produces an antibody heavy chain specific for a target antigen and wherein the variable domain of the heavy chain is the product of recombination between a VH, D and JH6*02 and wherein the HCDR3 length is at least 20 amino acids (eg, 20, 21, 23 or 24).

3. A non-human vertebrate cell (optionally a mouse cell or a rat cell) whose genome comprises an immunoglobulin heavy chain locus comprising human gene segment JH6*02, one or more VH gene segments and one or more D gene segments upstream of a constant region; wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof for

producing (eg, in a subsequent progeny cell) an antibody heavy chain produced by recombination of the human JH6*02 with a D segment and a VH segment.

4. The cell of clause 3, which is an ES cell capable of differentiation into a progeny antibody-producing cell that expresses said heavy chain.

5. The vertebrate or cell of any preceding clause, wherein the heavy chain locus comprises a human JH6*02 recombination signal sequence (RSS) operably connected 5' to the JH6*02 gene segment.

   For example, the native RSS-JH6*02 sequence can be used to advantageously maintain the natural pairing between RSS and theis JH gene segment. In this respect, the following sequence is used:-

   ggtttttgt**ggggtgaggatggacattctg***c*cattgtgattactactactactacggtatggacgtctggggccaagggaccacggtcaccg tctcctcag (SEQ ID NO: 238)

RSSs have a common architecture: 9mer (eg, first underlined sequence above) followed by a 22bp spacer and then a 7mer (eg, second underlined sequence above). Spacers are 23bp +/- 1 normally, while the 9 and 7mer are more conserved.

6. The vertebrate or cell of clause 5, wherein the RSS is SEQ ID NO: 238 or a sequence having an identical 9mer and 7mer sequence flanking a sequence that is at least 70% identical to the 22mer sequence of SEQ ID NO: 238.

7. The vertebrate or cell of clause 6, wherein the RSS and JH6*02 are provided as SEQ ID NO: 237.

8. The vertebrate or cell of any preceding clause, wherein the JH6*02 is the only JH6-type gene segment in the genome.

9. The vertebrate or cell of any preceding clause, wherein the JH6*02 is the closest JH gene segment to the constant region in the locus.

10. The vertebrate or cell of any preceding clause, wherein the locus comprises one, more or all human D gene segments D3-9; D4-17; D3-10; D2-2; D5-24; D6-19; D3-22; D6-13; D5-12; D1-26; D1-20; D5-18; D3-16; D2-21; D1-14; D7-27; D1-1; D6-25; D2-14; and D4-23.

    For example, the locus comprises one, more or all of human D gene segments D3-9*01; D4-17*01; D3-10*01; D2-2*02; D5-24*01; D6-19*01; D3-22*01; D6-13*01; D5-12*01; D1-26*01; D1-20*01; D5-18*01; D3-16*02; D2-21*02; D1-14*01; D7-27*02; D1-1*01; D6-25*01; D2-15*01; and D4-23*01.

11. The vertebrate or cell of clause 10, wherein the locus comprises one, more or all human D gene segments D3-9, D3-10, D6-19, D4-17, D6-13, D3-22, D2-2, D2-25 and D3-3.

    These D segments were found to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

    In an example, the locus comprises one, more or all human D gene segments D3-9, D3-10, D6-19, D4-17, D6-13 and D3-22 (for example one, more or all of D3-9*01, D3-10*01, D6-19*01, D4-17*01, D6-13*01 and D3-22*01). These D segments were found in naïve repertoires to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

    In an example, the locus comprises one, more or all human D gene segments D3-10, D6-19 and D1-26 (for example, one, more or all of D3-10*01, D6-19*01 and D1-26*01). These D segments were found in immunised repertoires to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

    In an example, the locus comprises one, more or all human D gene segments D3-9 and D3-10 (for example, one, more or all of D3-9*01 and D3-10*01). These D segments were found in antigen-specific repertoires to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

12. The vertebrate or cell of any preceding clause, wherein the locus comprises a plurality of human D gene segments and the JH6*02 is in human germline configuration with respect to the 3'-most human D gene segment (or all of the human D segments comprised by the locus).

In an example, the 3'-most D gene segment is D7-27. In an example, the locus comprises all of human D gene segments from D1-1 to D7-27 as present in a germline human IgH locus (eg, as shown in the IMGT database).

Alternatively or additionally, the JH6*02 is in human germline configuration with respect to one, more or all of the Eμ, Sμ and constant region (eg, Cμ).

13. The vertebrate or cell of any preceding clause, wherein the locus comprises one, more or all of IGHV gene segments selected from V3-21, V3-13, V3-7, V6-1, V1-8, V1-2, V7-4-1, V1-3, V1-18, V4-4, V3-9, V3-23, V3-11 and V3-20.

In an example, the locus comprises one, more or all human IGHV gene segments V3-21, V3-13, V3-7, V6-1, V1-8, V1-2, V7-4-1, V1-3, V1-18, V4-4, V3-9, V3-23 (for example, one, more or all of V3-21*03, V3-13*01, V3-7*01, V6-1*01, V1-8*01, V1-2*02, V7-4-1*01, V1-3*01, V1-18*01, V4-4*01, V3-9*01 and V3-23*04). These segments were found in naive repertoires to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

In an example, the locus comprises one, more or all human IGHV gene segments V3-7, V3-11 and V4-4 (for example, one, more or all of V3-7*01, V3-11*01 and V4-4*02). These segments were found in immunised repertoires to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

In an example, the locus comprises one, more or all human IGHV gene segments V4-4, V1-8, V3-9, V3-11 and V3-20 (for example, one, more or all of V4-4*02, V1-8*01, V3-9*01, V3-11*01 and V3-20 (eg, *d01). These segments were found in antigen-specific repertoires to be productive in recombination with human JH6*02 to produce HCDR3s of at least 20 amino acids in length.

14. The vertebrate or cell of any preceding clause, wherein the locus comprises one, more or all of human D3-9*01, D3-10*01, D6-19*01, D6-13*01, D1-26*01, IGHV1-8*01, IGHV4-61*01, IGHV6-1*01, IGHV4-4*02, IGHV1-3*01, IGHV3-66*03, IGHV3-7*01 and IGHV3-9*01.

These are gene segments that very frequently combine with JH6*02 to produce productive heavy chains and antibodies.

For example, the locus comprises one, more or all of human IGHV1-8*01, D3-9*01 and D3-10*01. These gene segments were productive with JH6*02 to produce HCDR3s of at least 20 amino acids in more than 10 antibodies.

15. An antibody-producing cell (eg, a B-cell) that is a progeny of the cell of any one of clauses 3 to 14, wherein the antibody-producing cell comprises a heavy chain locus comprising a rearranged variable region produced by recombination of human JH6*02 with a D segment and a VH segment (eg, JH6*02 with human VH3-11 (eg, VH3-11*01) and D3-9; VH3-20 (eg, VH3-20*01) and D3-10; VH4-4 (eg, VH4-4*02) and D3-10; VH3-9 (eg, VH3-9*01) and D3-10; or VH1-8 (eg, VH1-8*01) and D310).

    Such a variable region would be the product of *in vivo* somatic hypermutation in a non-hman vertebrate or cell of the invention.

16. The cell of clause 15, which is a B-cell or hybridoma that expresses a target antigen-specific antibody comprising a heavy chain that comprises a rearranged variable region produced by recombination of human JH6*02 with a D segment and a VH segment (eg, JH6*02 with human VH3-11 (eg, VH3-11*01) and D3-9; VH3-20 (eg, VH3-20*01) and D3-10; VH4-4 (eg, VH4-4*02) and D3-10; VH3-9 (eg, VH3-9*01) and D3-10; or VH1-8 (eg, VH1-8*01) and D310).

    Such a variable region would be the product of *in vivo* somatic hypermutation in a non-hman vertebrate or cell of the invention

17. The vertebrate or cell of any preceding clause, wherein the antibody heavy chain specifically binds a target antigen.

18. The vertebrate or cell of any preceding clause, wherein the antibody heavy chain has a HCDR3 length of at least 20 amino acids.

    Optionally, the HCDR3 length is at least 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 amino acids. Additionally, in one example the length is no more than 35, 34, 33, 32 or 31 amino acids. For example, the HCDR3 length is 20, 21, 22, 23 or 24 amino acids.

19. The vertebrate or cell of any preceding clause, wherein the antibody heavy chain is a product of the recombination of JH6*02 with a human VH gene segment recited in clause 13 or 14 and/or a D gene segment recited in clause 10, 11 or 14.

20. The vertebrate or cell of any preceding clause, wherein all endogenous non-human vertebrate heavy chain variable region gene segments have been inactivated in the genome (Eeg, by gene segment deletion or inversion).

21. The vertebrate or cell of any preceding clause, wherein the genome is homozygous for said heavy chain locus.

22. A heavy chain (eg, comprised by an antibody) isolated from a vertebrate of any one of clauses 1, 2, 5 to 14 and 17 to 21 wherein the heavy chain comprises a HCDR3 of at least 20 amino acids.

23. The heavy chain of clause 22, wherein the HCDR3 is the product of recombination of human JH6*02 with a human VH gene segment recited in clause 13 or 14 and/or a D gene segment recited in clause 10, 11 or 14.

    In an example, the heavy chain is chimaeric where the C region is non-human. In an example, the heavy chain is human where the C region is human.

24. A heavy chain (eg, comprised by an antibody) whose VH variable domain is identical to the VH variable domain of the heavy chain of clause 22 or 23, and which comprises a human constant region or a human-mouse chimaeric constant region (eg, CH1 is human and the other constant domains are mouse).

25. The heavy chain of clause 22, 23 or 24, whose VH variable domain is specific for a target antigen.

26. A method for producing a heavy chain, VH domain or an antibody specific to a target antigen, the method comprising immunizing a non-human vertebrate according to any one of clauses 1, 2, 5 to 14 and 17 to 21 with the antigen and isolating the heavy chain, VH domain or an antibody specific to a target antigen or a cell producing the heavy chain, VH domain or an antibody, wherein the heavy chain, VH domain or an antibody comprises a HCDR3 that is derived from the recombination of human JH6*02 with a VH gene segment and a D gene segment.

27. A method for producing a human heavy chain or antibody comprising carrying out the method of clause 26, wherein the constant region of the locus is a non-human vertebrate (eg, mouse or rat) constant region, and then replacing the non-human constant region of the isolated heavy chain or antibody with a human constant region (eg, by engineering of the nucleic acid encoding the antibody).

28. A heavy chain, VH domain or an antibody produced by the method of clause 26 or 27.

    Optionally the HCDR3 length is at least 20 amino acids as herein described.

29. A B-cell or hybridoma expressing a heavy chain VH domain that is identical to the VH domain of the heavy chain of clause 22, 23 or 28.

30. A nucleic acid encoding the VH domain of the heavy chain of clause 22, 23 or 28, or encoding the heavy chain of clause 22, 23, 24, 25 or 28.

31. A vector (eg, a CHO cell or HEK293 cell vector) comprising the nucleic acid of clause 30; optionally wherein the vector is in a host cell (eg, a CHO cell or HEK293 cell).

32. A pharmaceutical composition comprising the antibody, heavy chain or VH domain (eg, comprised by an antibody) of any one of clauses 22 to 25 and 28, together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, heavy chain or antibody).

33. The antibody, heavy chain or VH domain (eg, comprised by an antibody) of any one of clauses 22 to 25 and 28 for use in medicine (eg, human medicine).

For example, the locus comprises 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 locus comprises 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 locus comprises the following human JH gene segment variants

IGHJ2*01
IGHJ3*02

IGHJ4*02
IGHJ5*02 and
IGHJ6*02

For example, the locus comprises 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

The present invention provides in a fifth configuration:-

Constant Regions Tailored to Human Use & Antibody Humanisation

Additional rational design and bioinformatics has led the inventors to realise that specific human constant region variants are conserved across many diverse human populations. The inventors realised that this opens up the possibility of making a choice to humanise antibodies, chains and variable domains by using such specific constant regions in products, rather than arbitrarily choosing the human constant region (or a synthetic version of a human constant region). This aspect of the invention also enables one to tailor antibody-based drugs to specific human ethnic populations, thereby more closely matching drug to patient (and thus disease setting) than has hitherto been performed. It can be a problem in the state of the art that antibodies are humanised with an arbitrary choice of human constant region (presumably derived from one (often unknown) ethnic population or non-naturally occurring) that does not function as well in patients of a different human ethnic population. This is important, since the constant region has the major role in providing antibody effector functions, eg, for antibody recycling, cellular and complement recruitment and for cell killing.

As discussed further in WO2011066501, human IgG sub-types IgG1, IgG2, gG3 and IgG4 exhibit differential capacity to recruit immune functions, such as antibody-dependent cellular cytotoxicity (ADCC, e.g., IgG1 and IgG3), antibody-dependent cellular phagocytosis (ADCP, e.g., IgG1, IgG2, IgG3 and IgG4), and complement dependent cytotoxicity (CDC, e.g., IgG1, IgG3). Sub-type-specific engagement of such immune functions is based on selectivity for Fc receptors on distinct immune cells and the ability to bind C1q and activate the assembly of a membrane attack complex (MAC). Among the various types, relative affinity for FcΥ receptors (e.g., FcΥRI, FcΥRIIa/b/c, FcΥRIIIa/b) is high for IgG1 and IgG3, however, there is minimal affinity for IgG2 (restricted to the FcΥRIIa 131H polymorphism), and IgG4 only has measurable affinity for FcΥRI. Using comparative sequence analysis and co-crystal structures, the key contact residues for receptor binding have been mapped to the amino acid residues spanning the lower hinge and CH2 region. Using standard protein engineering techniques, some success in enhancing or reducing the affinity of an antibody preparation for Fc receptors and the C1q component of complement has been achieved.

Among the isotypes, IgG2 is least capable of binding the family of Fc receptors. Using IgG2 as the starting point, efforts have been made to find a mutant with diminished effector functions but which retains FcRn binding, prolonged stability, and low immunogenicity. Improved mutants of this nature may provide improved antibody therapeutics with retained safety. Human IgG1 therapeutic

antibodies that bind to cell surface targets are able to engage effector cells that may mediate cell lysis of the target cell by antibody-dependent cellular cytotoxicity (ADCC) or complement dependent cytotoxicity (CDC). These mechanisms occur through interaction of the CH2 region of the antibody Fc domain to FcγR receptors on immune effector cells or with C1q, the first component of the complement cascade. Table 19 shows the activities of different human gamma sub-types. The skilled person may choose accordingly to promote or dampen-down activity depending upon the disease setting in humans of interest. For example, use of a human gamma-1 constant region is desirable when one wishes to isolated totally human heavy chains and antibodies that have relatively high complement activation activity by the classical pathway and FcϒR1 recognition in human patients. See also Mol Immunol. 2003 Dec;40(9):585-93; "Differential binding to human Fcgamma RIIa and Fcgamma RIIb receptors by human IgG wild type and mutant antibodies"; Armour KL *et al*, which is incorporated herein by reference.

IgG2 constant regions are well suited to producing antibodies and heavy chains according to the invention for binding to cytokines or soluble targets in humans, since IgG2 is essentially FcϒRI,III-silent, FcϒRIIa-active and has little Complement activity.

IgG1 constant regions have wide utility for human therapeutics, since IgG1 antibodies and heavy chains are FcϒRI,II,III- active and have complement activity. This can be enhanced by using a human gamma-1 constant region that has been activated by engineering as is known in the art.

The work of the inventors has therefore identified a collection of human constant region of different isotypes from which an informed choice can be made when humanising chimaeric antibody chains (or conjugating V domains, such as dAbs or *Camelid* VHH, to constant regions). The collection was identified on the basis of bioinformatics analysis of the 1000 Genomes database, the inventors selecting constant region variants that are frequently occurring across several human ethnic populations, as well as those that appear with relatively high frequency within individual populations (as assessed by the number of individuals whose genomes comprise the variant). By sorting through the myriad possible sequences on this basis, the inventors have provided a collection of human constant region variants that are naturally-occuring and which can be used when rationally designing

antibodies, heavy chains and other antibody-based formats that bear a human constant region. In particular, this is useful when humanising chimaeric heavy chains to produce totally human chains in which both the variable and constant regions are human. This is useful for compatibility with human patients receiving antibody-based drugs.

To this end, the invention provides the following aspects:-

1. A method of producing an antibody heavy chain, the method comprising
   (a) providing an antigen-specific heavy chain variable domain (eg, VH (such as a human VH or dAb) or VHH or a humanised heavy chain variable domain); and
   (b) combining the variable domain with a human heavy chain constant region to produce an antibody heavy chain comprising (in N- to C-terminal direction) the variable domain and the constant region;
   wherein
   the human heavy chain constant region is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region.

   Step (b) can be carried out, eg, using recombinant DNA technology using the corresponding nucleotide sequences.

   For the constant region according to any aspect of this configuration, either genomic DNA or equivalent (ie, having introns and exons and optionally also 5' UTR sequences, eg, with native or a non-native leader sequence) can be used for the constant region. For example, any of the "GENOMIC" sequences disclosed as SEQ ID NO: 365 onwards herein. Alternatively, an intronless sequence can be used, for example any of the "CDS" sequences disclosed as SEQ ID NO: 365 onwards herein (eg, with native or a non-native leader sequence).

   Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHAref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHA1a constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHA2a constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHA2b constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is IGHG1ref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHG2ref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHG2a constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHG3ref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHG3a constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHG3b constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHG4ref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHG4a constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHDref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHEref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHMref constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHMa constant region.

Optionally for any aspect of this configuration of the invention, the human heavy chain constant region is an IGHMb constant region.

Optionally, a derivative (eg, a mutant or conjugate) of the heavy chain or an antibody containing the heavy chain is produced. For example, a toxic payload can be conjugated (eg, for oncology applications). For example, one or more mutations can be introduced, as is known in the art, to inactivate or enhance Fc effector function.

2. The method of aspect 1, wherein the variable domain is a human variable domain.

A human variable domain is, for example, the product of recombination in a transgenic non-human vertebrate of human VH, D and JH gene segments. Alternatively, the variable domain is identified using *in vitro* display technology from a human VH library, eg, using phage display, ribosome display or yeast display, as is known in the art.

In another embodiment, the variable domain is a humanised variable domain, eg, comprising human frameworks with non-human (eg, mouse or rat) CDRs). Humanisation technology is conventional in the art, and will be readily known to the skilled person.

3. The method of any preceding aspect, wherein the variable domain has previously been selected from a non-human vertebrate that has been immunised with the antigen.

For example, the vertebrate (such as a mouse or rat) genome comprises a chimaeric heavy chain locus comprising a human variable region (human V, D and JH gene segments) operably connected upstream of a non-human vertebrate constant region so that the locus is able to rearrange for the expression of heavy chains comprising human variable domains and non-human vertebrate constant regions.

In alternative embodiments, the variable domain is selected using an *in vitro* technology such as phage display, ribosome display or yeast display. In this case the variable domain may be displayed with or without an constant region, provided that it is later combined with a human constant region as per the invention.

4. The method of any preceding aspect, comprising providing an expression vector (Eg, a mammalian expression vector, such as a CHO or HEK293 vector) comprising a nucleotide sequence encoding the constant region; inserting a nucleotide sequence encoding the variable domain into the vector 5' of the constant region sequence; inserting the vector into a host cell and expressing the heavy chain by the host cell; the method further comprising isolating a heavy chain (eg, as part of an antibody) comprising the variable domain and the human constant region.

The vector comprises regulatory elements sufficient to effect expression of the heavy chain when the vector is harboured by a host cell, eg, a CHO or HEK293 cell.

5. The method of any preceding aspect, further comprising obtaining a nucleotide sequence encoding the heavy chain.

6. An antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region.

7. A polypeptide comprising (in N- to C- terminal direction) a leader sequence, a human variable domain that is specific for an antigen and a human constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region; wherein (i) the leader sequence is not the native human variable domain leader sequence (eg, the leader sequence is another human leader sequence or a non-human leader sequence); and/or (ii) the variable domain comprises mouse AID-pattern somatic mutations or mouse terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations.

8. A nucleotide sequence encoding (in 5' to 3' direction) a leader sequence and a human antibody heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region; and the leader sequence being operable for expression (eg, in a mammalian CHO or HEK293 cell) of the heavy chain and wherein the leader sequence is not the native human variable domain leader sequence (eg, the leader sequence is another human leader sequence or a non-human leader sequence).

   In an example, the leader sequence is

   ATGGGCTGGTCCTGCATCATCCTGTTTCTGGTGGCCACCGCCACCGGCGTGCACAGC

   Which translates to

   MGWSCIILFLVATATGVHS

9. A nucleotide sequence encoding (in 5' to 3' direction) a promoter and a human antibody heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref,

IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region; and the promoter being operable for expression (eg, in a mammalian CHO or HEK293 cell) of the heavy chain and wherein the promoter is not the native human promoter.

In one embodiment, the promoter sequence is a human IGK 3-15 promoter.

10. The antibody, polypeptide or nucleotide sequence of any one of aspects 6 to 9, wherein the variable domain comprises mouse AID-pattern somatic mutations and/or mouse terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations.

For example, one way, in any aspect of this configuration of the invention, to provide mouse AID-pattern somatic mutations and/or mouse terminal deoxynucleotidyl transferase (TdT)-pattern junctional mutations is to select a variable domain from a non-human vertebrate or cell. For example, a vertebrate or cell as disclosed herein.

11. A vector (eg, a CHO cell or HEK293 cell vector) comprising the nucleic acid of aspect 8, 9 or 10; optionally wherein the vector is in a host cell (eg, a CHO cell or HEK293 cell).

12. A pharmaceutical composition comprising the antibody or polypeptide of any one of aspects 6, 7 and 10, together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

13. The antibody or polypeptide of any one of aspects 6, 7 and 10 for use in treating and/or preventing a medical condition in a human patient.

14. Use of the antibody or polypeptide of any one of aspects 6, 7 and 10 for the manufacture of a medicament for treating and/or preventing a medical condition in a human patient.

15. The antibody, polypeptide or use of aspect 13 or 14, wherein the human is a member of a human population selected from population numbers 1-14, wherein the populations are numbered as follows (population labels being according to 1000 Genomes Project nomenclature)
1= ASW;

2= CEU;

3=CHB;

4=CHS;

5=CLM;

6=FIN;

7=GBR;

8=IBS;

9=JPT;

10=LWK;

11=MXL;

12=PUR;

13=TSI;

14=YRI.

16. The antibody, polypeptide or use of aspect 15, wherein the constant region is a

(i) IGHA1a constant region and the human population is selected from any population number 1-14;

(ii) IGHA2a constant region and the human population is selected from any population number 1-14;

(iII) IGHA2b constant region and the human population is selected from any population number 1-14;

(iV) IGHG2a constant region and the human population is selected from any population number 1-9 and 11-13;

(v) IGHG3a constant region and the human population is selected from any population number 1-14;

(vi) IGHG3b constant region and the human population is selected from any population number 1-8 and 11-13;

(vii) IGHG4a constant region and the human population is selected from any population number 1-9 and 11-13;

(viii) IGHMa constant region and the human population is selected from any population number 1-14; or

(ix) IGHMb constant region and the human population is selected from any population number 1-14;

Wherein the populations are numbered as follows (population labels being according to 1000 Genomes Project nomenclature)

1= ASW;

2= CEU;

3=CHB;

4=CHS;

5=CLM;

6=FIN;

7=GBR;

8=IBS;

9=JPT;

10=LWK;

11=MXL;

12=PUR;

13=TSI;

14=YRI.

17. A vector (eg, a CHO cell or HEK293 cell vector) comprising a IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region nucleotide sequence that is 3' of a cloning site for the insertion of a human antibody heavy chain variable domain nucleotide sequence, such that upon insertion of such a variable domain sequence the vector comprises (in 5' to 3' direction) a promoter, a leader sequence, the variable domain sequence and the constant region sequence so that the vector is capable of expressing a human antibody heavy chain when present in a host cell.

The present invention provides in a Sixth configuration:-

Multiple Variants in the Same Genome *Cis* or *Trans*

The inventors' analysis has revealed groupings of naturally-occurring human antibody gene segment variants as set out in Table 13 and Table 14. This revealed the possibility of producing transgenic genomes in non-human vertebrates and cells wherein the genomes contain more than the natural

human complement of specific human gene segments. In one example, this can be achieved by providing more than the natural human complement of a specific gene segment type on one or both of the respective Ig locus (eg, one or both chromosomes harbouring IgH in a mouse genome or mouse cell genome).

To this end, this configuration of the invention provides the following (as set out in numbered paragraphs):-

1. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 3 human variable region gene segments of the same type (eg, at least 3 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 3 human Vκ1-39 gene segments, at least 3 human D2-2 gene segments or at least 3 human Jκ1 gene segments), wherein at least two of the human gene segments are variants that are not identical to each other.

   For example, the genome comprises a variable region that comprises V, D and J gene segments (for the variable region of a heavy chain locus) or V and J gene segments (for the variable region of a light chain locus) upstream of a constant region for expression of heavy or light chains respectively.

   In an alternative, the skilled person can choose to provide more than the wild type human complement of a specific gene segment type by providing several copies of one variant type of the human gene segment. Thus, there is provided

   A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 3 human variable region gene segments of the same type (eg, at least 3 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 3 human Vκ1-39 gene segments, at least 3 human D2-2 gene segments or at least 3 human Jκ1 gene segments), wherein the human gene segments are identical variants.

   For example, the genome comprises a variable region that comprises V, D and J gene segments (for the variable region of a heavy chain locus) or V and J gene segments (for the variable region of a light chain locus) upstream of a constant region for expression of heavy or light chains respectively.

2. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *cis* at the same Ig locus.

   In an alternative, the skilled person can choose to provide more than the wild type human complement of a specific gene segment type by providing several copies of one variant type of the human gene segment. Thus, there is provided

   A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 non-endogenous variable region gene segments of the same variant type (eg, at least 2 human JH6*02 gene segments) *cis* at the same Ig locus.

3. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *trans* at the same Ig locus; and optionally a third human gene segment of the same type, wherein the third gene segment is *cis* with one of said 2 different gene segments.

   In an alternative, the skilled person can choose to provide more than the wild type human complement of a specific gene segment type by providing several copies of one variant type of the human gene segment. Thus, there is provided

   A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different human variable region gene segments of the same variant type (eg, at least 2 human JH6*02 gene segments) *trans* at the same Ig locus; and optionally a third human gene segment of the same variant type, wherein the third gene segment is *cis* with one of said 2 different gene segments.

4. A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human variable region gene segments, wherein the plurality comprises at least 2 human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), a first of said different gene segments is provided in the genome of a first vertebrate of the population, and a second of said different gene segments being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second gene segment.

5. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

6. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 3 human variable region gene segments of the same type (eg, at least 3 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 3 human Vκ1-39 gene segments, at least 3 human D2-2 gene segments or at least 3 human Jκ1 gene segments), wherein at least two of the human gene segments are variants that are not identical to each other.

7. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *cis* at the same Ig locus.

8. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *trans* at the same Ig locus; and optionally a third human gene segment of the same type, wherein the third gene segment is *cis* with one of said 2 different gene segments.

9. A method of providing an enhanced human immunoglobulin variable region gene segment repertoire, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human variable region gene segments, wherein the method comprises providing at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein a first of said different gene segments is provided in the genome of a first vertebrate of the population, and a second of said different gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second gene segment.

10. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

11. The vertebrate, cell or method of any preceding paragraph, wherein at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

12. The vertebrate, cell or method of any preceding paragraph, wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically

related over at least 3 generations.

13. The vertebrate, cell or method of any preceding paragraph, wherein the gene segments are derived from the genome sequence of two or more different human individuals; optionally wherein the different human individuals are from different human populations.

14. The vertebrate, cell or method of paragraph 13, wherein the individuals are not genetically related.

15. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same Ig locus in said genome.

16. The method of paragraph 15, wherein the different human individuals are from different human populations.

17. The method of paragraph 15, wherein the individuals are not genetically related.

18. The vertebrate, cell or method of preceding paragraph, wherein at least one of the different segments is a synthetic mutant of a human germline gene segment.

19. The vertebrate, cell or method of any preceding paragraph, wherein each of said gene segments occurs in 10 or more different human populations.

20. The vertebrate, cell or method of preceding paragraph, wherein each of said gene segments has a human frequency of 5% or greater (eg, 10, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, or 95% or greater).

    In this respect, the skilled person can be guided by the information provided in Table 14.

    Frequency can, for example, be cumulative frequency in the 1000 Genomes database.

21. The vertebrate, cell or method of paragraph 20, wherein each of said gene segments occurs in 10 or more different human populations.

22. The vertebrate, cell or method of any preceding paragraph, wherein each of said gene segments occurs in the 1000 Genomes database in more than 50 individuals.

23. The vertebrate, cell or method of preceding paragraph, wherein each of said gene segments (i) has a human frequency of 5% or greater (eg, 10, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, or 95% or greater); and (ii) occurs in 10 or more different human populations.

    In this respect, the skilled person can be guided by the information provided in Table 14.

    Frequency can, for example, be cumulative frequency in the 1000 Genomes database.

24. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from Table 14 (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence (eg, IGHJ6 ref; SEQ ID NO: 244).

    Table 14 lists commonly-occurring natural human variants. It can be seen that these occur across many human populations and thus usefully have wide applicability for human antibody-based drugs.

For example, the gene segments are provided as targeted insertions into an endogenous non-human vertebrate Ig locus. Alternatively, random integration (eg, using YACs) as is know in the art can be performed.

For example, the genome comprises a variable region that comprises V, D and J gene segments (for the variable region of a heavy chain locus) or V and J gene segments (for the variable region of a light chain locus) upstream of a constant region for expression of heavy or light chains respectively.

In another embodiment, the invention enables the skilled person to select two or more different naturally-occurring human gene segment variants for combination into the genome of a non-human vertebrate or cell. A reference sequence need not be included. It may be desirable to use one or more rare gene segments to increase diversity of the repertoire. Additionally or alternatively, it may be desirable to include a mixture of frequent and rare variants of the same type to provide repertoire diversity. The variants may be chosen additionally or alternatively to tailor the gene segment inclusion to one or more specific human populations as indicated by the information provided in Table 13 or Table 14.

Thus, the invention provides

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the gene segments are gene segments selected from Table 13 or Table 14; and optionally wherein one or more of the gene segments appears in Table 14 (eg, IGHJ6-a) or is a reference sequence (eg, IGHJ6 ref; SEQ ID NO: 244).

25. The vertebrate or cell of paragraph 24, wherein the genome comprises a third human gene segment of said type, the third gene segment being different from the first and second gene segments.

26. The vertebrate or cell of paragraph 24 or 25, wherein the first and second gene segments are *cis* on the same chromosome; and optionally the third gene segment is also *cis* on said chromosome.

27. The vertebrate or cell of paragraph 26, wherein the gene segments are targeted insertions into an endogenous non-human Ig locus.

    For example, the gene segments are heavy chain gene segments and the non-human locus is an IgH locus. For example, the gene segments are light chain (kappa or lambda) gene segments and the non-human locus is an IgL locus.

28. The vertebrate or cell of paragraph 24 or 25, wherein the first and second gene segments are *trans* on different chromosomes.

    Thus, the chromosomes are the same type (eg, both mouse chromosome 6 or rat chromosome 4).

29. The vertebrate or cell of any one of paragraphs 24 to 28, wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 (eg, selected from Table 13 or 14) and the second gene segment is the corresponding reference sequence.

30. A population of non-human vertebrates (eg, mice or rats) comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 (eg, Table 13 or 14) (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence (eg, SEQ ID NO: 7), wherein the first gene segment is provided in the genome of a first vertebrate of the population, and the second gene segment is provided in the genome of a second vertebrate of the population.

31. The population of paragraph 30, wherein the genome of the first vertebrate does not comprise the second gene segment.

32. The population of paragraph 30 or 31, wherein the population comprises a third human gene segment of said type, the third gene segment being different from the first and second gene segments and optionally wherein the first and third gene segments are present in the genome of the first vertebrate.

33. The population of paragraph 30, 31 or 32, wherein the gene segments are targeted insertions into an endogenous non-human Ig locus in the respective genome.

    For example, the gene segments are heavy chain gene segments and the non-human locus is an IgH locus. For example, the gene segments are light chain (kappa or lambda) gene segments and the non-human locus is an IgL locus.

34. The population of of any one of paragraphs 30 to 33, wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 (eg, Table 13 or 14) and the second gene segment is the corresponding reference sequence.

35. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 (eg, Table 13 or 14) (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence (eg, SEQ ID NO: 7).

36. A method of providing an enhanced human immunogolobulin gene segment repertoire, the method comprising providing a population according to any one of paragraphs 30 to 33.

<u>Variants Prevalent in Few Populations</u>

In another aspect, it is of note that certain human gene segment variants may appear relatively frequently in one or a small number of populations, but is not found prevalently across many different human populations. There is thinking that specific germline gene segment repertoires have evolved in individual human ethnic populations due to iterative exposure to antigens (eg, disease pathogen antigens) to which the population is often exposed. Repeated exposure and mutation may have lead to the evolution of gene segment variants that can provide an effective response to the antigen (pathogen) in the population, and this may explain the conservation of the gene segments in those populations (as opposed to other human ethnic populations that may not have frequently encountered the antigen). With this in mind, the inventors identified gene segment variants from their analysis that are relatively prevalent in a small number of human populations, and not across many populations. The inventors realized that inclusion of one or more of such gene segments in the configurations of the invention (eg, in transgenic Ig loci, vertebrates and cells) would be useful for producing antibodies, Ig chains and variable domains that can address antigens (eg, disease-causing antigens or pathogens) to which the small number of human populations may become exposed. Such products would be useful for treating and/or preventing disease or medical conditions in members of such a population. This aspect could also be useful for addressing infectious disease pathogens that may have been common in the small number of populations, but which in the future or relatively recently in evolution has become a more prevalent disease-causing pathogen in other human populations (ie, those not listed in Table 13 against the gene segment variant(s) in question). To this end, from the 1000 Genomes database the inventors have identified the gene segment variants listed in Table 20.

Thus, according to any configuration or aspect described herein, one, more or all of the gene segments used in the present invention can be a gene segment listed in Table 20A, 20B, 20C or 20D.

Multiple JH Gene Segment Variants

A specific application of this configuration is the provision of multiple human JH gene segments as follows.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 3 human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein at least two of the human JH gene segments are variants that are not

identical to each other.

In an example, any cell of the invention is an isolated cell. An "isolated" cell is one that has been identified, separated and/or recovered from a component of its production environment (eg, naturally or recombinantly). Preferably, the isolated cell is free of association with all other components from its production environment, eg, so that the cell can produce an antibody to an FDA-approvable or approved standard. Contaminant components of its production environment, such as that resulting from recombinant transfected cells, are materials that would typically interfere with research, diagnostic or therapeutic uses for the resultant antibody, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified: (1) to greater than 95% by weight of antibody as determined by, for example, the Lowry method, and in some embodiments, to greater than 99% by weight; (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Ordinarily, however, an isolated cell will be prepared by at least one purification step.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous JH gene segments (eg, human gene segments) of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *cis* at the same Ig (eg, IgH, eg, endogenous IgH, eg, mouse or rat IgH) locus. In an example, the genome comprises a human VH, D and JH repertoire comprising said different JH gene segments. Optionally the non-endogenous JH gene segments are non-mouse or non-rat, eg, human JH gene segments. In an example one or more or all of the non-endogenous gene segments are synthetic.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *trans* at the same Ig (eg, IgH, eg, endogenous IgH, eg, mouse or rat IgH) locus; and optionally a third human JH gene segments of the same type, wherein the third JH is *cis* with one of said 2 different JH gene segments.

A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human JH gene segments, wherein the plurality comprises at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), a first of said different JH gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JH gene segments being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JH gene segment.

A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous (eg, human) JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations (eg, 3, 4, 5 or 6 generations). Optionally the non-endogenous JH gene segments are human JH gene segments. In an example one or more or all of the non-endogenous gene segments are synthetic.

A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 3 human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein at least two of the human JH gene segments are variants that are not identical to each other.

A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous (eg, human) JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *cis* at the same Ig (eg, IgH, eg, endogenous IgH, eg, mouse or rat IgH) locus). Optionally the non-endogenous JH gene segments are non-mouse or non-rat, eg, human JH gene segments. In an example one or more or all of the non-endogenous gene segments are synthetic.

A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *trans* at the same Ig (eg, IgH, eg, endogenous IgH, eg, mouse or rat IgH) locus; and optionally a third human JH

gene segments of the same type, wherein the third JH is *cis* with one of said 2 different JH gene segments.

A method of providing an enhanced human immunoglobulin JH gene segment repertoire, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human JH gene segments, wherein the method comprises providing at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein a first of said different JH gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JH gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JH gene segment.

A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous (eg, human) JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations (eg, 3, 4, 5, or 6 generations). Optionally the non-endogenous JH gene segments are human JH gene segments. In an example one or more or all of the non-endogenous gene segments are synthetic.

In an example of the vertebrate or cell or the method of the invention at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

In an example of the vertebrate or cell or the method of the invention the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations (eg, 3, 4, 5, or 6 generations).

In an example of the vertebrate or cell or the method of the invention the JH gene segments are derived from the genome sequence of two or more different human individuals; optionally wherein the different human individuals are from different human populations.

In an example of the vertebrate or cell or the method of the invention the individuals are not genetically related (eg, going back 3, 4, 5, or 6 generations).

In an example of the vertebrate or cell or the method of the invention at least one of the different JH segments is a synthetic mutant of a human germline JH gene segment.

The invention also provides a method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations (eg, 3, 4, 5, or 6 generations); optionally wherein at least 2 or 3 of said different gene segments are provided at the same IgH locus in said genome.

In an example of the vertebrate or cell or the method of this embodiment of the invention the genome comprises a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

In an example of the population of the invention, the population comprises a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

A non-human vertebrate (eg, a mouse or rat) or a non-human cell (eg, an ES cell or a B-cell) having a genome comprising a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

A population of non-human vertebrates (eg, mice or rats) comprising a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

In an example of the vertebrate or the population, at least one of said JH gene segments is SEQ ID NO: 1, 2, 3 or 4. For example, at least one of said JH gene segments is SEQ ID NO: 1 and at least one, two or more of said supplementary JH gene segments is a variant according to any example above. For example, at least one of said JH gene segments is SEQ ID NO: 2 and at least one, two or more of said supplementary JH gene segments is a variant according to any one of the examples above. For example, at least one of said JH gene segments is SEQ ID NO: 2 and at least one, two or more of said supplementary JH gene segments is a variant according to any one of the examples above.

In an embodiment, the non-human vertebrate or vertebrate cell of the invention comprises a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the JH gene repertoire (eg, a human JH gene segment repertoire) comprising

a plurality of JH1 gene segments provided by at least 2 different JH1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH2 gene segments provided by at least 2 different JH2 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH3 gene segments provided by at least 2 different JH3 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH4 gene segments provided by at least 2 different JH4 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH5 gene segments provided by at least 2 different JH5 gene segments in *cis* at the same Ig locus in said genome; and/or
a plurality of JH6 gene segments provided by at least 2 different JH6 gene segments in *cis* at the same Ig locus in said genome;
optionally wherein the JH gene segments are derived from the genome sequence of two or more different human individuals.

Optionally said at least 2 different JH gene segments are human gene segments or synthetic gene segments derived from human gene segments.

Optionally, the Ig locus is a IgH locus, eg, an endogenous locus, eg, a mouse or rat IgH locus.

In an embodiment, the non-human vertebrate or vertebrate cell of the invention comprises a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the JH gene repertoire (eg, a human JH gene segment repertoire) comprising
a plurality of JH1 gene segments provided by at least 3 different JH1 gene segments;
a plurality of JH2 gene segments provided by at least 3 different JH2 gene segments;
a plurality of JH3 gene segments provided by at least 3 different JH3 gene segments;
a plurality of JH4 gene segments provided by at least 3 different JH4 gene segments;
a plurality of JH5 gene segments provided by at least 3 different JH5 gene segments; and/or
a plurality of JH6 gene segments provided by at least 3 different JH6 gene segments;
optionally wherein the JH gene segments are derived from the genome sequence of two or three different human individuals;
optionally wherein at least 2 or 3 of said different gene segments are provided in *cis* at the same Ig locus in said genome.

Optionally said at least 3 different JH gene segments are human gene segments or synthetic gene segments derived from human gene segments.

Optionally, the Ig locus is a IgH locus, eg, an endogenous locus, eg, a mouse or rat IgH locus.

Optionally in the vertebrate or cell the different human individuals are from different human populations.

Optionally in the vertebrate or cell the individuals are not genetically related (eg, Going back 3, 4, 5 or 6generations).

Optionally in the vertebrate or cell at least one of the different JH segments is a synthetic mutant of a human germline JH gene segment.

In an embodiment of a non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell) according to the invention, the vertebrate or cell genome comprises human VH, D and JH gene repertoires, the JH gene repertoire (eg, a human JH gene repertoire) comprising
a plurality of JH1 gene segments provided by at least 2 different human JH1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH2 gene segments provided by at least 2 different human JH2 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH3 gene segments provided by at least 2 different human JH3 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH4 gene segments provided by at least 2 different human JH4 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH5 gene segments provided by at least 2 different human JH5 gene segments, optionally in *cis* at the same Ig locus in said genome; and/or
a plurality of JH6 gene segments provided by at least 2 different human JH6 gene segments, optionally in *cis* at the same Ig locus in said genome;
wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations (eg, 3, 4, 5 or 6 generations).

Optionally said at least 2 different JH gene segments are human gene segments or synthetic gene segments derived from human gene segments.

Optionally, the Ig locus is a IgH locus, eg, an endogenous locus, eg, a mouse or rat IgH locus.

Optionally in the vertebrate or cell the human individuals are from different human populations.

JH5

An embodiment provides a vertebrate, cell or population of the invention whose genome comprises a plurality of JH5 gene segments, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a nucleotide mutation at one or more positions

corresponding to positions

106,330,024
106,330,027
106,330,032
106,330,041
106,330,044
106,330,045
106,330,062
106,330,063
106,330,065
106,330,066
106,330,067
106,330,068 and
106,330,071
on human chromosome 14.

In the vertebrate, cell or population optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a guanine at a position corresponding to position 106,330,067 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,071 on human chromosome 14 (optionally the additional mutation being a guanine); (ii) position 106,330,066 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (iii) position 106,330,068 on human chromosome 14 (optionally the additional mutation being a thymine).

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a guanine at a position corresponding to position 106,330,071 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,063 on human chromosome 14 (optionally the additional mutation being an adenine); and/or (ii) position 106,330,067 on human chromosome 14 (optionally the additional mutation being a guanine).

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a cytosine at a position corresponding to position 106,330,045 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises an adenine at a position corresponding to position 106,330,044 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,066 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,330,068 on human chromosome 14 (optionally the additional mutation being a thymine).

Optionaly the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a guanine at a position corresponding to position 106,330,066 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,067 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,330,068 on human chromosome 14 (optionally the additional mutation being a thymine).

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a thymine at a position corresponding to position 106,330,068 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,067 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,330,066 on human chromosome 14 (optionally the additional mutation being a guanine).

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a cytosine at a position corresponding to position 106,330,027 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises an adenine at a position corresponding to position 106,330,024 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a thymine at a position corresponding to position 106,330,032 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a thymine at a position corresponding to position 106,330,041 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises an adenine or thymine at a position corresponding to position 106,330,063 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the variant comprises additionally a mutation at a position corresponding to position 106,330,071 on human chromosome 14 (optionally the additional mutation being a guanine).

Optionally the plurality comprises a human JH5 gene variant of SEQ ID NO: 1,wherein the variant comprises a cytosine at a position corresponding to position 106,330,062 on human chromosome

14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

Optionally the genome comprises SEQ ID NO:1; optionally in *cis* at the same Ig locus as one, two or more of the variants.

JH6

An embodiment provides a vertebrate, cell or population of the invention whose genome comprises a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a nucleotide mutation at one or more positions corresponding to positions

106,329,411
106,329,413
106,329,414
106,329,417
106,329,419
106,329,426
106,329,434
106,329,435, and
106,329,468
on human chromosome 14.

Optionally the genome of the vertebrate, cell or population comprises a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a guanine at a position corresponding to position 106,329,435 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,329,468 on human chromosome 14 (optionally the additional mutation being a guanine); (ii) position 106,329,419 on human chromosome 14 (optionally the additional mutation being an adenine); (iii) position 106,329,434 on human chromosome 14 (optionally the additional mutation

being a cytosine) and/or position 106,329,414 on human chromosome 14 (optionally the additional mutation being a guanine); (iv) position 106,329,426 on human chromosome 14 (optionally the additional mutation being an adenine); (v) position 106,329,413 on human chromosome 14 (optionally the additional mutation being an adenine); (vi) position 106,329,417 on human chromosome 14 (optionally the additional mutation being a thymine); (vii) position 106,329,411 on human chromosome 14 (optionally the additional mutation being a thymine); (viii) position 106,329,451 on human chromosome 14 (optionally the additional mutation being an adenine); (ix) position 106,329,452 on human chromosome 14 (optionally the additional mutation being a cytosine); and/or (x) position 106,329,453 on human chromosome 14 (optionally the additional mutation being a cytosine).

Optionally the variant comprises additionally mutations at positions corresponding to position 106,329,451 on human chromosome 14, the additional mutation being an adenine; position 106,329,452 on human chromosome 14, the additional mutation being a cytosine; and position 106,329,453 on human chromosome 14, the additional mutation being a cytosine.

The vertebrate, cell or population optionally comprises a plurality of JH6 gene segments,, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a guanine at a position corresponding to position 106,329,468 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

Optionally the variant comprises additionally a mutation at a position corresponding to position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

Optionally the vertebrate, cell or population comprises a plurality of JH6 gene segments,, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a thymine at a position corresponding to position 106,329,417 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

Optionally the variant comprises additionally a mutation at a position corresponding to position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

Optionally the vertebrate, cell or population comprises a plurality of JH6 gene segments,, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a

cytosine at a position corresponding to position 106,329,434 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,329,414 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

Optionally the vertebrate, cell or population comprises a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a thymine at a position corresponding to position 106,329,411 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

Optionally the variant comprises additionally a mutation at a position corresponding to position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

Optionally the vertebrate, cell or population comprises a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant that is an antisense sequence of a variant described above.

Optionally the genome comprises SEQ ID NO:2; optionally *cis* at the same Ig locus as one, two or more of the JH6 variants.

JH2

An embodiment provides a vertebrate, cell or population of the invention whose genome comprises a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises a nucleotide mutation at one or more positions corresponding to positions

106,331,455
106,331,453, and
106,331,409

on human chromosome 14.

Optionally the vertebrate, cell or population comprises said plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises a guanine at a position corresponding to position 106,331,455 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 3.

Optionally the variant comprises additionally a mutation at a position corresponding to (i) position 106,331,453 on human chromosome 14 (optionally the additional mutation being an adenine); and/or (ii) position 106,331,409 on human chromosome 14 (optionally the additional mutation being an adenine); (iii) position 106,329,434 on human chromosome 14 (optionally the additional mutation being an adenine).

Optionally the vertebrate, cell or population comprises a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises an adenine at a position corresponding to position 106,331,453 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 3.

Optionally the variant comprises additionally a mutation at a position corresponding to position 106,331,409 on human chromosome 14 (optionally the additional mutation being an adenine).

Optionally the vertebrate, cell or population comprises a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises an adenine at a position corresponding to position 106,331,409 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 3.

Optionally the vertebrate, cell or population comprises a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant that is an antisense sequence of a variant described above.

Optionally the genome comprises SEQ ID NO:3; optionally *cis* at the same Ig locus as one, two or more of the JH2 variants.

Optionally the vertebrate, cell or population genome comprises two or more different JH gene segments selected from SEQ ID NOs: 1 to 3 and variants described above; optionally wherein said JH gene segments are *cis* at the same immunoglobulin Ig locus.

Multiple Human D Gene Segment Variants

A specific application of this configuration is the provision of multiple human D gene segments as follows (as set out in numbered clauses, starting at clause number 154).

154. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 3 human D gene segments of the same type (eg, D2-2 gene segments), wherein at least two of the human D gene segments are variants that are not identical to each other (eg, D2-2ref and D2-2a).

In an example of any aspect of the sixth configuration of the invention (V, D, J or C), one or more or all of the variants are naturally-occurring human gene segments.

In an example of any aspect of the sixth configuration of the invention (V, D, J or C), one or more of the variants may be a synthetic variant of a human gene segment.

155. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous D gene segments of the same type type (eg, D2-2ref and D2-2a) *cis* at the same Ig locus.

156. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different human D gene segments of the same type (eg, D2-2ref and D2-2a) *trans* at the same Ig locus; and optionally a third human D gene segment (eg, (eg, D2-2ref, D2-2a or D2-2b) of the same type, wherein the third D is *cis* with one of said 2 different D gene segments.

157. A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human D gene segments, wherein the plurality comprises at least 2 different human D gene segments of the same type (eg, D2-2 gene segments), a first of said different D gene segments (eg, D2-2ref) is provided in the genome of a first vertebrate of the population, and a second of said different D gene segment (eg, D2-2a) being provided in the genome of a second vertebrate of the

population, wherein the genome of the first vertebrate does not comprise the second D gene segment.

158. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous D gene segments of the same type (eg, human D2-2 gene segments), wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

159. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 3 human D gene segments of the same type (eg, D2-2 gene segments), wherein at least two of the human D gene segments are variants that are not identical to each other (eg, D2-2ref and D2-2a).

160. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous D gene segments of the same type (eg, human D2-2 gene segments) *cis* at the same Ig locus.

161. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different human D gene segments of the same type (eg, D2-2ref and D2-2a) *trans* at the same Ig locus; and optionally a third human D gene segment (eg, D2-2ref, D2-2a or D2-2b) of the same type, wherein the third D is *cis* with one of said 2 different D gene segments.

162. A method of providing an enhanced human immunoglobulin D gene segment repertoire, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human D gene segments, wherein the method comprises providing at least 2 different human D gene segments of the same type (eg, D2-2ref and D2-2a), wherein a first of said different D gene segments is provided in the genome of a first vertebrate of the population, and a second of said different D gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not

comprise the second D gene segment.

163. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous D gene segments of the same type (eg, D2-2ref and D2-2a), wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

164. The vertebrate or cell of clause 154, 156 or 158, or the method of clause 159, 161 or 163, wherein at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

165. The vertebrate or cell of clause 154, 155 or 156, or the method of any one of clauses 159 to 162 and 164, wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

166. The vertebrate or cell of any one of clauses 154 to 157, or the method of any one of clauses 159 to 162 and 165, wherein the D gene segments are derived from the genome sequence of two or more different human individuals; optionally wherein the different human individuals are from different human populations.

167. The vertebrate, cell or method of clause 166, wherein the individuals are not genetically related.

168. The vertebrate, cell or method of any one of clauses 154 to 167, wherein at least one of the different D segments is a synthetic mutant of a human germline D gene segment.

169. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human D gene segments of the same type (eg, D2-2ref and D2-2a), wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same IgH locus in said genome.

170. The vertebrate or cell of any one of clauses 154 to 158 and 164 to 168, wherein the genome comprises a substantially complete functional repertoire of human D gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

171. The population of clause 157, wherein the population comprises a substantially complete functional repertoire of human D gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

172. A non-human vertebrate (eg, a mouse or rat) or a non-human cell (eg, an ES cell or a B-cell) having a genome comprising a substantially complete functional repertoire of human D gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

173. A population of non-human vertebrates (eg, mice or rats) comprising a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

174. The vertebrate or cell of clause 172 or the population of clause 173, comprising first and second D gene segments selected from

D2-2ref and D2-2a; or

D2-21ref and D2-21a; or

D3-10ref and D3-10a; or

D3-16ref and D3-16a; or

D2-8ref and D2-8a; or

D3-3ref and D3-3a; or

D4-23ref and D4-23a; or

D6-13ref and D6-13a; or

D3-9ref and D3-9a; or

D4-4ref and D4-4a; or

D7-27ref and D7-27a;

Optionally wherein the first and/or second D gene segment is present in two or more copies.

For example, there are provided two or three copies of the first gene segment, optionally with one, two or three copies of the second gene segment. Copies can be arranged in *cis* or *trans.*

175. The vertebrate, cell or population of clause 174, comprising human gene segments D2-2ref and D2-2a; and D3-3ref and D3-3a; and optionally also D2-15.

In an example, the vertebrate, cell or population comprises one or more D segments selected from human D3-3, D2-15, D3-9; D4-17; D3-10; D2-2; D5-24; D6-19; D3-22; D6-13; D5-12; D1-26; D1-20; D5-18; D3-16; D2-21; D1-14; D7-27; D1-1; D6-25; D2-14 and D4-23 (eg, selected from D3-9*01; D4-17*01; D3-10*01; D2-2*02; D5-24*01; D6-19*01; D3-22*01; D6-13*01; D5-12*01; D1-26*01; D1-20*01; D5-18*01; D3-16*02; D2-21*02; D1-14*01; D7-27*02; D1-1*01; D6-25*01; D2-15*01; and D4-23*01), together with the reference sequence(s) of said selected segment(s). These were found in variable domains having a HCDR3 length of at least 20 amino acids (see examples herein).

176. A non-human vertebrate or vertebrate cell according to clause 155, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the D gene repertoire comprising one or more of

a plurality of D2-2 gene segments provided by at least 2 different D2-2 gene segments in *cis* at the same Ig locus in said genome;

a plurality of D2-21 gene segments provided by at least 2 different D2-21 gene segments in *cis* at the same Ig locus in said genome;

a plurality of D3-10 gene segments provided by at least 2 different D3-10 gene segments in *cis* at the same Ig locus in said genome;

a plurality of D3-16 gene segments provided by at least 2 different D3-16 gene segments in *cis* at the same Ig locus in said genome;

a plurality of D2-8 gene segments provided by at least 2 different D2-8 gene segments in *cis* at the same Ig locus in said genome;

a plurality of D3-3 gene segments provided by at least 2 different D3-3 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D4-23 gene segments provided by at least 2 different D4-23 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D6-13 gene segments provided by at least 2 different D6-13 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D3-9 gene segments provided by at least 2 different D3-9 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D4-4 gene segments provided by at least 2 different D4-4 gene segments in *cis* at the same Ig locus in said genome; and
a plurality of D7-27 gene segments provided by at least 2 different D7-27 gene segments in *cis* at the same Ig locus in said genome;

optionally wherein the D gene segments are derived from the genome sequence of two or more different human individuals.

177. A non-human vertebrate or vertebrate cell according to clause 155, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the D gene repertoire comprising one or more of

a plurality of D2-2 gene segments provided by at least 2 different D2-2 gene segments in *trans* in said genome;
a plurality of D2-21 gene segments provided by at least 2 different D2-21 gene segments in *trans* in said genome;
a plurality of D3-10 gene segments provided by at least 2 different D3-10 gene segments in *trans* in said genome;
a plurality of D3-16 gene segments provided by at least 2 different D3-16 gene segments in *trans* in said genome;
a plurality of D2-8 gene segments provided by at least 2 different D2-8 gene segments in *trans* in said genome;
a plurality of D3-3 gene segments provided by at least 2 different D3-3 gene segments in *trans* in said genome;
a plurality of D4-23 gene segments provided by at least 2 different D4-23 gene segments in *trans* in said genome;

a plurality of D6-13 gene segments provided by at least 2 different D6-13 gene segments in *trans* in said genome;
a plurality of D3-9 gene segments provided by at least 2 different D3-9 gene segments in *trans* in said genome;
a plurality of D4-4 gene segments provided by at least 2 different D4-4 gene segments in *trans* in said genome; and
a plurality of D7-27 gene segments provided by at least 2 different D7-27 gene segments in *trans* in said genome;

optionally wherein the D gene segments are derived from the genome sequence of two or more different human individuals.

178. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell), according to clause 154, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the D gene repertoire comprising one or more of

a plurality of D2-2 gene segments provided by at least 3 different D2-2 gene segments;
a plurality of D2-21 gene segments provided by at least 3 different D2-21 gene segments;
a plurality of D3-10 gene segments provided by at least 3 different D3-10 gene segments;
a plurality of D3-16 gene segments provided by at least 3 different D3-16 gene segments;
a plurality of D2-8 gene segments provided by at least 3 different D2-8 gene segments;
a plurality of D3-3 gene segments provided by at least 3 different D3-3 gene segments;
a plurality of D4-23 gene segments provided by at least 3 different D4-23 gene segments;
a plurality of D6-13 gene segments provided by at least 3 different D6-13 gene segments;
a plurality of D3-9 gene segments provided by at least 3 different D3-9 gene segments;
a plurality of D4-4 gene segments provided by at least 3 different D4-4 gene segments; and
a plurality of D7-27 gene segments provided by at least 3 different D7-27 gene segments;

optionally wherein the D gene segments are derived from the genome sequence of two or three different human individuals;

optionally wherein at least 2 or 3 of said different gene segments are provided in *cis* at the same

Ig locus in said genome.

179. The vertebrate or cell of clause 176, 177 or 178, wherein the different human individuals are from different human populations.

180. The vertebrate or cell of any one of clauses 176 to 179, wherein the individuals are not genetically related.

181. The vertebrate or cell of any one of clauses 176 to 180, wherein at least one of the different D segments is a synthetic mutant of a human germline D gene segment.

182. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell) according to clause 158, comprising a genome comprising human VH, D and JH gene repertoires, the D gene repertoire comprising of one or more of

a plurality of D2-2 gene segments provided by at least 2 different D2-2 gene ; optionally in *cis* in said genome;
a plurality of D2-21 gene segments provided by at least 2 different D2-21 gene ; optionally in *cis* in said genome;
a plurality of D3-10 gene segments provided by at least 2 different D3-10 gene ; optionally in *cis* in said genome;
a plurality of D3-16 gene segments provided by at least 2 different D3-16 gene ; optionally in *cis* in said genome;
a plurality of D2-8 gene segments provided by at least 2 different D2-8 gene ; optionally in *cis* in said genome;
a plurality of D3-3 gene segments provided by at least 2 different D3-3 gene ; optionally in *cis* in said genome;
a plurality of D4-23 gene segments provided by at least 2 different D4-23 gene ; optionally in *cis* in said genome;
a plurality of D6-13 gene segments provided by at least 2 different D6-13 gene ; optionally in *cis* in said genome;
a plurality of D3-9 gene segments provided by at least 2 different D3-9 gene ; optionally in *cis* in said genome;
a plurality of D4-4 gene segments provided by at least 2 different D4-4 gene ; optionally in *cis* in

said genome; and

a plurality of D7-27 gene segments provided by at least 2 different D7-27 gene ; optionally in *cis* in said genome;

wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

183. The vertebrate or cell of clause 182, wherein the human individuals are from different human populations.

184. The vertebrate, cell or population of any one of clauses 154 to 183, wherein one or more of the D gene segments is a variant of a human germline D gene segment, wherein the variant gene segment encodes an amino acid sequence that differs by 1, 2 or 3 amino acids from the corresponding amino acid sequence encoded by the human germline D gene segment, provided in that said amino acid sequence encoded by the variant does not include a stop codon when said corresponding amino acid sequence does not include a stop codon.

Optionally, the variant and germline D gene segments encode the respective amino acid sequences in reading frame 2 (IMGT numbering). See Briney *et al* 2012.

185. The vertebrate, cell or population of clause 184, wherein said corresponding amino acid sequence encoded by the human germline D gene segment is a hydrophilic or hydrophobic sequence (according to J Mol Biol. 1997 Jul 25;270(4):587-97; Corbett SJ *et al*; Table 2).

186. The vertebrate, cell or population of clause 184 or 185, comprising said variant and said germline human D gene segments; optionally wherein the variant and germline human D gene segments are *cis* on the same chromosome.

187. The vertebrate, cell or population of any one of clauses 184 to 186, wherein germline human D gene segment is a D2, D3, D5 or D6 family gene segment; optionally a D2-2, D2-15, D3-3, D3-9, D3-10, D3-22, D5-5, D5-18, D6-6, D6-13, D6-19 gene segment.

These D segments are usable in all three reading frames.

Optionally a variant of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or all of these human germline D gene segments is used.

188. The vertebrate, cell or population of any one of clauses 154 to 187, comprising a plurality of D2-2 gene segments, wherein the plurality comprises D2-2 gene segments that vary from each other at one or more nucleotide positions corresponding to positions

106,382,687 and
106,382,711
on human chromosome 14.

189. The vertebrate, cell or population of clause 188, wherein the plurality comprises a human D2-2 gene segment ((optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,382,687 on human chromosome 14; and optionally no further mutation from the sequence of D2-2ref.

190. The vertebrate, cell or population of clause 188 or 189, wherein the plurality comprises a human D2-2 gene segment comprising a cytosine at a position corresponding to position 106,382,687 on human chromosome 14; and optionally no further mutation from the sequence of D2-2a.

191. The vertebrate, cell or population of any one of clauses 188 to 190, wherein the plurality comprises a human D2-2 gene segment comprising an adenine at a position corresponding to position 106,382,711 on human chromosome 14; and optionally no further mutation from the sequence of D2-2b.

192. The vertebrate, cell or population of any one of clauses 188 to 191, wherein the plurality comprises a human D2-2 gene segment comprising an thymine at a position corresponding to position 106,382,711 on human chromosome 14; and optionally no further mutation from the sequence of D2-2ref.

193. The vertebrate, cell or population of any one of clauses 154 to 192, comprising a plurality of D7-27 gene segments, wherein the plurality comprises D7-27 gene segments that vary from each other at a nucleotide position corresponding to position 106,331,767 on human chromosome 14.

194. The vertebrate, cell or population of clause 193, wherein the plurality comprises a human D7-27 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,331,767 on human chromosome 14; and optionally no further mutation from the sequence of D7-27ref.

195. The vertebrate, cell or population of clause 193 or 194, wherein the plurality comprises a human D7-27 gene segment comprising a guanine at a position corresponding to position 106,331,767 on human chromosome 14; and optionally no further mutation from the sequence of D7-27a.

196. The vertebrate, cell or population of any one of clauses 154 to 195, comprising a plurality of D4-23 gene segments, wherein the plurality comprises D4-23 gene segments that vary from each other at a nucleotide position corresponding to position 106,350,740 on human chromosome 14.

197. The vertebrate, cell or population of clause 196, wherein the plurality comprises a human D4-23 gene segment (optionally two copies and/or in homozygous state) comprising an adenine at a position corresponding to position 106,350,740 on human chromosome 14; and optionally no further mutation from the sequence of D4-23ref.

198. The vertebrate, cell or population of clause 196 or 197, wherein the plurality comprises a human D4-23 gene segment (optionally two copies and/or in homozygous state) comprising an guanine at a position corresponding to position 106,350,740 on human chromosome 14; and optionally no further mutation from the sequence of D4-23a.

199. The vertebrate, cell or population of any one of clauses 154 to 197, comprising a plurality of D2-21 gene segments, wherein the plurality comprises D2-21 gene segments that vary from each other at a nucleotide position corresponding to position 106,354,418 on human chromosome 14.

200. The vertebrate, cell or population of clause 199, wherein the plurality comprises a human D2-21 gene segment (optionally two copies and/or in homozygous state) comprising an adenine

at a position corresponding to position 106,354,418 on human chromosome 14; and optionally no further mutation from the sequence of D2-21ref.

201. The vertebrate, cell or population of clause 199 or 200, wherein the plurality comprises a human D2-21 gene segment (optionally two copies and/or in homozygous state) comprising a guanine at a position corresponding to position 106,354,418 on human chromosome 14; and optionally no further mutation from the sequence of D2-21a.

202. The vertebrate, cell or population of any one of clauses 154 to 201, comprising a plurality of D3-16 gene segments, wherein the plurality comprises D3-16 gene segments that vary from each other at a nucleotide position corresponding to position 106,354,418 on human chromosome 14.

203. The vertebrate, cell or population of clause 202, wherein the plurality comprises a human D3-16 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,361,515 on human chromosome 14; and optionally no further mutation from the sequence of D3-16ref.

204. The vertebrate, cell or population of clause 202 or 203, wherein the plurality comprises a human D3-16 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,361,515 on human chromosome 14; and optionally no further mutation from the sequence of D3-16a.

205. The vertebrate, cell or population of any one of clauses 154 to 204, comprising a plurality of D6-13 gene segments, wherein the plurality comprises D6-13 gene segments that vary from each other at a nucleotide position corresponding to position 106,367,013 on human chromosome 14.

206. The vertebrate, cell or population of clause 205, wherein the plurality comprises a human D6-13 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,367,013 on human chromosome 14; and optionally no further mutation from the sequence of D6-13ref.

207. The vertebrate, cell or population of clause 205 or 206, wherein the plurality comprises a human D6-13 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,367,013 on human chromosome 14; and optionally no further mutation from the sequence of D6-13a.

208. The vertebrate, cell or population of any one of clauses 154 to 207, comprising a plurality of D3-10 gene segments, wherein the plurality comprises D3-10 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,370,370 and
106,370,371
on human chromosome 14.

209. The vertebrate, cell or population of clause 208, wherein the plurality comprises a human D3-10 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,370,370 on human chromosome 14; and optionally no further mutation from the sequence of D3-10ref.

210. The vertebrate, cell or population of clause 208 or 209, wherein the plurality comprises a human D3-10 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,370,370 on human chromosome 14; and optionally no further mutation from the sequence of D3-10a.

211. The vertebrate, cell or population of clause 208, 209 or 210 wherein the plurality comprises a human D3-10 gene segment (optionally two copies and/or in homozygous state) comprising an adenine at a position corresponding to position 106,370,371 on human chromosome 14; and optionally no further mutation from the sequence of D3-10ref.

212. The vertebrate, cell or population of any one of clauses 208 to 211, wherein the plurality comprises a human D3-10 gene segment (optionally two copies and/or in homozygous state) comprising a guanine at a position corresponding to position 106,370,371 on human chromosome 14; and optionally no further mutation from the sequence of D3-10b.

213. The vertebrate, cell or population of any one of clauses 154 to 212, comprising a plurality of D3-9 gene segments, wherein the plurality comprises D3-9 gene segments that vary from each

other at a nucleotide position corresponding to position 106,370,567 on human chromosome 14.

214. The vertebrate, cell or population of clause 213, wherein the plurality comprises a human D3-9 gene segment (optionally two copies and/or in homozygous state) comprising an adenine at a position corresponding to position 106,370,567 on human chromosome 14; and optionally no further mutation from the sequence of D3-9ref.

215. The vertebrate, cell or population of clause 213 or 214, wherein the plurality comprises a human D3-9 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,370,567 on human chromosome 14; and optionally no further mutation from the sequence of D3-9a.

216. The vertebrate, cell or population of any one of clauses 154 to 215, comprising a plurality of D2-8 gene segments, wherein the plurality comprises D2-8 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,373,085;
106,373,086 and
106,373,089
on human chromosome 14.

217. The vertebrate, cell or population of clause 216, wherein the plurality comprises a human D2-8 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,373,085 on human chromosome 14.

218. The vertebrate, cell or population of clause 216 or 217, wherein the plurality comprises a human D2-8 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,373,085 on human chromosome 14; and optionally no further mutation from the sequence of D2-8b.

219. The vertebrate, cell or population of clause 216, 217 or 218 wherein the plurality comprises a human D2-8 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,373,086 on human chromosome 14; and

optionally no further mutation from the sequence of D2-8ref.

220. The vertebrate, cell or population of any one of clauses 216 to 219, wherein the plurality comprises a human D2-8 gene segment comprising a thymine at a position corresponding to position 106,373,086 on human chromosome 14; and optionally no further mutation from the sequence of D2-8ref.

221. The vertebrate, cell or population of any one of clauses 154 to 220, comprising a plurality of D4-4 gene segments, wherein the plurality comprises D4-4 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,379,086; and
106,379,089
on human chromosome 14.

222. The vertebrate, cell or population of clause 221, wherein the plurality comprises a D4-4 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,379,086 on human chromosome 14; and optionally no further mutation from the sequence of D4-4ref.

223. The vertebrate, cell or population of clause 221 or 222, wherein the plurality comprises a human D4-4 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,379,086 on human chromosome 14; and optionally no further mutation from the sequence of D4-4a.

224. The vertebrate, cell or population of clause 221, 222 or 223 wherein the plurality comprises a human D4-4 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,379,089 on human chromosome 14; and optionally no further mutation from the sequence of D4-4ref or a cytosine at a position corresponding to position 106,379,086 on human chromosome 14.

225. The vertebrate, cell or population of any one of clauses 221 to 224, wherein the plurality comprises a human D4-4 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,373,089 on human

chromosome 14; and optionally no further mutation from the sequence of D4-4a.

226. The vertebrate, cell or population of any one of clauses 154 to 225, comprising a plurality of D3-3 gene segments, wherein the plurality comprises D3-3 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,380,241; and
106,380,246
on human chromosome 14.

227. The vertebrate, cell or population of clause 226, wherein the plurality comprises a D3-3 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,380,241 on human chromosome 14; and optionally no further mutation from the sequence of D3-3ref.

228. The vertebrate, cell or population of clause 226 or 227, wherein the plurality comprises a human D3-3 gene segment (optionally two copies and/or in homozygous state) comprising a cytosine at a position corresponding to position 106,380,241 on human chromosome 14; and optionally no further mutation from the sequence of D3-3a.

229. The vertebrate, cell or population of clause 226, 227 or 228 wherein the plurality comprises a human D3-3 gene segment (optionally two copies and/or in homozygous state) comprising an adenine at a position corresponding to position 106,380,246 on human chromosome 14; and optionally no further mutation from the sequence of D3-3ref.

230. The vertebrate, cell or population of any one of clauses 226 to 229, wherein the plurality comprises a human D3-3 gene segment (optionally two copies and/or in homozygous state) comprising a thymine at a position corresponding to position 106,380,246 on human chromosome 14; and optionally no further mutation from the sequence of D3-3a.

Multiple Human JL Gene Segment Variants

A specific application of this configuration is the provision of multiple human JL gene segments (Jκ and/or Jλ) as follows (as set out in numbered paragraphs, starting at paragraph number 80).

80. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 3 human JL gene segments of the same type (eg, Jκ1), wherein at least two of the human JL gene segments are variants that are not identical to each other.

81. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1) *cis* at the same Ig locus.

82. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different human JL gene segments of the same type (eg, Jκ1) *trans* at the same Ig locus; and optionally a third human JL gene segment of the same type, wherein the third JL is *cis* with one of said 2 different JL gene segments.

83. A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human JL gene segments, wherein the plurality comprises at least 2 different human JL gene segments of the same type (eg, Jκ1), a first of said different JL gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JL gene segments being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JL gene segment.

84. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1), wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

85. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 3 human JL gene segments of the same type (eg, Jκ1), wherein at least two of the human JL gene segments are variants that are not identical to each other.

86. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1) *cis* at the same Ig

locus.

87. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different human JL gene segments of the same type(eg, Jκ1) *trans* at the same Ig locus; and optionally a third human JL gene segment of the same type, wherein the third JL is *cis* with one of said 2 different JL gene segments.

88. A method of providing an enhanced human immunoglobulin JL gene segment repertoire, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human JL gene segments, wherein the method comprises providing at least 2 different human JL gene segments of the same type (eg, Jκ1), wherein a first of said different JL gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JL gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JL gene segment.

89. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1), wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

90. The vertebrate or cell of paragraph 80, 82 or 84, or the method of paragraph 85, 82 or 89, wherein at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

91. The vertebrate or cell of paragraph 80, 81 or 82, or the method of paragraph 85, 86 or 87, wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

92. The vertebrate or cell of paragraph 80, 81 or 82, or the method of paragraph 85, 86 or 87, wherein the JL gene segments are derived from the genome sequence of two or more different human individuals; optionally wherein the different human individuals are from different human

populations.

93. The vertebrate, cell or method of paragraph 92, wherein the individuals are not genetically related.

94. The vertebrate, cell or method of any one of paragraphs 80 to 93, wherein at least one of the different JL segments is a synthetic mutant of a human germline JL gene segment.

95. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human JL gene segments of the same type (eg, Jκ1), wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same IgL locus in said genome.

96. The vertebrate or cell of any one of paragraphs paragraph 80 to 82 and 84, wherein the genome comprises a substantially complete functional repertoire of human Jκ and/or Jλ gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

97. The population of paragraph 83, wherein the population comprises a substantially complete functional repertoire of human JL gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

98. A non-human vertebrate (eg, a mouse or rat) or a non-human cell (eg, an ES cell or a B-cell) having a genome comprising a substantially complete functional repertoire of human Jκ and/or Jλ gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

99. A population of non-human vertebrates (eg, mice or rats) comprising a substantially complete functional repertoire of human Jκ and/or Jλ gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

100. A non-human vertebrate or vertebrate cell according to paragraph 81, comprising a genome that comprises VL and JL gene repertoires comprising human gene segments, the JL gene repertoire comprising

a plurality of human Jκ1 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ2 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ3 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ4 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ5 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ1 gene segments provided by at least 2 different human Jλ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ2 gene segments provided by at least 2 different human Jλ2 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ3 gene segments provided by at least 2 different human Jλ3 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ4 gene segments provided by at least 2 different human Jλ4 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ5 gene segments provided by at least 2 different human Jλ5 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ6 gene segments provided by at least 2 different human Jλ6 gene segments in *cis* at the same Ig locus in said genome; or
a plurality of human Jλ7 gene segments provided by at least 2 different human Jλ7 gene segments in *cis* at the same Ig locus in said genome;

optionally wherein the JL gene segments are derived from the genome sequence of two or more different human individuals.

101. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell), according to paragraph 80, comprising a genome that comprises VL and JL gene repertoires comprising human gene segments, the JL gene repertoire comprising

a plurality of human Jκ1 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jκ1 gene segments;
a plurality of human Jκ2 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jκ1 gene segments;
a plurality of human Jκ3 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jκ1 gene segments;
a plurality of human Jκ4 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jκ1 gene segments;
a plurality of human Jκ5 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jκ1 gene segments;
a plurality of human Jλ1 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jλ1 gene segments;
a plurality of human Jλ2 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jλ2 gene segments;
a plurality of human Jλ3 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jλ3 gene segments;
a plurality of human Jλ4 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jλ4 gene segments;
a plurality of human Jλ5 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jλ5 gene segments;
a plurality of human Jλ6 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jλ6 gene segments; or
a plurality of human Jλ7 gene segments provided by at least 3 (eg, 3, 4, 5, 6, or 7) different human Jλ7 gene segments;

optionally wherein the JL gene segments are derived from the genome sequence of two or three

different human individuals;

optionally wherein at least 2 or 3 of said different gene segments are provided in *cis* at the same Ig locus in said genome.

102. The vertebrate or cell of paragraph 104 or 105, wherein the different human individuals are from different human populations.

103. The vertebrate or cell of any one of paragraphs 104 to 106, wherein the individuals are not genetically related.

104. The vertebrate or cell of any one of paragraphs 104 to 107, wherein at least one of the different JL segments is a synthetic mutant of a human germline JL gene segment.

105. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell) according to paragraph 84, comprising a genome comprising human VL and JL gene repertoires, the JL gene repertoire comprising

a plurality of human Jκ1 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ2 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ3 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ4 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ5 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ1 gene segments provided by at least 2 different human Jλ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ2 gene segments provided by at least 2 different human Jλ2 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ3 gene segments provided by at least 2 different human Jλ3 gene segments, optionally in *cis* at the same Ig locus in said genome;

a plurality of human Jλ4 gene segments provided by at least 2 different human Jλ4 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ5 gene segments provided by at least 2 different human Jλ5 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ6 gene segments provided by at least 2 different human Jλ6 gene segments, optionally in *cis* at the same Ig locus in said genome; or
a plurality of human Jλ7 gene segments provided by at least 2 different human Jλ7 gene segments, optionally in *cis* at the same Ig locus in said genome;
wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

106. The vertebrate or cell of paragraph 109, wherein the human individuals are from different human populations.

The skilled person will realise that standard molecular biology techniques can be used to provide vectors comprising synthetic combinations of immunoglobulin gene segments (eg, V, D and/or J) for use in 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 of gene segments can be made using standard recombineering techniques in E coli to construct BAC vectors harbouring the synthetic combination 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 and in EP1034260 and EP1204740 the disclosures of which are explicitly incorporated herein.

In one embodiment, it 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 Infulenza virus; it is possible, therefore, to include two or more polymorphic DNA versions of the VH segment VH1-69 in

the locus of the invention. The examples below illustrate how such a transgenic locus can be constructed in which diversity is extended by extending the VH1-69 gene segment repertoire based on naturally-occurring VH1-69 polymorphic variants.

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 one embodiment each said locus constant region is a heavy chain endogenous non-human vertebrate (optionally host mouse or rat) constant region.

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

The invention provides a monoclonal or polyclonal antibody composition prepared by immunisation of at least one vertebrate (eg, mouse or rat) according to the invention, optionally wherein the antigen is an antigen of an infectious disease pathogen (eg, a bacterial or viral pathogen antigen), 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 also provides a monoclonal or polyclonal antibody mixture produced by the method of the invention 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). In an aspect of the invention, the monoclonal or polyclonal antibody mixture is provided for therapy and/or prophylaxis of a disease or condition in a human, eg, 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 the invention, or a mutant or derivative antibody thereof in the manufacture of a medicament for the treatment and/or prevention of a disease or condition in a human, eg, 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 method of the invention. 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. 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.

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

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 9).

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_λII gene family member, V_λVII 4A, V_λII 2.1, V_λVII 4A, a V_λ1 gene family member, a V_λ3gene family member, IGLV1S2, V_λ3-cML70, lalh2, lalvl, la3h3, Kv325, a VκI gene family member, κI-15A (KL012), V_κII family member, a V_κIII family member, a VκI gene family member, κI-15A (KL012), V_κII A2 (optionally the A2a variant), V_κ A27 (Humkv325) and a gene segment at least 80% identical thereto.

(b) a V_λ gene segment selected from a V_λII gene family member, V_λVII 4A, V_λII 2.1, V_λVII 4A, a V_λ1 gene family member, a V_λ3gene family member, IGLV1S2, V_λ3-cML70, lalh2, lalvl, la3h3 and a gene segment at least 80% identical thereto.

(c) a V_κ gene segment selected from Kv325, a VκI gene family member, κI-15A (KL012), V_κII family member, a V_κIII family member, a VκI gene family member, κI-15A (KL012), V_κII A2 (optionally the A2a variant), V_κ A27 (Humkv325) and a gene segment at least 80% identical thereto.

(d) a V_H gene segment a V_HIII gene family member (optionally, a VHIIIa or VHIIIb 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.

(e) a J_λ gene segment selected from J_λ2, J_λ3 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_λ gene segment selected from a V_λII gene family member, V_λVII 4A, V_λII 2.1, V_λVII 4A and a gene segment at least 80% identical thereto.

(b) a V_κ gene segment selected from a VκI gene family member, κI-15A (KL012), V_κII family member, a V_κIII family member, a VκI gene family member, κI-15A (KL012), V_κII A2 (optionally the A2a variant), V_κ A27 (Humkv325) and a gene segment at least 80% identical thereto.

(c) a V_H gene segment a VH3 gene family member (optionally, a VHIIIa or VHIIIb 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 and a gene segment at least 80% identical thereto.

(d) a J_λ gene segment selected from J_λ2, J_λ3 and a gene segment at least 80% identical thereto.

(e) a J_H gene segment selected from J_H2, J_H3, J_H4 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_λ gene segment selected from a V_λII gene family member, V_λVII 4A, V_λII 2.1, V_λVII 4A and a gene segment at least 80% identical thereto.

(b) a V_κ gene segment selected from a V_κII family member, a V_κIII family member, a VκI gene family member, κI-15A (KL012), V_κII A2 (optionally the A2a variant), V_κ 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_HIII 9.1 (VH3-15), V_HIII VH26 (VH3-23), V_H3-21, LSG6.1, LSG12.1, DP77 (V3-21) and a gene segment at least 80% identical thereto.

(d) a J_λ gene segment selected from J_λ2, J_λ3 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_HIII 9.1 (VH3-15), V_H H11, V_HIII VH26 (VH3-23) a gene segment at least 80% identical thereto, eg, V_HIII 9.1 + J_H3; or V_H H11 + J_H4; or V_HIII VH26 + J_H2.

(b) a V_κ gene segment selected from a VκI gene family member, κI-15A (KL012) and a gene segment at least 80% identical thereto.

(c) a V_λ gene segment selected from a V_λII gene family member, V_λII 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_λ gene segment selected from a V_λ1 gene family member, a V_λ3gene family member, IGLV1S2, V_λ3-cML70, lalh2, lalvl, la3h3 and a gene segment at least 80% identical thereto.

(c) a Vk 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_λ gene segment selected from J_λ2, J_λ3 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_λ 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 $J_H 3$, $J_H 5$, $J_H 6$ 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_λ 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_H gene segment selected from Hv1051 and a gene segment at least 80% identical thereto.

(b) a Vk 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_H 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_H 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_L 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.

Antibodies For Therapy & Prophylaxis of Patients of Specific Ancestry

The inventors, having undertaken the extensive Bioinformatics analysis exercise described herein, realised that the output of that analysis has made it possible to identify specific gene segments that are useful to produce antibody- and VH domain-based drugs that are tailored specifically to a patient's ancestry (ie, genotype). That is, antibodies can be selected on the basis that they are made *in vivo* in a transgenic non-human vertebrate (eg, mouse or rat with transgenic IgH loci) and particularly derived from gene segments that are relatively prevalent in members of the patient's population, ie, from individuals of the same human ancestry. Since variant distributions differ across different populations (see Table 13), this presumably reflects the effects of evolution, adaptation and conservation of useful variant gene types in those populations. Thus, by tailoring the antibody-based drugs according to the invention, it is possible to match the drug to the population gene biases, thus with the aim of making better drugs for that specific population of humans. Better can, for example, mean more efficacious, better neutralising, higher target antigen affinity, less immunogenic, less patient reactions to the drug etc. This can be determined empirically, as is standard in drug research and development processes.

Thus, the invention provides the following embodiments (numbered from clause 345 onwards):-

345. An isolated antibody for administration to a Chinese patient, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen

and a constant region, wherein the constant region is a human constant region selected from a constant region (eg, an IGHG constant region) inTable 13 found in a Chinese population and with a cumulative frequency of at least 1 or 5%;, and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); and/or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

In another embodiment, the invention provides

An isolated antibody for administration to a Chinese patient, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the constant region is a human constant region selected from a constant region (eg, an IGHG constant region) present in a Chinese population with a cumulative frequency of at least 5%;, and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat);and/ or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

In an example, the constant region is found in the 1000 Genomes database. In an example, the constant region is found in Table 13.

346. The antibody of clause 345 wherein the constant region is a IGHG1a, IGHG2a, IGHG3a, IGHG3b or IGHG4a constant region.

347. The antibody of clause 345 or 346, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

In another embodiment, the invention provides

The antibody of clause 345 or 346, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

348. The antibody of clause 345, 346 or 347, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

In another embodiment, the invention provides

The antibody of clause 345, 346 or 347, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

349. The antibody of clause 345, 346, 347 or 348 wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

In another embodiment, the invention provides

The antibody of clause 345, 346, 347 or 348 wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

350. An isolated VH domain identical to a variable domain as recited in any one of clauses 347 to 349, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

In an embodiment, there is provided an isolated VH domain identical to a variable domain as recited in any one of clauses 347 to 349 which is part of a conjugate, conjugated with a label (eg, for imaging in the patient) or a toxin (eg, a radioactive toxic payload, such as for cancer treatment in the patient) or a half-life-extending moiety (eg, PEG of human serum albumin).

351. A pharmaceutical composition comprising the antibody or variable domain of any one of clauses 345 to 350 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

352. An isolated antibody for administration to a Chinese patient, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen

and a constant region, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%; and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); and/or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

In another embodiment, the invention provides

An isolated antibody for administration to a Chinese patient, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH present in a Chinese population with a cumulative frequency of at least 5%; and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); and/or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

353. The antibody of clause 352, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

In another embodiment, the invention provides

The antibody of clause 352, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

354. The antibody of clause 352 or 353, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

In another embodiment, the invention provides

The antibody of clause 352 or 353, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

355. An isolated VH domain identical to a variable domain as recited in any one of clauses 352 to 354, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

In an embodiment, there is provided a VH domain identical to a variable domain as recited in any one of clauses 352 to 354 which is part of a conjugate, conjugated with a label (eg, for imaging in the patient) or a toxin (eg, a radioactive toxic payload, such as for cancer treatment in the patient) or a half-life-extending moiety (eg, PEG of human serum albumin).

356. A pharmaceutical composition comprising the antibody or variable domain of any one of clauses 352 to 355 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

357. An antibody heavy chain or VH domain (eg, provided as part of an antibody) for therapy and/or prophylaxis of a disease or medical condition in a Chinese patient, wherein the heavy chain is a heavy chain produced by the following steps (or is a copy of such a heavy chain):-

(a) Selection of an antigen-specific antibody heavy chain or VH domain from a non-human vertebrate (eg, a mouse or a rat), wherein the heavy chain or VH domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%;

(b) Optional humanisation of the heavy chain by combining the variable domain of the heavy chain with a human constant region; or optional humanisation of the selected VH domain by combining with a human constant region.

In another embodiment, the invention provides

An antibody heavy chain or VH domain (eg, provided as part of an antibody) for therapy and/or prophylaxis of a disease or medical condition in a Chinese patient, wherein the heavy chain is a heavy chain produced by the following steps (or is a copy of such a heavy chain):-

(a) Selection of an antigen-specific antibody heavy chain or VH domain from a non-human vertebrate (eg, a mouse or a rat), wherein the heavy chain or VH domain is derived from the

recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH present in a Chinese population with a cumulative frequency of at least 5%;

(b) Optional humanisation of the heavy chain by combining the variable domain of the heavy chain with a human constant region; or optional humanisation of the selected VH domain by combining with a human constant region.

In an example, the VH gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

358. The antibody heavy chain or VH domain of clause 357, wherein the human constant region is as recited in clause 345 or 346.

359. An antibody heavy chain or VH domain as recited in clause 357 or 358 for use in a medicament for therapy and/or prophylaxis of a disease or medical condition in a Chinese patient.

360. A method of treating and/or preventing a disease or medical condition in a Chinese patient, the method comprising administering to the patient a therapeutically or prophylactically-effective amount of the antibody heavy chain or VH domain as recited in clause 357 or 358.

361. An isolated antibody for administration to a patient of European, East Asian, West African, South Asian or Americas ancestry, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the constant region is a human constant region selected from a constant region (eg, an IGHG constant region) in Table 13 found in a population of European, East Asian, West African, South Asian or Americas ancestry respectively and with a cumulative frequency of at least 1 or 5%;, and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

In another embodiment, the invention provides

An isolated antibody for administration to a patient of European, East Asian, West African, South Asian or Americas ancestry, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the constant region is a human constant region selected from a constant region (eg, an IGHG constant region) present in a population of European, East Asian, West African, South Asian or Americas ancestry respectively with a cumulative frequency of at least 1 or 5%;, and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a

non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

In an example, the constant region is found in the 1000 Genomes database. In an example, the constant region is found in Table 13.

362. The antibody of clause 361 wherein the constant region is a IGHG1a, IGHG2a, IGHG3a, IGHG3b or IGHG4a constant region and the patient is of European ancestry.

363. The antibody of clause 361 or 362, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment, the VH gene segment being selected from a VH in Table 13 found in said population and with a cumulative frequency of at least 1 or 5%.

In another embodiment, the invention provides

The antibody of clause 361 or 362, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

364. The antibody of clause 361, 362 or 363, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in said population and with a cumulative frequency of at least 1 or 5%.

In another embodiment, the invention provides

The antibody of clause 361, 362 or 363, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

365. The antibody of clause 361, 362, 363 or 364 wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in said population and with a cumulative frequency of at least 1 or 5%.

In another embodiment, the invention provides

The antibody of clause 361, 362, 363 or 364 wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH present in a Chinese population with a cumulative frequency of at least 5%.

In an example, the gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

366. An isolated VH domain identical to a variable domain as recited in any one of clauses 363 to 365, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

367. A pharmaceutical composition comprising the antibody or variable domain of any one of clauses 361 to 366 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

368. An isolated antibody for administration to a patient of European, East Asian, West African or Americas ancestry, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a population of European, East Asian, West African, South Asian or Americas ancestry respectively and with a cumulative frequency of at least 1 or 5%; and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

In another embodiment the invnention provides:-

An isolated antibody for administration to a patient of European, East Asian, West African or Americas ancestry, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH present in a population of European, East Asian, West African, South Asian or Americas ancestry respectively with a cumulative frequency of at least 1 or 5%; and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg,

mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

In an example, the VH gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

369. The antibody of clause 368, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in said population and with a cumulative frequency of at least 1 or 5%.

In another example there is provided

The antibody of clause 368, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D present in said population with a cumulative frequency of at least 1 or 5%.

In an example, the D gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

370. The antibody of clause 368 or 369, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in said population and with a cumulative frequency of at least 1 or 5%.

In another example there is provided

The antibody of clause 368 or 369, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH present in said population and with a cumulative frequency of at least 1 or 5%.

In an example, the JH gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

371. An isolated VH domain identical to a variable domain as recited in any one of clauses 368 to 370, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

372. A pharmaceutical composition comprising the antibody or variable domain of any one of clauses 368 to 371 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

373. An antibody heavy chain or VH domain (eg, provided as part of an antibody) for therapy and/or prophylaxis of a disease or medical condition in a patient of European, East Asian, West African, South Asian or Americas ancestry, wherein the heavy chain is a heavy chain produced by

the following steps (or is a copy of such a heavy chain):-

(a) Selection of an antigen-specific antibody heavy chain or VH domain from a non-human vertebrate (eg, a mouse or a rat), wherein the heavy chain or VH domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in said population and with a cumulative frequency of at least 1 or 5%;

(b) Optional humanisation of the heavy chain by combining the variable domain of the heavy chain with a human constant region; or optional humanisation of the selected VH domain by combining with a human constant region.

In another embodiment, there is provided:-

An antibody heavy chain or VH domain (eg, provided as part of an antibody) for therapy and/or prophylaxis of a disease or medical condition in a patient of European, East Asian, West African, South Asian or Americas ancestry, wherein the heavy chain is a heavy chain produced by the following steps (or is a copy of such a heavy chain):-

(a) Selection of an antigen-specific antibody heavy chain or VH domain from a non-human vertebrate (eg, a mouse or a rat), wherein the heavy chain or VH domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH present in said population with a cumulative frequency of at least 1 or 5%;

(b) Optional humanisation of the heavy chain by combining the variable domain of the heavy chain with a human constant region; or optional humanisation of the selected VH domain by combining with a human constant region.

In an example, the VH gene segment is found in the 1000 Genomes database. In an example, the gene segment is found in Table 13.

374. The antibody heavy chain or VH domain of clause 373, wherein the human constant region is as recited in clause 361 or 362.

375. An antibody heavy chain or VH domain as recited in clause 373 or 374 for use in a medicament for therapy and/or prophylaxis of a disease or medical condition in a patient of said ancestry.

376. A method of treating and/or preventing a disease or medical condition in a patient of European, East Asian, West African, South Asian or Americas ancestry, the method comprising administering to the patient a therapeutically or prophylactically-effective amount of the antibody heavy chain or VH domain as recited in clause 373 or 374.

In embodiments herein, a Chinese patient can be a Han Chinese patient.

In embodiments herein, a patient of European ancestry can be a patient of Northern or Western European ancestry, Italian ancestry, British or Scottish ancestry, Finnish ancestry or Iberian ancestry.

In embodiments herein, a patient of East Asian ancestry can be a patient of Han Chinese ancestry, Japanese ancestry Chinese Dai ancestry, Vietnamese ancestry or Kinh ancestry.

In embodiments herein, a patient of West African ancestry can be a patient of Yoruba ancestry, Luhya ancestry, Gambian ancestry or Malawian ancestry.

In embodiments herein, a patient of Americas ancestry can be a patient of African American ancestry, African Caribbean ancestry, Mexican ancestry, Puerto Rican ancestry, Colombian ancestry or Peruvian ancestry.

In embodiments herein, a patient of South Asian ancestry can be a patient of Ahom ancestry, Kayadtha ancestry, Reddy ancestry, Maratha ancestry, or Punjabi ancestry.

In an example of any aspect, the cumulative frequency is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90 or 95%.

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 prophetic Examples.

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 psuedogenes. Using recombineering 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 recombineering 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 variant 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 10-2, 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 immunoglobin 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 Cμ 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 variants of V gene segments. One source of such variants include sequence databases. In this example, 13 distinct variants of the VH1-69 gene segment are provided.

These variant sequences and comparisons are drawn from the "IMmunoGeneTics" IMGT Information System (www.imgt.com) database. Figure 5 is a diagram of the alignment of variants *02 through *13 with the *01 variant. The VH1-69*01 nucleotide and amino acid sequence is provided at the top of the figure. Where the remaining variants are identical to the *01 variant sequence a dash is inserted below the sequence. Nucleotide differences are noted alongside the appropriate variant 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 variant in comparison to the *01 variant. 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

Human IgH JH Gene Variants Selected from the 1000 Genomes Database

Data is presented for human JH2, 5 and 6 variants. In Tables 10A, 11A and 12A samples from humans from various populations are listed where the sequence analysis of the inventors has revealed the presence of polymorphisms in one or both IgH JH alleles. The population codes are explained in Table 8 above. The polymorphisms are nucleotide variants from JH2, 5 and 6 reference sequences (SEQ ID NOs: 1, 2 and 3 respectively; see below). All references are sequences taken from the Ensembl database (*www.**ensembl**.org)*. The JH5 reference is human IgH J5-001 disclosed in that database. The JH6 reference is human IgH J6-001 disclosed in that database. The JH2 reference is human IgH J2-001 disclosed in that database.

The reference nucleotide and encoded amino acid sequences are shown on the next page. Alignments with encoded amino acid sequences are also provided, including the corresponding position numbers on human chromosome 14.

Variant Frequencies are shown in Tables 10A, 11A and 12A and these relate to the frequency of the variants in the 1000 Genomes Database (release current at October 2011).

Tables 10B, 11B and 12B show the non-synonymous nucleotide polymorphisms in the human JH variants, as sorted by the present inventors from the 1000 Genomes database. Position numbers corresponding to nucleotide positions on human chromosome 14 are shown for variant positions (chromosome 14 being the chromosome bearing the IgH locus in humans). Thus, for example, the first entry in Table 11B is "14:106330027:A/C" which refers to a position in a variant JH5 sequence wherein the position corresponds to position 106,330,027 on human chromosome 14, such position being A (adenine) in the reference sequence. The "C" indicates that the present inventors observed a mutation to cytosine at this position in the variants found in the 1000 Genomes database. This change leads to a change at the amino acid level of the encoded sequence (ie, a "non-synonymous" change), in this case a change from a serine (found in the reference) to an alanine in the variant.

Example 5:

Human Antibody Gene Segment Variant Identification & Population Analysis

The genomic coding region coordinates for each target gene for variant analysis were identified from the Ensembl WWW site (www.ensembl.org) using coordinates from the GRCh.p8 Human Genome assembly (www.ncbi.nlm.nih.gov/projects/genome/assembly/grc). Using the collected gene location coordinates, variant data was extracted from the public ftp site of the 1000 Genomes Project using the Perl 'Variant Pattern Finder' (VPF - www.1000genomes.org/variation-pattern-finder-api-documentation).

Data extracted by VPF was post processed using software to extract all non-synonymous (NSS) variants with their associated genotype calls. Genotypes calls were assembled to form unique haplotypes, representing groups of NSS variants associated with 1000 Genome population groups and frequency of occurrence within those populations.

The output of the analysis results in tables such as in Table 13. The main body of the table describes each haplotype in turn giving a unique ID for that gene (in the range a-z,aa-zz), the population frequencies and occurrence in individuals and unique population groups; one or more subsequent columns describe the DNA base calls at each location that form the haplotype giving both the base from the reference sequence or the variant base call.

Table 13 was constructed in this manner. The table can be read as follows:

The first four columns (left to right) consist of (1) the haplotype ID letter ('ref' indicates reference - the DNA base call at each genomic location from the GRCh37 Human Reference Assembly) (2) the observed cumulative frequency of the haplotype among the different populations (3) the number of individuals in which a specific haplotype was observed (4) the number of unique population groups that the identified individuals belong to (the actual population group identifiers are displayed as a string of ID's in the most right hand column for each haplotype. For example haplotype 'a' has a population ID string of '3,4,9,13').

The populations are numbered as follows (population labels being according to 1000 Genomes Project nomenclature)
1= ASW;
2= CEU;
3=CHB;
4=CHS;
5=CLM;
6=FIN;
7=GBR;
8=IBS;

9=JPT;
10=LWK;
11=MXL;
12=PUR;
13=TSI;
14=YRI.

Subsequent columns detail a single point variant and have the following format (top to bottom) (1) the human genomic location of the variant (format [chromosome number]: [location] e.g. '14:106204113'); (2) The identifier for the point variant as defined in DbSNP (www.ncbi.nlm.nih.gov/projects/SNP/); (3) One or additional rows show the amino acid change as result of the variant for a specific transcript (denoted by the Ensembl transcript ID in the most right-hand column for each row), the format is the amino acid in the reference sequence followed by '->' and the amino acid caused by the substitution of the variant in the reference sequence (e.g. 'Gly->Arg' means a that the translated reference sequence would result in a glycine at that location, whereas the substitution of the identified variant would result in translated protein containing arginine) using the IUPAC three letter amino acid codes (http://pac.iupac.org/publications/pac/pdf/1972/pdf/3104x0639.pdf). Subsequent rows (one per haplotype) show the DNA base at each location, bases matching the reference sequence are shown in black on white back ground, bases varying from the reference are shown as white text on a black background.

The most right-hand column contains the Ensembl transcript ID's (e.g. 'ENST00000390542') for each of the gene transcript and relates to the amino acid changes to the left of this column. Because the transcripts are differing lengths each variant position may or may not have an associated amino acid change at the that position.

Example 6:

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

A functional human gene segment repertoire (from V_H2-26 to J_H6, 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 WO2011004192, 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 (see an extract of the UCSC databse in Figure 10, identified BACs being circled) 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, RP11-141I7, 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_H2-5, V_H7-4-1, V_H4-4, V_H1-3, V_H1-2, V_H6-1, 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_H1, J_H2, J_H3, J_H4, J_H5 and J_H6 (in 5' to 3' order), wherein the JH6 was chosen to be the human JH6*02 variant. The insertion was made between positions 114666435 and 114666436 on mouse chromosome 12, which is upstream of the mouse Cμ region. The mouse V_H, D and J_H 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_H gene segments were inserted upstream (5') of the 5'-most V_H 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_H3-13, V_H3-11, V_H3-9, V_H1-8, V_H3-7. The mouse V_H, D and J_H 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_H gene segments were inserted upstream (5') of the 5'-most V_H 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_H2-26, V_H1-24, V_H3-23, V_H3-21, V_H3-20, V_H1-18, and V_H3-15. The mouse V_H, D and J_H 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^R into the mouse J_H-Cμ intron (to produce the "HA" allele).

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 Freunds 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 Freunds 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

Screening

Culture supernatants of the hybridoma wells above were screened using homogenious time resolved fluorescence assay (htrf) using Europium cryptate labelled anti-mouse IgG (Cisbio anti-mouse Ig Europium Cryptate) and a biotin tagged target antigen with a commercially available streptavidin conjucated donor (Cisbio; streptaviding conjugated D2) or by IgG-specific 384 well ELISA. Positive wells identified by htrf 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 conflueny, supernatants were re-tested using htrf 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.

Bionformatics 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 °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		
	JH6*02%	Total Count	JH6*02%	Total Count	% HCDR3>20
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.

SEQ ID NO: 1 (JH5 Reference)

```
T  T  G  A  C  C  A  A  G  C  T  G  G  G  A  C  C  C  C
G  G  T  C  C  C  T  T  G  G  G  A  C  C  C  A  G  T  G  G
A  G  A  G  G  G  A  G  T  C
```

JH5 Alignment:

(top line=SEQ ID NO: 1, Middle line=-SEQ ID NO:5, Bottom line=SEQ ID NO:6)

SEQ ID NO: 2 (JH6 Reference)

```
A T G A T G A T G A T G G T A
C C T G C A G A C C C C C T
G G T G C C A G T G G C A G
C
```

JH6 Alignment:

(top line=SEQ ID NO: 2, Middle line=-SEQ ID NO:7, Bottom line=SEQ ID NO:8)

G G G A C C A C G G T C A C C G T C T C C T C A G
G T T V T V S S

SEQ ID NO: 3 (JH2 Reference)

A T G A C C A T G A A G C T A G A G A C

C C C G G C A C C G T G G G A C C A G T

G A C A G A G G A G T C

JH2 Alignment:

(top line=SEQ ID NO: 3, Middle line=-SEQ ID NO:9, Bottom line=SEQ ID NO:10)

100-5J Hg

-064 133 601 -55 133 601 35 133 601 35 133 601

T G A C G A G A A G C T A G A G A C C C C G G C A C C G T G G G A C C A G T G A C A G A G G A G T
A C T G G T A C T T C G A T C T C T G G G G C C G T G G C A C C C T G G T C A C T G T C T C C T C A G
Y W Y F D L W G R G T L V T V S S

TABLES

In the tables, the notation is illustrated by the following example

IGLV1-40	V1-40*02	X53936	g9>c|c10>g,L4>V|

Polymorphic variant IGV lambda V1-40*02 has Genbank Accession No. X53936 and when compared to the *01 variant, the V1-40*02 variant has mutations at positions 9, 10 and 4. For example, at position 9, a "C" appears instead of a "G" that is present in the *01 variant. The "|" is simply a notation separator, and does not indicate any mutation. For example the "g282 I" notation indicates no change (ie, position 282 is a g). "del#" means that the residue at that position is absent.

Table 1: Human IgH V Polymorphic Variants

All variant sequences appearing in Genbank under the listed Accession Numbers are incorporated herein in their entirety by reference as though explicitly recited herein and for possible inclusion in any of the claims herein.

subgroup	IGHV	IGHV	Accession number	Description of mutations
1	IGHV1-2	IGHV1-2*01	X07448	SEQ ID NO: 11
		IGHV1-2*02	X62106	|c163>t,R55>W| |g233>t,t234>g,S78>M|t299>c,V100>A|
		IGHV1-2*03	X92208	c44>t,P15>L|c163>t,R55>W| |g233>t,t234>g,S78>M|t299>c,V100>A |
		IGHV1-2*04	Z12310	|c163>t,R55>W|a223>t,R75>W|g233>t,t234>g,S78>M|t299>c,V100 >A|
	IGHV1-3	IGHV1-3*01	X62109	SEQ ID NO: 12 c6 |t12 |t167 ,I56 |a208 ,K70 |a291 |c296 ,T99 |
		IGHV1-3*02	X62107	c6>t|t12>g|t167>g,I56>S|a208>g,K70>E|a291>g|c296>t,T99>M|
	IGHV1-8	IGHV1-8*01	M99637	SEQ ID NO: 13
	IGHV1-18	IGHV1-18*01	M99641	g282 |
		IGHV1-18*02	X60503	g282>a|

IGHV1-24	IGHV1-24*01	M99642	SEQ ID NO: 14
IGHV1-45	IGHV1-45*01	X92209	SEQ ID NO: 15 g139 ,G47 |t237 |
	IGHV1-45*02	AB019438	|t237>c|
	IGHV1-45*03	Z17391	g139>a,G47>R|t237>c|
IGHV1-46	IGHV1-46*01	X92343	SEQ ID NO: 16 c92 ,T31 |g315 |
	IGHV1-46*02	J00240	c92>a,T31>N|
	IGHV1-46*03	L06612	|g315>t|
IGHV1-58	IGHV1-58*01	M29809	SEQ ID NO: 17 g115 ,V39 |
	IGHV1-58*02	AB019438	g115>a,V39>M|

IGHV1-69	IGHV1-69*01	L22582	SEQ ID NO: 18	
	IGHV1-69*02	Z27506	g6>c|g18>a| |g100>a,A34>T|g163>a,G55>R|t178>c,F60>L|c185>t,T62>| | |g244>a,E82>K|	
	IGHV1-69*03	X92340	|g291>t,E97>D|	
	IGHV1-69*04	M83132	g6>c| |g163>a,G55>R|t178>c,F60>L|c185>t,T62>| |g244>a,E82>K|	
	IGHV1-69*05	X67905	g6>c| |g238>a,A80>T|	
	IGHV1-69*06	L22583	|g244>a,E82>K|	
	IGHV1-69*07	Z29978	|g163>a,G55>R|	
	IGHV1-69*08	Z14309	g6>c|g18>a| |g100>a,A34>T|g163>a,G55>R|t178>C,F60>L| |g244>a,E82>K|	
	IGHV1-69*09	Z14307	|g163>a,G55>R|t178>c,F60>L|c185>t,T62>| |g244>a,E82>K|	
	IGHV1-69*10	Z14300	g6>c| |g54>a| |t178>c,F60>L|c185>t,T62>| |g244>a,E82>K|	
	IGHV1-	Z14296	g6>c| |g163>a,G55>R|t178>c,F60>L|	

		69*11		g6>c				
		IGHV1-69*12	Z14301					
		IGHV1-69*13	Z14214(st)	g6>c| |g54>a|				
	IGHV1-c	IGHV1-c*01	Z18904	SEQ ID NO: 19				
	IGHV1-f	IGHV1-f*01	Z12305	SEQ ID NO: 20 c201 |				
		IGHV1-f*02	Z29977	c201>t|				
2	IGHV2-5	IGHV2-5*01	X62111	SEQ ID NO: 21				
		IGHV2-5*02	Z14072	|a175>g,N59>D|				
		IGHV2-5*03	X93619	|a175>g,N59>D| |c234>t|				
		IGHV2-5*04	L21963	|c299>g,A100>G|c314>t,A105>V|a317>g,c318>g,H106>R|				
		IGHV2-	L21964	|a175>g,N59>D|a202>g,S68>G|				

	5*05		
	IGHV2-5*06	L21966	|g39>a| |a175>g,N59>D|a202>g,S68>G|
	IGHV2-5*07	L21968	|c299>g,A100>G |c314>t,A105>V|
	IGHV2-5*08	L21971	a4>g,I2>V|a31>g,T11>A| |g60>a|g103>a,V35>M|g106>c,G36>R|g118>a ,G40>S|a175>g,N59>D|
	IGHV2-5*09	L21972	a4>g,I2>V| |a175>g,N59>D|a202>g,S68> G|
	IGHV2-5*10	X69690	|a175>g,N59>D| |a317>g,c318>g,H106>R|
IGHV2-26	IGHV2-26*01	M99648	SEQ ID NO: 22
IGHV2-70	IGHV2-70*01	L21969	SEQ ID NO: 23 22 a2 ,Q1 |g4 ,V2 |g14 ,R5 |g31 ,A11 |a60 |t67 ,C23 |a70 ,T24 |t10 6 ,C36 |t116 ,V39 |a138 |c157 ,L53 |t164 ,L55 |a197 ,Y66 |t210 |g2 16 ,K72 |a297 |c299 ,A100 |a301 ,c302 ,T101 |g303 |c309 |c314 ,A1 05 |g317 ,g318 ,R106 |t320 ,I107 |
	IGHV2-70*02	X92241	|a297>g| |a301>g,c302>t,T101>V|

IGHV2-70*03	X92238	|g14>a,R5>K| |t164>g,L55>R|a197>t,Y66>F| |t106>c,C36>R| |a297>g| |a301>g,c302>t,T101>V|	
IGHV2-70*04	Z12330	|g14>a,R5>K| |t164>g,L55>R|a197>t,Y66>F| |t106>c,C36>R|	
IGHV2-70*05	Z27502	|t106>c,C36>R|t116>c,V39>A| |t164>g,L55>R|a197>t,Y66>F|	
IGHV2-70*06	X92239	|g14>a,R5>K| |t164>g,L55>R|a197>t,Y66>F|t210>c| |t106>c,C36>R| |a297>g| |a301>g,c302>t,T101>V|	
IGHV2-70*07	X92243	|a297>g| |a301>g,c302>t,T101>V| |a138>g|	
IGHV2-70*08	X92245	|a297>g| |a70>g,T24>A| |a301>g,c302>t,T101>V| |t 164>g,L55>R|	
IGHV2-70*09	L21962	|g4>a,V2>I|g14>a,R5>K|g31>a,A11>T|a60>g|t67>c,C23>R| |g303>a| |c314>t,A105>V| |g216>c,K72>N| |c299>g,A100>G|	
IGHV2-70*10	L21965	|g14>a,R5>K| c157>a,L53>I|t164>g,L55>R| |t106>c,C36>R| |	
IGHV2-70*11	L21967	a2>g,Q1>R| 64>g,L55>R| |t1	

		IGHV2-70*12	L21970	|g4>a,V2>||g14>a,R5>K|g31>a,A11>T|a60>g| |g303>a| |g317a,g318>c,R106>H|t320>g,I107>R|
		IGHV2-70*13	AB019437	|c309>t|
3	IGHV3-7	IGHV3-7*01	M99649	SEQ ID NO: 24
		IGHV3-7*02	X92288	g144>a|
	IGHV3-9	IGHV3-9*01	M99651	SEQ ID NO: 25
	IGH3-11	IGHV3-11*01	M99652	SEQ ID NO: 26 |g13 ,V5 |g32 ,G11 |g47 ,G16 |g178 ,G60 |a184 ,c185 ,T62 |t188 ,I63 |t196 ,Y66 |a206 ,D69 |g240 |c243 ,D81 |c296 ,g297 ,T99 |g301 ,g303 ,V101 |t304 ,a305 ,Y102 |t307 ,a308 ,c309 ,Y103 |t312 |c314 ,A105 |
		IGHV3-11*02	M15496	9^10>ins^t| |g32>del#,G11>del#|g47>del#,G16>del#| |a206>del#,D69>del#| |c243>g,D81>E|c296>del#,g297>del#,T99>del#|g301>t,g303>a,V101>L|t304>c,a305>t,Y102>L|t307>a,a308>c,c309>t,Y103>T|t312>c|c314>a,A105>E|

		IGHV3-11*03	X92287	|g13>t,V5>L| |g178>a,G60>S|a184>t,c185>a,T62> Y|t188>c,I63>T|t196>a,Y66>N| |g240>a|
	IGHV3-13	IGHV3-13*01	X92217	SEQ ID NO: 27 g9 ,Q3 |t52 ,S18 |g77 ,S26 |g95 ,S32 |t165 |t167 ,I56 |c222 |g224 , R75 |
		IGHV3-13*02	M99653	g9>t,Q3>H|t52>g,S18>A| |g95>a,S32>N|t165>c|t167>a,I56>N|c222> g|
		IGHV3-13*03	U29582	|c77>g,S26>C| |g224>a,R75>Q|
	IGHV3-15	IGHV3-15*01	X92216	SEQ ID NO: 28 g23 ,g24 ,G8 |g32 ,G11 |a40 ,K14 |a81 |t89 ,F30 |g119 ,S40 |t159 | c163 ,R55 |a169 ,K57 |a178 ,T60 |g181 ,D61 |g196 ,D66 |c210 |a242 ,D81 |a275 ,N92 |c279 |a297 |a320 |
		IGHV3-15*02	M99654	|g32>c,G11>A|
		IGHV3-15*03	M99408	g23>c,g24>c,G8>A|g32>c,G11>A|a40>c,K14>Q| |t89>g,F30>C| |a178>g,T60>A|g181>a,D61>N| |c210>t|a242>t,D81 >V|
		IGHV3-15*04	M99402	|a169>g,K57>E|

	IGHV3-15*05	M99403	|c279>t|
	IGHV3-15*06	M99404	|t159>c| |g196>a,D66>N|
	IGHV3-15*07	M99406	|a81>t| |g119>a,S40>N|t159>c|
	IGHV3-15*08	M99400	g23>c ,G8>A| |a40>c,K14>Q| |t89>g,F30>C| |c163>t, R55>C| |a178>g,T60>A|g181>a,D61>N| |c210>t| |a275 >t,N92>I| |a297>g|a320>g|
IGHV3-16	IGHV3-16*01	M99655	SEQ ID NO: 29 a6 |a9 |
	IGHV3-16*02	AB019440	a6>g|a9>g|
IGHV3-19	IGHV3-19*01	M99656	SEQ ID NO: 30
IGHV3-20	IGHV3-20*01	M99657	SEQ ID NO: 31
IGHV3-21	IGHV3-21*01	AB019439	SEQ ID NO: 32 g9 |
	IGHV3-	M99658	g9>a|

	21*02		
IGHV3-23	IGHV3-23*01	M99660	SEQ ID NO: 33 c164 ,A55 |a169 ,g170 ,S57 |g172 ,t174 ,G58 |a175 ,S59 |g181 ,G61 |c201 |c203 ,A68 |c237 |c243 |
	IGHV3-23*02	M35415	|c203>g,A68>G|c237>a|
	IGHV3-23*03	U29481	c164>t,A55>V|a169>t,g170>a,S57>Y|g172>a,t174>c,G58>S|a175>g,S59>G |g181>a,G61>S|c201>t| |c243>t|
	IGHV3-23*04	AJ879486	t13>g|
	IGHV3-23*05	AY757302	a154>t, g155>a, S>Y|g157>a, t159>a, G>S|g166>a, G>S|
IGHV3-30	IGHV3-30*01	M83134	SEQ ID NO: 34 g18 ,E6 |a27 |a49 ,R17 |c75 |g80 ,G27 |c101 ,t102 ,A34 |t135 |a138 |a150 |g163 ,V55 |t169 ,c170 ,a171 ,S57 |c201 |g202 ,A68 |a229 ,T 77 |c257 ,T86 |g267 |a275 ,N92 |t288 |a293 ,D98 |g317 ,R106 |
	IGHV3-30*02	L26401	|a49>g,R17>G|c75>g| |c101>g,t102>c,A34>G| |a150 >g|g163>t,V55>F|t169>c,c170>g,a171>g,S57>R|c201>t| |g317>a,R106>K|
	IGHV3-30*03	M99663	|c101>g,t102>c,A34>G| |a150>g| |c201>t|

IGHV3-30*04	L06615	|a150>g|	
IGHV3-30*05	M77323	|c101>g,t102>c,A34>G| |a293>g,D98>G|	
IGHV3-30*06	L06617	|c75>g| |c101>c,t102>c,A34>G|	
IGHV3-30*07	L06614	|t288>c|	
IGHV3-30*08	M62737	g18>c,E6>D| |g80>c,G27>A|	
IGHV3-30*09	M77300	|a229>g,T77>A| |a150>g|	
IGHV3-30*10	M77326	g202>a,A68>T| —	
IGHV3-30*11	M77331	|c75>g|	
IGHV3-30*12	M77338	|a27>g| |c75>g| |c101>g,t102>c,A34>G| |t288>c|	
IGHV3-30*13	M77339	|c101>g,t102>c,A34>G| |c257>g,T86>R|	

	IGHV3-30*14	M77324	|a150>g| |g267>t|
	IGHV3-30*15	M77327	|a275>g,N92>S|
	IGHV3-30*16	M77328	|t135>c|
	IGHV3-30*17	M77329	|a138>g|
	IGHV3-30*18	X92214	|c101>g,t102>g,A34>G| |a150>g| |g317>a,R106>K| |c201>t|
	IGHV3-30*19	L06616	|c75>g| |c101>g,t102>g,A34>G| |a150>g|
IGHV3-30-3	IGHV3-30-3*01	X92283	SEQ ID NO: 35 c75 |g317 ,R106 |
	IGHV3-30-3*02	M77302	c75>g|g317>a,R106>K|
IGHV3-33	IGHV3-33*01	AB019439	SEQ ID NO: 36 g6 |g150 |g170 ,g171 ,W57 |t177 |t212 ,V71 |t246 |a251 ,K84 |a263 ,Y88 |g317 ,R106 |

IGHV3-33*02	M99665	g6>a| |t212>c,V71>A| |a251>c,K84>T|a263>t,Y88>F|	
IGHV3-33*03	M77305	|t246>c| |g317>a,R106>K|	
IGHV3-33*04	M77335	|g150>a| |t177>c|	
IGHV3-33*05	M77334	|g170>c,g171>a,W57>S|	
IGHV3-35*01	M99666	SEQ ID NO: 37	IGHV3-35
IGHV3-38*01	M99669	SEQ ID NO: 38 c302 ,A101 |	IGHV3-38
IGHV3-38*02	AB019439	c302>t,A101>V|	
IGHV3-43*01	M99672	SEQ ID NO: 39 t32 ,V11 |a100 ,T34 |g138 |t172 ,W58 |	IGHV3-43
IGHV3-43*02	Z18901	t32>g,V11>G|a100>g,T34>A|g138>a|t172>g,W58>G|	
IGHV3-47*01	Z18900	SEQ ID NO: 40 c58 |c101 ,A34 |t149 ,L50 |t267 ,L89 |t270 ,H90 |t310 |	IGHV3-47

	IGHV3-47*02	AB019438	c58>a|c101>t,A34>V|t149>c,L50>P| |t270>a,H90>Q|	
	IGHV3-47*03	M99674	c58>a|c101>t,A34>V|t149>c,L50>P|t267>del#,L89>del#|t270>a,H90>Q|t310>g|	
IGHV3-48	IGHV3-48*01	M99675	SEQ ID NO: 41 g48 |c96 |a100 ,g101 ,c102 ,S34 |a178 ,S60 |t246 |c287 ,A96 |g303 |	
	IGHV3-48*02	AB019438	|c287>a,A96>D|	
	IGHV3-48*03	U03893	g48>a|c96>t|a100>g,g101>a,c102>a,S34>E|a178>g,S60>G|t246>c| |g303>t|	
IGHV3-49	IGHV3-49*01	M99676	SEQ ID NO: 42 g50 ,R17 |t93 |g94 ,D32 |g100 ,A34 |t124 ,F42 |a202 ,T68 |g245 ,G82 |	
	IGHV3-49*02	M99401	g50>c,R17>P|t93>g|g94>t,D32>Y|g100>c,A34>P|t124>g,F42>V|a202>g,T68>A|g245>a,G82>D|	
	IGHV3-49*03	AB019438	|a202>g,T68>A|g245>a,G82>D|	
	IGHV3-49*04	AM940220	|t124>g ,F42>V|a202>g,T68>A|g245>a,G82>D|	

		IGHV3-49*05	AM940221	|a202>g,T68>A|g245>a,G82>D|
	IGHV3-53	IGHV3-53*01	M99679	t19 ,S7 |g133 ,A45 |c210 |g315 |a318 |
		IGHV3-53*02	Z12342	SEQ ID NO: 43 t19>a,S7>T|
		IGHV3-53*03	J03617	|g133>c,A45>P|c210>t|g315>t|a318>g|
	IGHV3-64	IGHV3-64*01	M99682	SEQ ID NO: 44 g1 ,E1 |g26 ,G9 |g70 ,A24 |t198 |t201 |a205 ,N69 |t210 |c265 ,t267 ,L89 |g274 ,g275 ,G92 |c279 |t296 ,M99 |c314 ,A105 |g317 ,R106 |
		IGHV3-64*02	AB019437	|g26>a,G9>E| |a205>g,N69>D|
		IGHV3-64*03	M77298	|g70>t,A24>S|t198>c|t201>c|a205>g,N69>D|t210>a|c265>g,t267>c,L89>V|g274>a ,G92>S|c279>t|t296>c,M99>T|c314>t,A105>V|g317>a,R106>K|
		IGHV3-64*04	M77299	g1>c,E1>Q| |g70>t,A24>S|t198>c|t201>c|a205>g,N69>D|t210>a|,t267>g, |g274>a,g275>a,G92>N| |t296>c,M99>T|
		IGHV3-64*05	M77301	|g70>t,A24>S|t198>c|t201>c|a205>g,N69>D|t210>a|c265>g,,L89>V|g274>a ,G92>S|c279>t|t296>c,M99>T|c314>t,A105>V|g317>a,R106>K|

IGHV3-66	IGHV3-66*01	X92218	SEQ ID NO: 45 g24 |g37 ,V13 |a81 |g175 ,G59 |a223 |c288 |g319 |
	IGHV3-66*02	Z27504	|a223>c|c288>t|
	IGHV3-66*03	AB019437	g24>a|g37>a,V13>||a81>g|g175>t,G59>C|a223>c|c288>t|
	IGHV3-66*04	X70208	|g319>c|
IGHV3-72	IGHV3-72*01	X92206	SEQ ID NO: 46 t186 |
	IGHV3-72*02	Z29979	t186>c|
IGHV3-73	IGHV3-73*01	X70197	SEQ ID NO: 47 t21 |
	IGHV3-73*02	AB019437	t21>c|
IGHV3-74	IGHV3-74*01	L33851	SEQ ID NO: 48 c21 |g197 ,c198 ,S66 |
	IGHV3-	Z17392	c21>t|

						74*02		
						IGHV3-74*03	J00239	|g197>c,c198>g,S66>T|
		IGHV3-d				IGHV3-d*01	Z18898	SEQ ID NO: 49
		IGHV3-h				IGHV3-h*01	AJ879484	SEQ ID NO: 50
						IGHV3-h*02	AJ879485	|g303>t|
	4	IGHV4-4				IGHV4-4*01	X05713	SEQ ID NO: 51
						IGHV4-4*02	X92232	|c46>t,P16>S| >Y| |g308>a,C103
						IGHV4-4*03	X92252	|g308>a,C103>Y|
						IGHV4-4*04	X92253	|g73>a,V25>I| |g308>a, C103>Y|

	IGHV4-4*05	X92254	c20>t,S7>L| |g308>a,C103>Y|
	IGHV4-4*06	Z75355	|a234>g,I78>M|a245>c,K82>T |g308>a,C103>Y|
	IGHV4-4*07	X62112	|c46>t,P16>S|g50>a,G17>E|g70>a,A24>T| |c93>t|a97>t,g98>a,t99>c,S33>Y|a100>t,N34>Y|t120>c|g124>a,V42>I|c129>g|c136>g,a138>c,P46>A|g147>a|g163>c,a164>g,a165>t,E55>R|c172>a,a173>c,t174>c,H58>T|g207>c|a234>g,I78>M|a245>c,K82>T|g308>a,C103>Y| |a234>g,I78>M|a245>c,K82>T|g308>a,C103>Y|
IGHV4-28	IGHV4-28*01	X05714	SEQ ID NO: 52 g48 |g49 ,c51 ,D17 |c185 ,T62 |g297 |c299 ,A100 |a319 |
	IGHV4-28*02	M83133	g48>a|g49>c,c51>g,D17>Q|c185>t,T62>I|
	IGHV4-28*03	X92233	|a319>g|
	IGHV4-28*04	X56358	|g297>c|c299>g,A100>G|
	IGHV4-28*05	X92260	|c185>t,T62>I|

IGHV4-30-2	IGHV4-30-2*01	L10089	SEQ ID NO: 53 t163 ,a164 ,c165 ,Y55 |c172 ,H58 |a237 |g245 ,R82 |c288 |g291 |c300 |c315 |g319 |
	IGHV4-30-2*02	M95122	|c288>t| |c315>g|
	IGHV4-30-2*03	X92229	t163>a,a164>g,c165>t,Y55>S|c172>t,H58>Y|a237>c|g245>c,R82>T|c288>t|g291>a|c300>t|c315>g|g319>c|
	IGHV4-30-2*04	Z75351	|g245>c,R82>T| |g291>a| |c315>g|
IGHV4-30-4	IGHV4-30-4*01	Z14238	SEQ ID NO: 54 g18 ,E6 |a48 |c49 ,g51 ,Q17 |c134 ,P45 |a166 ,I56 |t228 |c288 |a291 |
	IGHV4-30-4*02	Z14239	|a48>g|c49>g,g51>c,Q17>D| |c288>a|
	IGHV4-30-4*03	X92274	|a291>g|
	IGHV4-30-4*04	X92275	g18>c,E6>D| |a166>t,I56>F|
	IGHV4-30-4*05	Z75353	|t228>c|

		IGHV4-30-4*06	Z75360	|c134>a,P45>H|
	IGHV4-31	IGHV4-31*01	L10098	SEQ ID NO: 55 a8 ,Q3 |g37 ,V13 |c69 |c84 |g100 ,G34 |a164 ,Y55 |t224 ,L75 |a237 |a244 ,c245 ,T82 |t249 |t285 |t285-c287 |a295 ,T99 |t304 ,Y102 |
		IGHV4-31*02	M99683	|c69>t| |t224>g,L75>R|
		IGHV4-31*03	Z14237	|t224>g,L75>R|
		IGHV4-31*04	M95120	a8>g,Q3>R| |t224>g,L75>R|
		IGHV4-31*05	M95121	|t224>g,L75>R| |t28 5-c287>del(3nt)#|a295>g,T99>A|
		IGHV4-31*06	X92270	|g100>a,G34>S| |t224>g,L75>R|
		IGHV4-31*07	X92271	|c84>a| |t224>g,L75>R|
		IGHV4-31*08	X92272	|t224>g,L75>R|a237>c| |t249>c |
		IGHV4-	X92273	|t224>g,L75>R| | ,c245>a,T82>K|t24

			31*09	9>c|t285>c|
		Z14235	IGHV4-31*10	|g37>t,V13>L| |a164>g,Y55>C|t224>g,L75>R| |a244>c, ,T82>P|t249>c| |t304>g,Y102>D|
	IGHV4-34	AB019439	IGHV4-34*01	SEQ ID NO: 56 a12 |g15 |c16 ,Q6 |g20 ,W7 |g25 ,A9 |t37 ,L13 |g48 |g49 ,E17 |t85 ,S29 |t88 ,F30 |a118 ,S40 |c129 |c137 ,P46 |g147 |g163 ,a165 ,E55 |a169 ,a170 ,t171 ,N57 |c172 ,H58 |a180 |t199 ,Y67 |g207 |g226 ,t2 27 ,c228 ,V76 |a234 ,I78 |c249 |c263 ,S88 |g270 ,K90 |a274 ,S92 |c2 85 |t300 |a308 ,Y103 |
		M99684	IGHV4-34*02	|g15>a|
		X92255	IGHV4-34*03	|t300>c|
		X92236	IGHV4-34*04	|t199>a,Y67>N| | ,t227>c , V76>A|
		X92237	IGHV4-34*05	|a118>t,S40>C| |t199>a,Y67>N| c137>t,P46>L| | ,t227>c ,V76>A|

IGHV4-34*06	X92256	|a274>g,S92>G| |t300>c|
IGHV4-34*07	X92258	| ,t171>c | |t300>c|
IGHV4-34*08	M95113	|t85>a,S29>T|
IGHV4-34*09	Z14241	a12>g| |c16>g,Q6>E|g20>c,W7>S|g25>c,A9>P|t37>g,L13>V|g48>a|g49>c,E17>Q| |g147>a| | c228>t | |c249>t| |c285>t|t300>c|
IGHV4-34*10	Z14242	a12>g| |c16>g,Q6>E|g20>c,W7>S|g25>c,A9>P|t37>g,L13>V| |g147>a| |g226>a, ,V76>I|a234>g,I78>M| |c263>a,S88>Y| |t300>c|
IGHV4-34*11	X05716	|t88>g,F30>V| |c129>g| |g163>t,a165>t,E55>Y|a169>t ,N57>Y|c172>t,H58>Y| a180>g|t199>a,Y67>N|g207>c| t227>c ,V76>A| |g270>c,K90>N| |t300>c|a308>g,Y103>C|

	IGHV4-34*12	X56591	| a170>t ,N57>I|	
	IGHV4-34*13	Z75356	|g207>c|	
IGHV4-39	IGHV4-39*01	AB019439	SEQ ID NO: 57 a2 ,Q1 |t56 ,L19 |c237 |g258 ,Q86 |t262 ,S88 |a291 |t300 |c319 |	
	IGHV4-39*02	X05715	|g258>c,Q86>H| |c319>g|	
	IGHV4-39*03	X92259	|t300>c|	
	IGHV4-39*04	X92297	|a291>g|	
	IGHV4-39*05	M95116	|t56>c,L19>P|	
	IGHV4-39*06	Z14236	a2>g,Q1>R| |c237>a| |t262>c,S88>P|a291>g|t300>c|	
	IGHV4-39*07	AM940222	|c237>a| |a291>g|t300>c| c319>g	

IGHV4-55	IGHV4-55*01	M99685	SEQ ID NO: 58 c44 ,P15 |c237 |a251 ,K84 |c255 ,N85 |a263 ,Y88 |c288 |c290 ,A97 |g291 |t319 |
	IGHV4-55*02	X92223	|c237>a|
	IGHV4-55*03	X92263	|c237>a| |a263>c,Y88>S|
	IGHV4-55*04	X92265	c44>t,P15>L|c237>a|
	IGHV4-55*05	X92266	c44>t,P15>L|
	IGHV4-55*06	X92267	|c255>g,N85>K| |c288>t|
	IGHV4-55*07	X92268	|a251>g,K84>R| |a263>c,Y88>S| |g291>a|
	IGHV4-55*08	X92234	|c237>a| |t319>g|
	IGHV4-55*09	X92235	|a263>c,Y88>S| |c290>t,A97>V| |t319>a|

IGHV4-59	IGHV4-59*01	AB019438	SEQ ID NO: 59 g12 |g16 ,E6 |c20 ,S7 |c25 ,P9 |g37 |g51 ,E17 |a70 ,T24 |t76 ,c77 , S26 |a88 ,I30 |c135 |c136 ,a138 ,P46 |a147 |t163 ,a164 ,t165 ,Y55 | t172 ,a173 ,c174 ,Y58 |a196 ,N66 |c207 |c228 |a234 ,I78 |a237 |c24 9 |g258 |c285 |t288 |g291 |c300 |g319 |a320 |
	IGHV4-59*02	M29812	|a88>g,I30>V|
	IGHV4-59*03	M95114	|g258>a|
	IGHV4-59*04	M95117	| ,c174>t |a196>t,N66>Y|c207>g| |a234>g,I78>M| |t288>c|g291>a|c300>t|
	IGHV4-59*05	M95118	|c135>g| ,a138>g | |t163>c,a164>g, ,Y55>R| ,c174>t |a196>t,N66>Y|c207> g| |a237>c| |t288>c|g291>a|c300>t|
	IGHV4-59*06	M95119	|t76>a ,S26>T| |c136>g, a138>t,P46>A|a147>c| ,t165>c, | |a196>t,N66>Y|c2 07>g|c228>t| |c249>t| |c285>t|t288>c|
	IGHV4-59*07	X56360	|g51>c,E17>D|

	IGHV4-59*08	X87091	g291>a| |t288>c|
	IGHV4-59*09	Z75359	|a320>g|
	IGHV4-59*10	Z14243	g12>a|g16>c,E6>Q|c20>g,S7>W|c25>g,P9>A|g37>t| |a70>g,T24>A| ,c77>a,S26>Y| |c136>g,a138>c,P46>A|a147>g|t163>c,a164>g, ,Y55>R|t172>a,a173>c, Y58>T| |a234>g,I78>M| |t288>c| |g319>t|
IGHV4-61	IGHV4-61*01	M29811	SEQ ID NO: 60 g4 ,V2 |g48 |g49 ,E17 |g88 ,V30 |g97 ,G33 |a100 ,S34 |a118 ,S40 |c 136 ,a138 ,P46 |g162 |t163 ,a164 ,Y55 |t172 ,a173 ,Y58 |c245 ,T82 | g258 ,Q86 |t288 |g289-g291 |g291 |g315 |g319 |
	IGHV4-61*02	L10097	|g48>a|g49>c,E17>Q|g88>a,V30>I| |c136>g,a138 >c,P46>A| |t163>c,a164>g,Y55>R|t172>a,a173>c,Y58>T| | t288>c| |g291>a|
	IGHV4-61*03	X92230	|g258>c,Q86>H|
	IGHV4-61*04	X92250	|g162>a| |g289-g291>del(3nt)|

	IGHV4-61*05	X56356	g4>c,V2>L| |g88>a,V30>I|g97>a,G33>S| |c245>a,T82>K| |a118>g,S40>G| |t288>c|		
	IGHV4-61*06	Z75347	|t288>c| |g315>c|		
	IGHV4-61*07	Z75348	|g319>c|		
	IGHV4-61*08	AB019437	|a100>g,S34>G|		
	IGHV4-b*01	Z12367	SEQ ID NO: 61		
	IGHV4-b*02	X56365	g70>a,A24>T|		
	IGHV5-51*01	M99686	SEQ ID NO: 62		
	IGHV5-51*02	M18806	|t116>c,I39>T| |c148>t|		
	IGHV5-51*03	X56368	c45>g|		
	IGHV5-51*04	X56367	c45>g| |t247>c,S83>P|		
		IGHV4-b			
		IGHV5-51			
	5				

		IGHV5-51*05	Z27449	|g139>a,G47>R|
	IGHV5-a	IGHV5-a*01	X92227	SEQ ID NO: 63 t21 |c148 |c225 ,H75 |c288 ,A96 |
		IGHV5-a*02	X92279	|c148>t| |c288>del#,A96>del#|
		IGHV5-a*03	X56375	t21>c|
		IGHV5-a*04	X56376	|c225>g,H75>Q|
6	IGHV6-1	IGHV6-1*01	X92224/J04097	SEQ ID NO: 64
		IGHV6-1*02	Z14223	a27>g|
7	IGHV7-4-1	IGHV7-4-1*01	L10057	SEQ ID NO: 65
		IGHV7-4-1*02	X62110	t274>a,C92>S|
		IGHV7-4-1*03	X92290	t274>a,C92>S|g278>c,c279>g,S93>T|

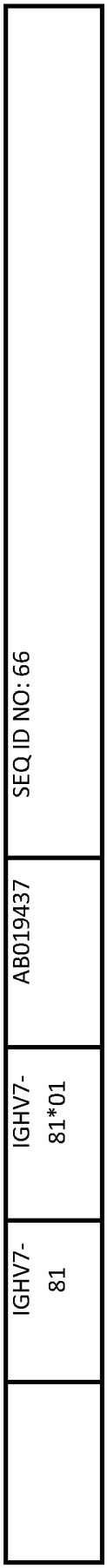

SEQ ID NO: 66
AB019437
IGHV7-81*01
IGHV7-81

Table 2: Human IgH D Polymorphic Variants

All variant sequences appearing in Genbank under the listed Accession Numbers are incorporated herein in their entirety by reference as though explicitly recited herein and for possible inclusion in any of the claims herein.

IGHD subgroup	IGHD name	Variant name	Genbank Accession number	Sequence
1	**IGHD1-1**	IGHD1-1*01	X97051	SEQ ID NO: 67
	IGHD1-7	IGHD1-7*01	X13972	SEQ ID NO: 68
	IGHD1-14	IGHD1-14*01	X13972	SEQ ID NO: 69
	IGHD1-20	IGHD1-20*01	X97051	SEQ ID NO: 70
	IGHD1-	IGHD1-	X97051	SEQ ID NO: 71

2	**26**	26*01		SEQ ID NO: 72 AGGATATTGTAGTAGTACCAGCTGCTATGCC
	IGHD2-2	IGHD2-2*01	J00232	
		IGHD2-2*02	X97051	G>A
		IGHD2-2*03	M35648	A>T
	IGHD2-8	IGHD2-8*01	X13972	SEQ ID NO: 73 AGGATATTGTACTAATGGTGTATGCTATACC
		IGHD2-8*02	J00233	AA>GG
	IGHD2-15	IGHD2-15*01	J00234	SEQ ID NO: 74
	IGHD2-21	IGHD2-21*01	J00235	SEQ ID NO: 75 AGCATATTGTGGTGGTGATTGCTATTCC
		IGHD2-	X97051	T>C

3				
IGHD3-3	IGHD3-3*01	X13972	SEQ ID NO: 76 GTATTACGATTTTTGGAGTGGTTATTATACC	
	IGHD3-3*02	X93618	CG>GC	
IGHD3-9	IGHD3-9*01	X13972	SEQ ID NO: 77	
IGHD3-10	IGHD3-10*01	X13972	SEQ ID NO: 78 GTATTACTATGGTTCGGGGAGTTATTATAAC	
	IGHD3-10*02	X93615	G>[OMITTED]	
IGHD3-16	IGHD3-16*01	X93614	SEQ ID NO: 79 GTATTATGATTACGTTTGGGGGAGTTATGCTTATACC	
	IGHD3-16*02	X97051	GC>CG	
IGHD3-	IGHD3-	X93616	SEQ ID NO: 80	
	21*02			

	22	22*01		
4	**IGHD4-4**	IGHD4-4*01	X13972	SEQ ID NO: 81
	IGHD4-11	D4-11*01	X13972	SEQ ID NO: 82
	IGHD4-17	IGHD4-17*01	X97051	SEQ ID NO: 83
	IGHD4-23	IGHD4-23*01	X97051	SEQ ID NO: 84
5	**IGHD5-5**	IGHD5-5*01	X13972	SEQ ID NO: 85
	IGHD5-12	IGHD5-12*01	X13972	SEQ ID NO: 86
	IGHD5-18	IGHD5-18*01	X97051</A<<B>	SEQ ID NO: 87
	IGHD5-24	IGHD5-24*01	X97051	SEQ ID NO: 88
6	**IGHD6-6**	IGHD6-6*01	X13972	SEQ ID NO: 89

	IGHD6-13	IGHD6-13*01	X13972	SEQ ID NO: 90
	IGHD6-19	IGHD6-19*01	X97051	SEQ ID NO: 91
	IGHD6-25	IGHD6-25*01	X97051	SEQ ID NO: 92
7	**IGHD7-27**	IGHD7-27*01	J00256	SEQ ID NO: 93

Underlined nucleotides in the *01 variant are positions that are different in other variants, eg, GC>CG for IGHD3-16*02 indicates that GC in the *01 variant is instead CG in the *02 variant. All other positions are identical to between the *01 and other variant. Similar notation is used in other tables below to denote changes between variants of gene segments.

Table 3: Human IgH J Polymorphic Variants

All variant sequences appearing in Genbank under the listed Accession Numbers are incorporated herein in their entirety by reference as though explicitly recited herein and for possible inclusion in any of the claims herein.

IGHV name	Variant name	Genbank Accession number	Sequence
IGHJ1	J1*01	J00256	SEQ ID NO: 94
IGHJ2	J2*01	J00256	SEQ ID NO: 95
IGHJ3	J3*01	J00256	SEQ ID NO: 96 T GAT GCT TTT GAT GTC TGG GGC CAA GGG ACA ATG GTC ACC GTC TCT TCA G
	J3*02	X86355	G>A
IGHJ4	J4*01	J00256	SEQ ID NO: 97 AC TAC TTT GAC TAC TGG GGC CAA GGA ACC CTG GTC ACC GTC TCC TCA G

	J4*02	X86355	AA>AG
	J4*03	M25625	AC>GC GA>GG
IGHJ5	J5*01	J00256	SEQ ID NO: 98 AC AAC TGG TTC GAC TCC TGG GGC CAA GGA ACC CTG GTC ACC GTC TCC TCA G
	J5*02	X86355	T>C A>G
IGHJ6	J6*01	J00256	SEQ ID NO: 99 AT TAC TAC TAC TAC GGT ATG GAC GTC TGG GGG CAA GGG ACC ACG GTC ACC GTC TCC TCA G
	J6*02	X86355	GG>GC
	J6*03	X86356	GGT>TAC GG>GC

			CA>AA
	J6*04	AJ8794887	GG>GC CA>AA

Table 4: Human Ig Vk Polymorphic Variants

All variant sequences appearing in Genbank under the listed Accession Numbers are incorporated herein in their entirety by reference as though explicitly recited herein and for possible inclusion in any of the claims herein.

IGKV subgroup	IGKV name	Variant name	Accession number	Sequence
1	IGKV1-5	IGKV1-5*01	Z00001	SEQ ID NO: 100 GAC ATC CAG ATG ACC CAG TCT CCT TCC ACC CTG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC CGG GCC AGT CAG AGT ATT AGT AGC TGG ... ... TTG GCC TGG TAT CAG CAG AAA CCA GGG AAA GCC CCT AAG

			CTC CTG ATC TAT GAT GCC TCC ... . AGT TTG GAA AGT GGG GTC CCA ... TCA AGG TTC AGC GGC AGT GGA ... ... TCT GGG ACA GAA TTC ACT CTC ACC ATC AGC AGC CTG CAG CCT GAT GAT TTT GCA ACT TAT TAC TGC CAA CAG TAT AAT AGT TAT TCT CC
	IGKV1-5*02	M23851	AC>AT
	IGKV1-5*03	X72813	GAT>AAG GCC>GCG TCC>TCT TTG>TTA
IGKV1-6	IGKV1-6*01	M64858	SEQ ID NO: 101
IGKV1-8	IGKV1-8*01	Z00014	SEQ ID NO: 102
IGKV1D-8	IGKV1D-8*01	Z00008	SEQ ID NO: 103
IGKV1-9	IGKV1-9*01	Z00013	SEQ ID NO: 104
IGKV1-12	IGKV1-	V01577	SEQ ID NO: 105

	12*01		GAC ATC CAG ATG ACC CAG TCT CCA TCT TCC GTG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGT CGG GCG AGT CAG GGT ATT AGC AGC TGG … … … … … … TTA GCC TGG TAT CAG CAG AAA CCA GGG AAA GCC CCT AAG CTC CTG ATC TAT GCT GCA TCC … … … … … … … AGT TTG CAA AGT GGG GTC CCA … TCA AGG TTC AGC GGC AGT GGA … … TCT GGG ACA GAT TTC ACT CTC ACC ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA ACT TAC TAT TGT CAA CAG GCT AAC AGT TTC CCT CC
IGKV1D-12	IGKV1D-12*01	X17263	TCC>TCT ACC>ACT
IGKV1-12/IGKV1D-12 (1)	IGKV1-12*02/1D-12*02	V01576	CC>TC
IGKV1-13	IGKV1-13*01	Z00010	SEQ ID NO: 106 GCC ATC CAG TTG ACC CAG TCT CCA TCC TCC CTG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC CGG GCA AGT CAG GGC ATT AGC AGT GCT … … … … … … TTA GCC TGA TAT CAG CAG AAA CCA GGG AAA GCT CCT AAG CTC CTG ATC TAT GAT GCC TCC … … … … … … … AGT TTG GAA AGT GGG GTC CCA … TCA AGG TTC AGC GGC AGT GGA … … TCT GGG ACA GAT TTC ACT CTC ACC ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA ACT TAT TAC TGT CAA CAG TTT AAT AAT TAC CCT CA

			TGA>TGG AAT>AGT	
	IGKV1-13*02	Z00006		
IGKV1D-13	IGKV1D-13*01	X17262	TGA>TGG	
IGKV1-16	IGKV1-16*01	J00248	SEQ ID NO: 107 GAC ATC CAG ATG ACC CAG TCT CCA TCC TCA CTG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGT CGG GCG AGT CAG GGC ATT AGC AAT TAT … … … … … … TTA GCC TGG TTT CAG CAG AAA CCA GGG AAA GCC CCT AAG TCC CTG ATC TAT GCT GCA TCC … … … … … … … AGT TTG CAA AGT GGG GTC CCA … TCA AGG TTC AGC GGC AGT GGA … … … TCT GGG ACA GAT TTC ACT CTC ACC ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA ACT TAT TAC TGC CAA CAG TAT AAT AGT TAC CCT CC	
	IGKV1-16*02	FM164406	AGG>AAG	
IGKV1D-16	IGKV1D-16*01	K01323	SEQ ID NO: 108 GAC ATC CAG ATG ACC CAG TCT CCA TCC TCA CTG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGT CGG GCG AGT CAG GGT ATT AGC AGC TGG … … … … … … TTA GCC TGG TAT CAG CAG AAA CCA GAG AAA GCC CCT AAG TCC CTG ATC TAT GCT GCA TCC … … … … … … … AGT TTG CAA AGT GGG GTC CCA … TCA AGG TTC AGC GGC AGT GGA … … TCT GGG ACA GAT TTC ACT CTC ACC ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA ACT TAT TAC	

				TGC CAA CAG TAT AAT AGT TAC CCT CC
	IGKV1D-16*02	V00558		AGT>AGG
IGKV1-17	IGKV1-17*01	X72808		SEQ ID NO: 109 GAC ATC CAG ATG ACC CAG TCT CCA TCC TCC CTG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC CGG GCA AGT CAG GGC ATT AGA AAT GAT … … … … … … TTA GGC TGG TAT CAG CAG AAA CCA GGG AAA GCC CCT AAG CGC CTG ATC TAT GCT GCA TCC … … … … … … … … AGT TTG CAA AGT GGG GTC CCA … TCA AGG TTC AGC GGC AGT GGA … … TCT GGG ACA GAA TTC ACT CTC ACA ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA ACT TAT TAC TGT CTA CAG CAT AAT AGT TAC CCT CC
	IGKV1-17*02	D88255		AGC> AAC
IGKV1D-17	IGKV1D-17*01	X63392		SEQ ID NO: 110 AAC ATC CAG ATG ACC CAG TCT CCA TCT GCC ATG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGT CGG GCG AGG CAG GGC ATT AGC AAT TAT … … … … … … TTA GCC TGG TTT CAG CAG AAA CCA GGG AAA GTC CCT AAG CAC CTG ATC TAT GCT GCA TCC … … … … … … … … AGT TTG CAA AGT GGG GTC CCA … TCA AGG TTC AGC GGC AGT GGA … … TCT GGG ACA GAA TTC ACT CTC ACA ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA ACT TAT TAC TGT CTA CAG CAT AAT AGT TAC CCT CC
	IGKV1D-	FM164407		AAC>GAC

	17*02		CAC>CGC
IGKV1-27	IGKV1-27*01	X63398	SEQ ID NO: 111
IGKV1-33	IGKV1-33*01	M64856	SEQ ID NO: 112
IGKV1D-33	IGKV1D-33*01	M64855	SEQ ID NO: 113
IGKV1-37	IGKV1-37*01	X59316	SEQ ID NO: 114
IGKV1D-37	IGKV1D-37*01	X71893	SEQ ID NO: 115
IGKV1-39	IGKV1-39*01	X59315	SEQ ID NO: 116 GAC ATC CAG ATG ACC CAG TCT CCA TCC TCC CTG TCT GCA TCT GTA GGA GAC AGA GTC ACC ATC ACT TGC CGG GCA AGT CAG AGC ATT AGC AGC TAT ... ... TTA AAT TGG TAT CAG CAG AAA CCA GGG AAA GCC CCT AAG CTC CTG ATC TAT GCT GCA TCC ... ... AGT TTG CAA AGT GGG GTC CCA ... TCA AGG TTC AGT GGC AGT GGA ... ... TCT GGG ACA GAT TTC ACT CTC ACC ATC AGC AGT CTG CAA CCT GAA GAT TTT GCA ACT TAC TAC TGT CAA CAG AGT TAC AGT ACC CCT CC
	IGKV1-39*02	X59318	TCC>TTC TAC>TAT

				CAA CAG AGT> CAG TGT GGT ACC>ACA
	IGKV1D-39	IGKV1D-39*01	X59312	SEQ ID NO: 117
	IGKV1D-42	IGKV1D-42*01	X72816	SEQ ID NO: 118
	IGKV1D-43	IGKV1D-43*01	X72817	SEQ ID NO: 119
2	IGKV2-24	IGKV2-24*01	X12684	SEQ ID NO: 120
	IGKV2D-24	IGKV2D-24*01	X63401	SEQ ID NO: 121
	IGKV2-28	IGKV2-28*01	X63397	SEQ ID NO: 122
	IGKV2D-28	IGKV2D-28*01	X12691	SEQ ID NO: 123
	IGKV2-29	IGKV2-29*01	X63396	SEQ ID NO: 124 GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC GTC ACC CCT GGA CAG CCG GCC TCC ATC TCC TGC AAG TCT AGT CAG AGC CTC CTG CAT AGT GAT GGA AAG ACC TAT ... TTG TAT TGG TAC CTG CAG AAG CCA GGC CAG

			TCT CCA CAG CTC CTG ATC TAT GAA GTT TCC ... ... AGC CGG TTC TCT GGA GTG CCA ... GAT AGG TTC AGT GGC AGC GGG ... ... TCA GGG ACA GAT TTC ACA CTG AAA ATC AGC CGG GTG GAG GCT GAG GAT GTT GGG GTT TAT TAC TGA ATG CAA GGT ATA CAC CTT CCT CC
	IGKV2-29*02	U41645	CTG> CTA TGA> TGC
	IGKV2-29*03	AJ783437	TGA> TGC
IGKV2D-29	IGKV2D-29*01	M31952	SEQ ID NO: 125 GAT ATT GTG ATG ACC CAG ACT CCA CTC TCT CTG TCC GTC ACC CCT GGA CAG CCG GCC TCC ATC TCC TGC AAG TCT AGT CAG AGC CTC CTG CAT AGT GAT GGA AAG ACC TAT ... TTG TAT TGG TAC CTG CAG AAG CCA GGC CAG CCT CCA CAG CTC CTG ATC TAT GAA GTT TCC ... ... AAC CGG TTC TCT GGA GTG CCA ... GAT AGG TTC AGT GGC AGC GGG ... ... TCA GGG ACA GAT TTC ACA CTG AAA ATC AGC CGG GTG GAG GCT GAG GAT GTT GGG GTT TAT TAC TGC ATG CAA AGT ATA CAG CTT CCT CC
	IGKV2D-29*02	U41644	CCT>TCT
IGKV2-30	IGKV2-30*01	X63403	SEQ ID NO: 126 GAT GTT GTG ATG ACT CAG TCT CCA CTC TCC CTG CCC GTC ACC CTT GGA CAG CCG GCC TCC ATC TCC TGC AGG TCT AGT CAA AGC CTC GTA TAC AGT GAT GGA AAC ACC TAC ... TTG AAT TGG TTT CAG CAG AGG CCA GGC CAA

			TCT CCA AGG CGC CTA ATT TAT AAG GTT TCT ... ... AAC CGG GAC TCT GGG GTC CCA ... GAC AGA TTC AGC GGC AGT GGG ... ... TCA GGC ACT GAT TTC ACA CTG AAA ATC AGC AGG GTG GAG GCT GAG GAT GTT GGG GTT TAT TAC TGC ATG CAA GGT ACA CAC TGG CCT CC
	IGKV2-30*02	FM164408	TAC>CAC
IGKV2D-30	IGKV2D-30*01	X63402	SEQ ID NO: 127
IGKV2-40	IGKV2-40*01	X59314	SEQ ID NO: 128 GAT ATT GTG ATG ACC CAG ACT CCA CTC TCC CTG CCC GTC ACC CCT GGA GAG CCG GCC TCC ATC TCC TGC AGG TCT AGT CAG AGC CTC TTG GAT AGT GAT GAT GGA AAC ACC TAT TTG GAC TGG TAC CTG CAG AAG CCA GGG CAG TCT CCA CAG CTC CTG ATC TAT ACG CTT TCC ... ... TAT CGG GCC TCT GGA GTC CCA ... GAC AGG TTC AGT GGC AGT GGG ... ... TCA GGC ACT GAT TTC ACA CTG AAA ATC AGC AGG GTG GAG GCT GAG GAT GTT GGA GTT TAT TAC TGC ATG CAA CGT ATA GAG TTT CCT TC
	IGKV2-40*02	X59317	GAC TGG>GAT TGT GGC>GAC
IGKV2D-40	IGKV2D-40*01	X59311	SEQ ID NO: 129

3	IGKV3-7	IGKV3-7*01	X02725	SEQ ID NO: 130 GAA ATT GTA ATG ACA CAG TCT CCA CCC ACC CTG TCT TTG TCT CCA GGG GAA AGA GTC ACC CTC TCC TGC AGG GCC AGT CAG AGT GTT AGC AGC AGC TAC ... ... TTA ACC TGG TAT CAG CAG AAA CCT GGC CAG GCG CCC AGG CTC CTC ATC TAT GGT GCA TCC ... ... ACC AGG GCC ACT AGC ATC CCA ... GCC AGG TTC AGT GGC AGT GGG ... ... TCT GGG ACA GAC TTC ACT CTC ACC ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA GTT TAT TAC TGT CAG CAG GAT CAT AAC TTA CCT CC
		IGKV3-7*02	X72812	ACC>TCC TAT>TAC GGC>GGG GCG>GCT AGC>GGC ACA>AGA CAT>TAT
		IGKV3-7*03	K02769	SEQ ID NO: 131 GAA ATT GTA ATG ACA CAG TCT CCA CCC ACC CTG TCT TTG TCT CCA GGG GAA AGA GTC ACC CTC TCC TGC AGG GCC AGT CAG AGT GTT AGC AGC AGC TAC ... ... TTA ACC TGG TAT CAG CAG AAA CCT GGC CAG GCG CCC AGG CTC CTC ATC TAT GGT GCA TCC ... ... ACC AGG GCC ACT AGC ATC CCA ... GCC AGG TTC AGT GGC AGT GGG ... ... TCT GGG ACA GAC TTC ACT CTC ACC ATC AGC AGC CTG CAG CCT GAA GAT TTT GCA GTT TAT

				TAC TGT CAG CAG GAT CAT AAC TTA CCT CC
		IGKV3-7*04	FM164409	CAT>TAT
	IGKV3D-7	IGKV3D-7*01	X72820	SEQ ID NO: 132
	IGKV3-11	IGKV3-11*01	X01668	SEQ ID NO: 133 GAA ATT GTG TTG ACA CAG TCT CCA GCC ACC CTG TCT TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC AGG GCC AGT CAG AGT GTT AGC AGC TAC ... ... TTA GCC TGG TAC CAA CAG AAA CCT GGC CAG GCT CCC AGG CTC CTC ATC TAT GAT GCA TCC ... ... AAC AGG GCC ACT GGC ATC CCA ... GCC AGG TTC AGT GGC AGT GGG ... ... TCT GGG ACA GAC TTC ACT CTC ACC ATC AGC AGC CTA GAG CCT GAA GAT TTT GCA GTT TAT TAC TGT CAG CAG CGT AGC AAC TGG CCT CC
		IGKV3-11*02	K02768	ACA>AGA
	IGKV3D-11	IGKV3D-11*01	X17264	SEQ ID NO: 134
	IGKV3-15	IGKV3-15*01	M23090	SEQ ID NO: 135
	IGKV3D-15	IGKV3D-	X72815	SEQ ID NO: 136

			GAA ATA GTG ATG ACG CAG TCT CCA GCC ACC CTG TCT GTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC AGG GCC AGT CAG AGT GTT AGC AGC AAC … … … … … … TTA GCC TGG TAC CAG CAG AAA CCT GGC CAG GCT CCC AGG CTC CTC ATC TAT GGT GCA TCC … … … … … … … … ACC AGG GCC ACT GGC ATC CCA … GCC AGG TTC AGT GGC AGT GGG … … TCT GGG ACA GAG TTC ACT CTC ACC ATC AGC AGC CTG CAG TCT GAA GAT TTT GCA GTT TAT TAC TGT CAG CAG TAT AAT AAC TGG CCT CC
15*01	IGKV3D-15*02	M23091	ACG>ATG TGG>TGA
	IGKV3-20*01	X12686	SEQ ID NO: 137 GAA ATT GTG TTG ACG CAG TCT CCA GGC ACC CTG TCT TTG TCT CCA GGG GAA AGA GCC ACC CTC TCC TGC AGG GCC AGT CAG AGT GTT AGC AGC AGC TAC … … … … … … TTA GCC TGG TAC CAG CAG AAA CCT GGC CAG GCT CCC AGG CTC CTC ATC TAT GGT GCA TCC … … … … … … … … AGC AGG GCC ACT GGC ATC CCA … GAC AGG TTC AGT GGC AGT GGG … … TCT GGG ACA GAC TTC ACT CTC ACC ATC AGC AGA CTG GAG CCT GAA GAT TTT GCA GTG TAT TAC TGT CAG CAG TAT GGT AGC TCA CCT CC
	IGKV3-20*01, Tou-kv325	X56593	ACC>CCC CTG>TGT TCT> CTT TTG> TGT TCT> CTC
IGKV3-20			

		L37729	ACG>ACA GGC>GCC GAC>GCA	
	IGKV3-20*02			
	IGKV3D-20	IGKV3D-20*01	X12687	SEQ ID NO: 138
4	IGKV4-1	IGKV4-1*01	Z00023	SEQ ID NO: 139
5	IGKV5-2	IGKV5-2*01	X02485	SEQ ID NO: 140
6	IGKV6-21	IGKV6-21*01	X63399	SEQ ID NO: 141
	IGKV6D-21	IGKV6D-21*01	X12683	SEQ ID NO: 142
	IGKV6D-41	IGKV6D-41*01	M27751 (3)	SEQ ID NO: 143

Table 5: Human Ig VA Polymorphic Variants

All variant sequences appearing in Genbank under the listed Accession Numbers are incorporated herein in their entirety by reference as though explicitly recited herein and for possible inclusion in any of the claims herein.

IGLV subgroup	**IGLV Gene name**	**IGLV variant name**	**Accession number**	**Sequences & Description of mutations**
1	IGLV1-36	V1-36*01	Z73653	SEQ ID NO: 144
	IGLV1-40	V1-40*01	M94116	SEQ ID NO: 145
		V1-40*02	X53936	g9>c|c10>g,L4>V|
		V1-40*03	Z22192	g9>c|c10>g,L4>V|a253>g,T85>A|
	IGLV1-41	V1-41*01	M94118	SEQ ID NO: 146
		V1-41*02	D87010	g295>t,E99>*|c332>t,P111>L|

	IGLV1-44	V1-44*01	Z73654	SEQ ID NO: 147	
	IGLV1-47	V1-47*01	Z73663	SEQ ID NO: 148	
		V1-47*02	D87016	g168>t,R56>S|	
	IGLV1-50	V1-50*01	M94112	SEQ ID NO: 149	
	IGLV1-51	V1-51*01	Z73661	SEQ ID NO: 150	
		V1-51*02	M30446	t162>c|c168>a,D56>E|	
2	IGLV2-8	V2-8*01	X97462	SEQ ID NO: 151	
		V2-8*02	L27695	g37>a,G13>R|	

	V2-8*03	Y12418	|c230>t,S77>F|		
IGLV2-11	V2-11*01	Z73657	SEQ ID NO: 152		
	V2-11*02	Z22198	t96>g|		
	V2-11*03	Y12415	|g132>a|		
IGLV2-14	V2-14*01	Z73664	SEQ ID NO: 153		
	V2-14*02	L27822	c87>t|t93>g|g94>a,G32>S|t103>c,a104>t,Y35>L| |t170>g,V57>G|t198>g,N66>K|		
	V2-14*03	Y12412	|g132>a|g168>t,E56>D|		
	V2-14*04	Y12413	|g168>t,E56>D|		
IGLV2-	V2-18*01	Z73642	SEQ ID NO: 154		

	18			
		V2-18*02	L27697	|t317>c,L106>S|
		V2-18*03	L27694	|t272>c,I91>T|t317>c,L106>S|
		V2-18*04	L27692	t227>c,F76>S|t249>c| |t317>c,L106>S|
	IGLV2-23	V2-23*01	X14616	SEQ ID NO: 155
		V2-23*02	Z73665	g170>t,G57>V|a339>c,L113>F|
		V2-23*03	D86994	|a339>c,L113>F|
	IGLV2-33	V2-33*01	Z73643	SEQ ID NO: 156
		V2-33*02	L27823	a3>g|

3	IGLV3-1	V2-33*03	L27691	|t96>c|a256>g,M86>V|
		V3-1*01	X57826	SEQ ID NO: 157
	IGLV3-9	V3-9*01	X97473	a52 ,T18 |a82 ,I28 |a88 ,g89 ,S30 |a95 ,N32 |g173 ,S58 |
		V3-9*02	X74288	a52>g,T18>A|a82>c,I28>L|a88>t,g89>a,S30>Y|a95>g ,N32>S |g173>a,S58>N|
		V3-9*03	X51754	a52>g,T18>A|a82>c,I28>L|a88>t,g89>a,S30>Y|a95>del#,N32>del# |g173>a,S58>N|
	IGLV3-10	V3-10*01	X97464	SEQ ID NO: 158
		V3-10*02	L29166	g166>a,E56>K|c207>a|t270>c|c299>a,A100>D|a319>g,T107>A|a325> t,g326>a,S109>Y|
	IGLV3-12	V3-12*01	Z73658	SEQ ID NO: 159
		V3-12*02	D86998	a259>g,T87>A|

IGLV3-16	V3-16*01	X97471	SEQ ID NO: 160
IGLV3-19	V3-19*01	X56178	SEQ ID NO: 161
IGLV3-21	V3-21*01	X71966	SEQ ID NO: 162
	V3-21*02	D87007	a27>g|a49>c,K17>Q|a160>g,I54>V|t166>g,Y56>D|c324>t|
	V3-21*03	M94115	a27>g| |a160>g,I54>V|t166>g,Y56>D|c324>t|
IGLV3-22	V3-22*01	Z73666	SEQ ID NO: 163
	V3-22*02	X71967	c53>a,T18>K|g88>a,E30>K|g140>del#,G47>del#|c148>t,c149>g,t150>a,P50>*|150^151>ins^tatac#,50^51>ins^Y#|g322>a,D108>N|c330>t|
IGLV3-25	V3-25*01	X97474	SEQ ID NO: 164

		V3-25*02	D86994	t14>c,g15>a,M5>T|	
		V3-25*03	L29165	t14>c,g15>a,M5>T|t294>c|	
	IGLV3-27	V3-27*01	D86994	SEQ ID NO: 165	
	IGLV3-32	V3-32*01	Z73645	SEQ ID NO: 166	
4	IGLV4-3	V4-3*01	X57828	SEQ ID NO: 167	
	IGLV4-60	V4-60*01	Z73667	SEQ ID NO: 168	
		V4-60*02	D87000	a288>t,L96>F|	
		V4-60*03	AF073885	t287>c,a288>t,L96>S|	
	IGLV4-	V4-69*01	Z73648	SEQ ID NO: 169	

					c198>t|
69	V4-69*02	U03868			
IGLV5-37	V5-37*01	Z73672			SEQ ID NO: 170
IGLV5-39	V5-39*01	Z73668			t81 ,g135>t|
	V5-39*02	AF216776			t81>c,g135>t|
IGLV5-45	V5-45*01	Z73670			SEQ ID NO: 171
	V5-45*02	Z73671			g22>t,A8>S|g75>a|
	V5-45*03	D86999			g22>t,A8>S|
IGLV5-48	V5-48*01	Z73649			SEQ ID NO: 172
5					

	IGLV5-52	V5-52*01	Z73669	SEQ ID NO: 173
6	IGLV6-57	V6-57*01	Z73673	SEQ ID NO: 174
7	IGLV7-43	V7-43*01	X14614	SEQ ID NO: 175
	IGLV7-46	V7-46*01	Z73674	SEQ ID NO: 176
		V7-46*02	D86999	|c275>t,S92>L|
		V7-46*03	Z22210	c271>del#|L91>del#|
8	IGLV8-61	V8-61*01	Z73650	SEQ ID NO: 177
		V8-61*02	U03637	|c223>t,R75>C|
		V8-61*03	AF266511	g210>c|

				SEQ ID NO: 178	
9	IGLV9-49	V9-49*01	Z73675		
		V9-49*02	Z73675	|g291>a|	
		V9-49*03	U03869	g153>a|	
				SEQ ID NO: 179	
10	IGLV10-54	V10-54*01	Z73676		
		V10-54*02	D86996	|a86>t,N29>I|	|a228>c,L76>F|g320>t,W107>L|
		V10-54*03	S70116	g33>c|	|t125>c,g126>t,L42>P|c127>g,Q43>E|
				SEQ ID NO: 180	
11	IGLV11-55	V11-55*01	D86996		

Table 6: Human IgH Jk Polymorphic Variants

All variant sequences appearing in Genbank under the listed Accession Numbers are incorporated herein in their entirety by reference as though explicitly recited herein and for possible inclusion in any of the claims herein.

IGKJ gene name	**Variant name**	**Accession number**	**Sequence**
IGKJ1	IGKJ1*01	J00242	SEQ ID NO: 181
IGKJ2	IGKJ2*01	J00242	SEQ ID NO: 182 TG TAC ACT TTT GGC CAG GGG ACC AAG CTG GAG ATC AAA C
	IGKJ2*02	Z70260 #c	TAC>TGC
	IGKJ2*03	U95246 #c	ACT>AGT

	IGKJ2*04	L40735 #c	TAC>TGC ACT>AGT
IGKJ3	IGKJ3*01	J00242	SEQ ID NO: 183
IGKJ4	IGKJ4*01	J00242	SEQ ID NO: 184 G CTC ACT TTC GGC GGA GGG ACC AAG GTG GAG ATC AAA C
	IGKJ4*02	AF103571 #c	ACT>ACG
IGKJ5	IGKJ5*01	J00242	SEQ ID NO: 185

Table 7: Human IgH Jλ Polymorphic Variants

All variant sequences appearing in Genbank under the listed Accession Numbers are incorporated herein in their entirety by reference as though explicitly recited herein and for possible inclusion in any of the claims herein.

IGLJ gene name	Variant name	Accession number	Sequence
IGLJ1	J1*01	X04457	SEQ ID NO: 186
IGLJ2	J2*01	M15641	SEQ ID NO: 187
IGLJ3	J3*01	M15642	SEQ ID NO: 188 T GTG GTA TTC GGC GGA GGG ACC AAG CTG ACC GTC CTA G
	J3*02	D87023(1)	GTG> TGG GTA> GTG
IGLJ4	J4*01	X51755	SEQ ID NO: 189
IGLJ5	J5*01	X51755	SEQ ID NO: 190 C TGG GTG TTT GGT GAG GGG ACC GAG CTG ACC GTC CTA G
	J5*02	D87017	ACC>ACG

IGLJ6	J6*01	M18338	SEQ ID NO: 191
IGLJ7	J7*01	X51755	SEQ ID NO: 192 T GCT GTG TTC GGA GGA GGC ACC CAG CTG ACC GTC CTC G
	J7*02	D87017	GTC>GCC

TABLE 8: 1000 GENOMES PROJECT HUMAN POPULATIONS

Below is a summary of the ethnic population origin of samples that the 1000 Genomes Project sequences.

Population

European ancestry

Utah residents (CEPH) with Northern and Western European ancestry (CEU)

Toscani in Italia (TSI)

British from England and Scotland (GBR)

Finnish from Finland (FIN)

Iberian populations in Spain (IBS)

East Asian ancestry

Han Chinese in Beijing, China (CHB)

Japanese in Toyko, Japan (JPT)

Han Chinese South (CHS)

Chinese Dai in Xishuangbanna (CDX)

Kinh in Ho Chi Minh City, Vietnam (KHV)

Chinese in Denver, Colorado (CHD) (pilot 3 only)

West African ancestry

Yoruba in Ibadan, Nigeria (YRI)

Luhya in Webuye, Kenya (LWK)

Gambian in Western Division, The Gambia (GWD)

Malawian in Blantyre, Malawi (MAB)

K00004-1 WO

Population

West African Population (TBD)

Americas

African Ancestry in Southwest US (ASW)

African American in Jackson, MS (AJM)

African Caribbean in Barbados (ACB)

Mexican Ancestry in Los Angeles, CA (MXL)

Puerto Rican in Puerto Rico (PUR)

Colombian in Medellin, Colombia (CLM)

Peruvian in Lima, Peru (PEL)

South Asian ancestry

Ahom in the State of Assam, India

Kayadtha in Calcutta, India

Reddy in Hyderabad, India

Maratha in Bombay, India

Punjabi in Lahore, Pakistan

Table 9: Immunoglobulin Gene Usage in Human Antibody Responses to Infectious Disease Pathogens

KAPPA V GENES • Vk II germline gene A2 + JK3 • Vk II family gene + JK4 • 94% identical to the A27 (Humkv325) germ line gene • a VκI gene family member; κI-15A (KL012) LAMBDA V GENES • Four Vλ VII family members that are 96-98% identical to each other • Vλ II family members (82, 89 and 91% homologous to Vλ2.1 gene) + VHIII segments closely homologous to germline gene 9.1 • V_λVII 4A • All with Jλ homologous to germline Jλ2 and Jλ3 VH GENES • VH 96% identical to the VH germ line gene segment DP77 (V3-21)	Haemophilus influenzae type b polysaccharide (Hib PS)	Haemophilus influenzae	1. Lonberg, Nat Biotech 2005; [human PBMCs] 2. Adderson et al, J Clin Invest 1992; [Human PBLs] 3. Chung et al, J Immunol 1993 4. Nadel et al, J Immunol 1998 5. Feeney et al, J Clin Invest 1996 6. Lucas et al, Infect Immun 1994; [Human PBLs] 7. Adderson et al, J Clin Invest 1993; [Human PBLs] 8. Granoff et al, J Clin Invest 1993; [human PBLs] 9. Azmi et al, Infect Immun 1994; [human tonsil cells]

• LSG6.1, LSG12.1, V_HIII VH26, V_HIII 9.1 <u>VH and VL COMBINATIONS</u> • V_HIII 9.1+ V_λVII 4A • V_HIII 9.1 + V_λII 2.1 • V_HIII 9.1 + V_κII A2 • V_HIII VH26 + V_λII 2.1			
• V_HIII 9.1; V_HIII H11; V_HIII VH26 • κI 15A • Vλ2.1	• Polysaccharide capsule of E coli K1 • Meningococcal B polysaccharide; Poly[α(2→8)-N-acetylneuramic acid	• E coli K1 • Neisseria meningitidis Group B	9. Azmi et al, Infect Immun 1994
• VHIII or VHIV family member • VλI or Vλ3 member • VH26 + Dk1 + JH6 with IGLV1S2 + Jλ2 • VH4.18 • VH2-1 (VH3) + D region Dxp'1 + JH5 with Vλ3 cML70 + Jλ3 • VH1GRR+ JH3 + Dn4r or D2r with IGLV1S2 + Jλ2	• HSV 120-kD glycoprotein • 116-, 105-, 64-kD glycoproteins of VZV	• Herpes family virus • Herpes simplex virus (HSV); HSV-1; HSV-2 • Varicella zoster virus (VZV)	10. Huang et al, J Clin Invest 1992; [human tonsils]

• For VZV Abs: ha3h2 (VH3) with la1h2 (Vλ); or ha1c1 (VH1) with la1vl (Vλ1) • For VZV Abs: ha4h3 (VH4) with la3h3 (Vλ3)			
• Hv1051 (VH) • Kv325 (Vk)		• Cytomegalovirus (CMV)	10. Huang et al, J Clin Invest 1992;
• 71-2 (VH) • Hv1f10 (VH) • VH4.11 • 71-4 (VH) • VH251 • VH1-69		• HIV	10. Huang et al, J Clin Invest 1992; 11. Wang & Palese, Science 2011
• VH1-69	Haemagglutinin (HA)	• Influenza virus, eg, Group 1 and/or Group 2 Infulenza A virus; eg, H1N1, H2N2, or H3N2 or H7N2 or H7N7 influenza virus	12. Ekiert et al, Science 2009 13. Throsby et al, PLoS One 2008 14. Sui et al, Nat Struct Mol Biol 2009 15. Ekiert et al, Science 2011

Table 10A:Human IgH JH5 Variant Occurrences

HUMAN POPULATION															
NA19835	NA12814; NA12763 and 1 other(s)	NA18617; NA18636 and 2 other(s)	HG00683; HG00590	HG01359	HG00189	HG00261; HG00136		NA18987	NA19448; NA19331		HG00740	NA20799		0.017	
	NA11994		HG00595; HG00584		HG00267; HG00276	HG00137			NA19456; NA19332			NA20542	NA18934; NA19147 and 1 other(s)	0.011	
				HG01494	HG00351	HG00234		NA19005		NA19719				0.005	
NA19921	NA12815		HG00650		HG00177								NA18517	0.005	
	NA12829; NA12890 and 1 other(s)				HG00358							NA20503	NA19093	0.005	

	NA12874		HG00312			HG00249				NA19391															0.004
	NA10851; NA12154																								0.003
																								NA19248	
																									0.003
							HG01521			NA19466														NA19095	
	NA07346	NA18610																							0.003
																									0.003
																								NA18867	
																									0.003
																									0.002
																									0.002
	NA06986																								0.002
																									0.002
	NA12830																								
																									0.002
NA19707	NA11918																								
																									0.002
																									0.001
																									0.001

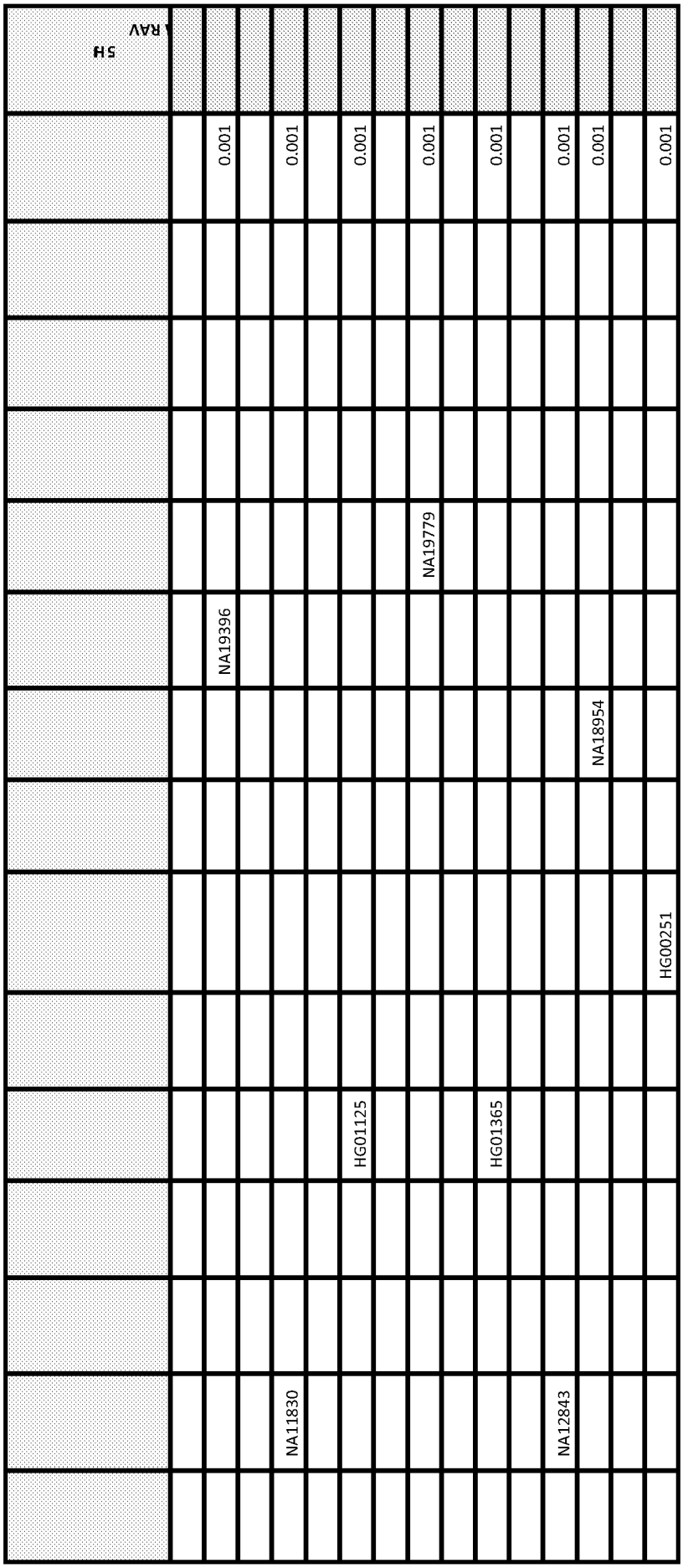
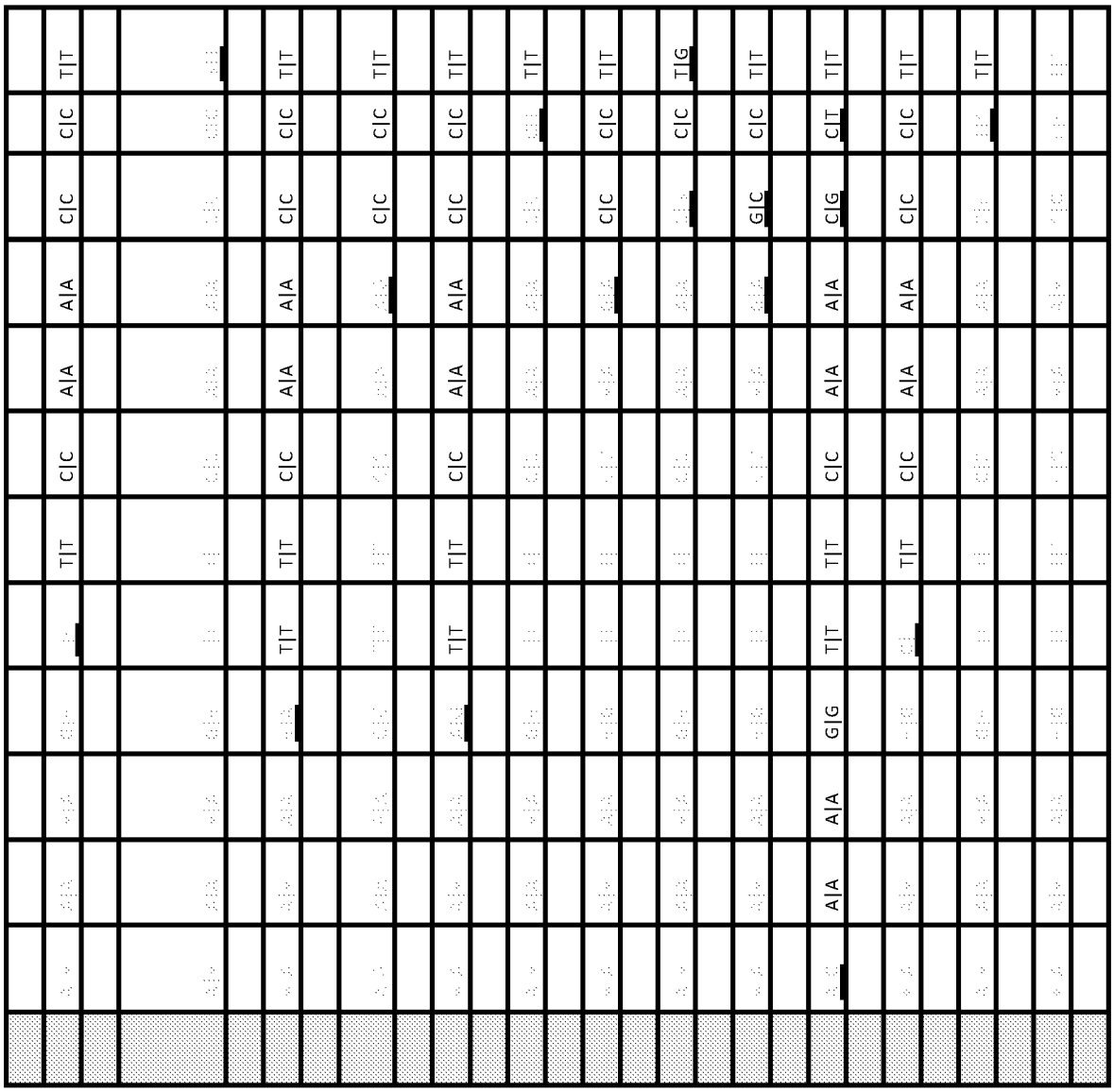

5H RAV
0.001
NA19396
NA19779
NA18954
HG00251
HG01125
HG01365
NA11830
NA12843

Table 10B:Non-Synonymous Human IgH JH5 Variants

A/S	GM DOC_SUOMYONVS_NON	67488400001			1	1
D/V	GM DOC_SUOMYONVS_NON	67488400001			1	1
Q/L	GM DOC_SUOMYONVS_NON	67488400001			1	1
I/T	GM DOC_SUOMYONVS_NON	67488400001			1	1
A/T	GM DOC_SUOMYONVS_NON	67488400001			1	1
G/D	GM DOC_SUOMYONVS_NON	67488400001			1	1
H/D	GM DOC_SUOMYONVS_NON	67488400001			1	1
C/F	GM DOC_SUOMYONVS_NON	67488400001			1	1
L/F	GM DOC_SUOMYONVS_NON	67488400001			1	1
C/W	GM DOC_SUOMYONVS_NON	67488400001			1	F
C/*	WDEMAG_POTS	67488400001			1	F
T/N	GM DOC_SUOMYONVS_NON	67488400001			1	1



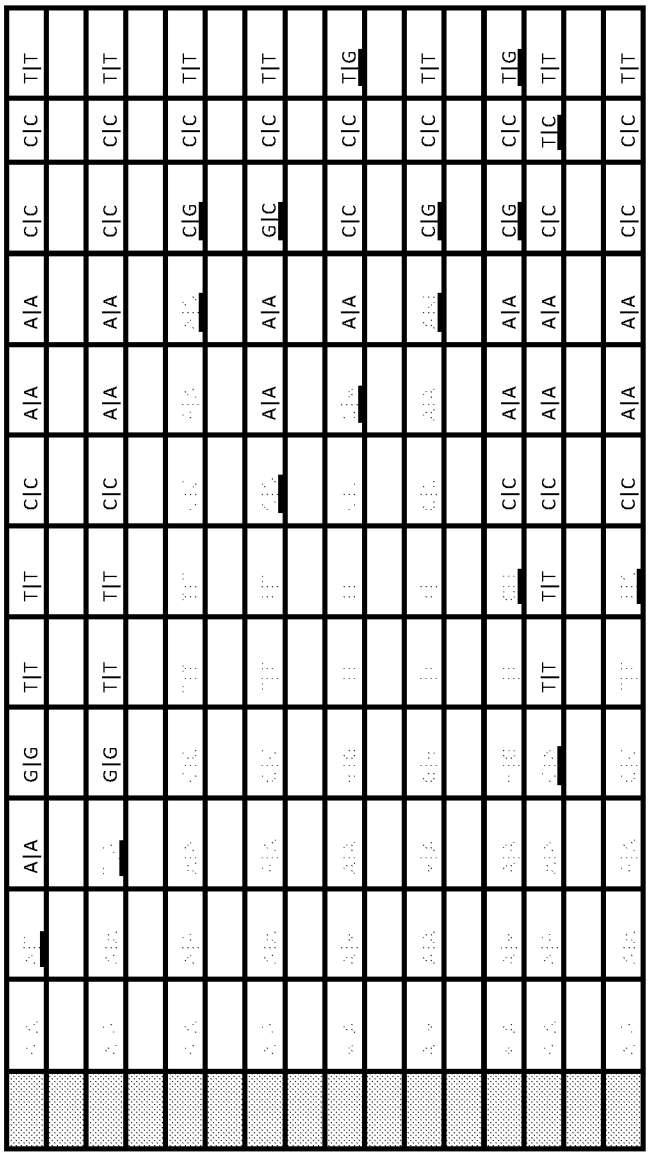


Table 11A:Human IgH JH6 Variant Occurrences

HUMAN POPULATION															
															6H
NA20289; NA20294 and 19 other(s)	NA12546; NA12046 and 41 other(s)	NA18532; NA18562 and 58 other(s)	HG00628; HG00690 and 57 other(s)	HG01456; HG01275 and 33 other(s)	HG00278; HG00313 and 57 other(s)	HG00259; HG00129 and 53 other(s)	HG01619; HG01626 and 7 other(s)	NA19083; NA18986 and 48 other(s)	NA19395; NA19321 and 20 other(s)	NA19777; NA19795 and 48 other(s)	HG00640; HG01167 and 22 other(s)	NA20798; NA20513 and 58 other(s)	NA19172; NA18909 and 8 other(s)	0.509	
NA20127; NA20414 and 11 other(s)	NA12348; NA10851 and 17 other(s)	NA18558; NA18579 and 9 other(s)	HG00464; HG00683 and 16 other(s)	HG01495; HG01550 and 7 other(s)	HG00358; HG00362 and 12 other(s)	HG00135; HG00245 and 6 other(s)	HG01620; HG01625 and 1 other(s)	NA18942; NA19004 and 11 other(s)	NA19038; NA19451 and 17 other(s)	NA19725; NA19773 and 6 other(s)	HG01108; HG01072 and 12 other(s)	NA20787; NA20754 and 14 other(s)	NA19116; NA18517 and 22 other(s)	0.173	
NA20314; NA20317 and 11 other(s)	NA06989; NA12044 and 10 other(s)	NA18608; NA18624 and 16 other(s)	HG00436; HG00404 and 13 other(s)	HG01357; HG01461 and 10 other(s)	HG00188; HG00282 and 10 other(s)	HG00116; HG00102 and 16 other(s)	HG01518	NA18974; NA18978 and 17 other(s)	NA19372; NA19474 and 16 other(s)	NA19758; NA19749 and 4 other(s)	HG01191; HG01048 and 7 other(s)	NA20522; NA20815 and 14 other(s)	NA18520; NA19209 and 16 other(s)	0.171	
	NA12827		HG00584		HG00336									0.003	
									NA19471; NA19401				NA18868	0.003	
													NA19121; NA19175	0.002	
	NA07346								NA19376					0.002	
													NA18934; NA18856	0.002	
		NA18637												0.001	

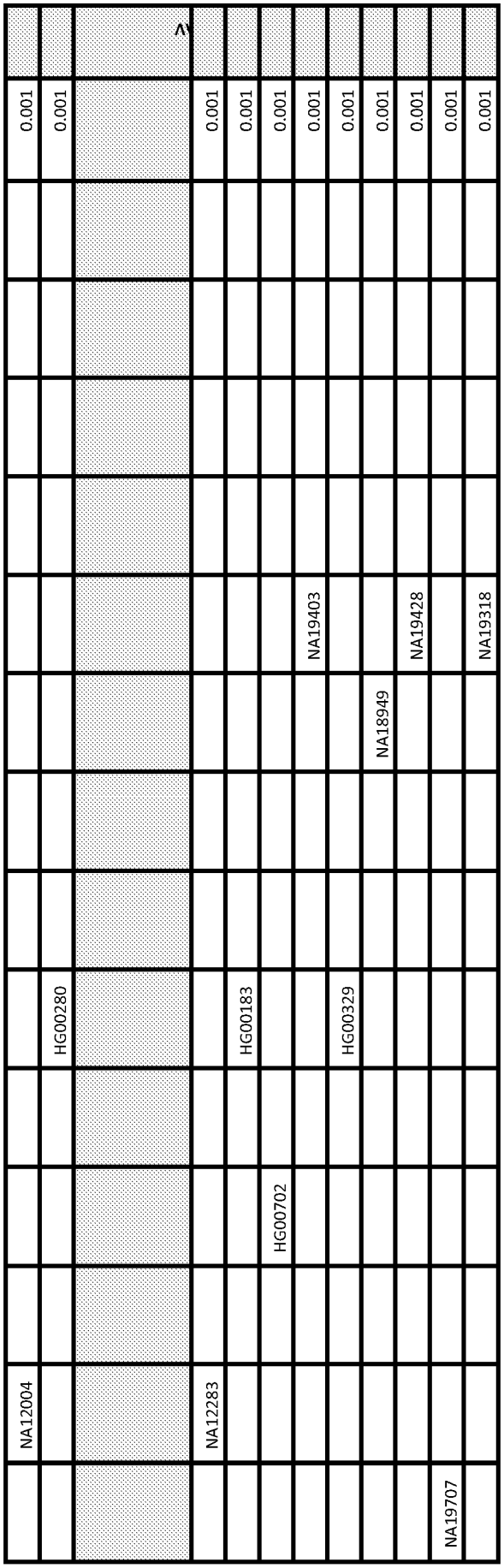

NA12004
HG00280
HG00183
HG00702
HG00329
NA18949
NA19403
NA19428
NA19318
NA19707
NA12283
0.001

Table 11B:

(i) Non-Synonymous Human IgH JH6 Variants

T/S	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
F/S	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
P/S	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
/V	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
I/T	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
S/F	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
R/K	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
Q/K	SNP	CDS_SYNONYMOUS_CODON	1	1	T				
H/Y	SNP	CDS_SYNONYMOUS_CODON	1	1	T				

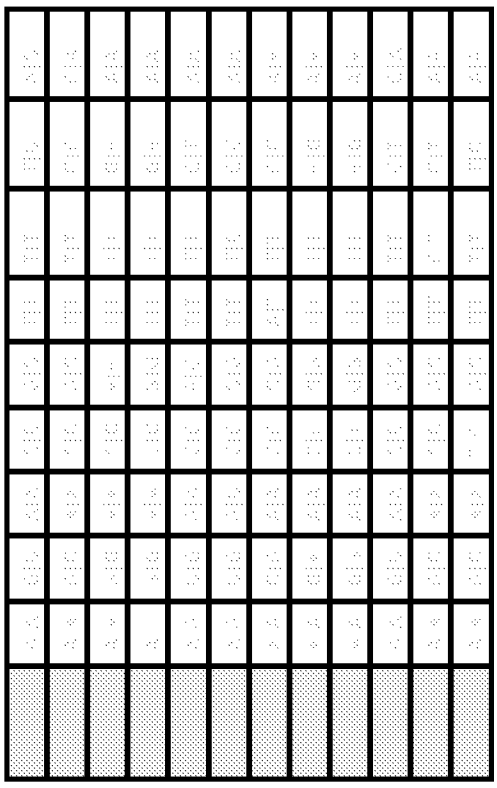
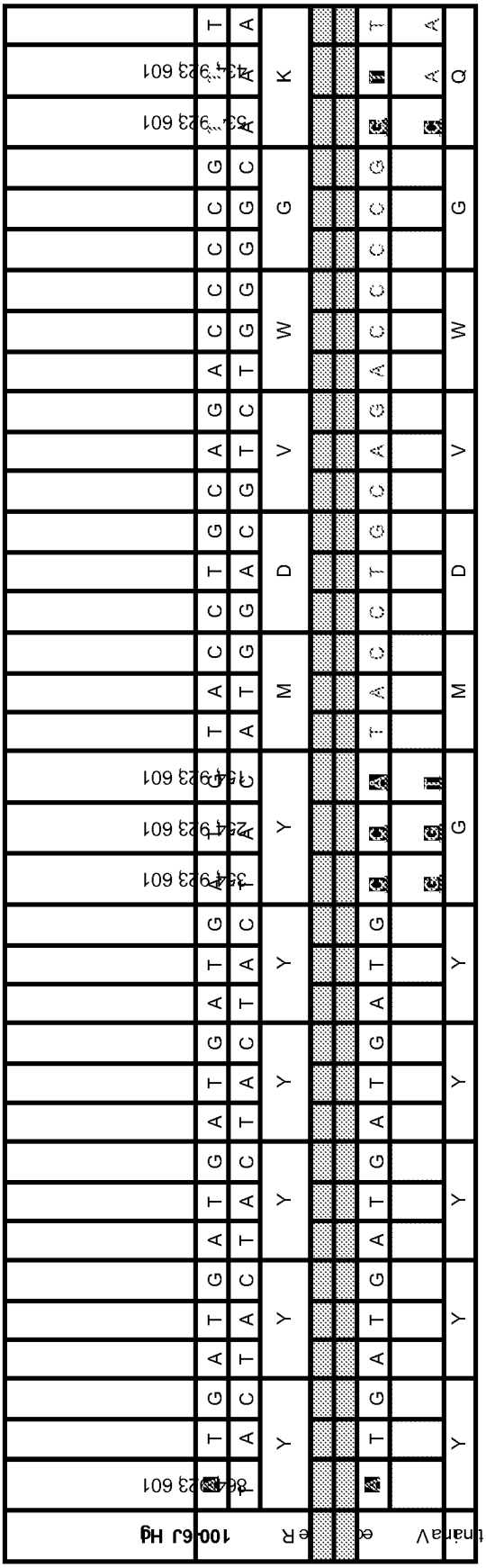


Table 11B(ii): A further variant (not listed in Table11B(i)) is the following:-



						624 923 601							914 923 601		714 923 601			414 923 601	314 923 601		114 923 601			304 923 601
C	C	C	T	G	G	[illegible]	G	C	C	A	G	T	[illegible]	G	[illegible]	A	G	[illegible]	[illegible]	G	[illegible]	G	T	[illegible]
G	G	G	A	C	C	A	C	G	G	T	C	A	C	C	G	T	C	T	C	C	T	C	A	G
G			T			T			V			T			V			S			S			
C	C	C	T	G	G	[illegible]	G	C	C	A	G	T	[illegible]	G	[illegible]	A	G	A	[illegible]	G	[illegible]	G	T	C
G			T			T			V			T			V			S			S			

SEQ ID NO: 4 (Nucleotide sequence of JH6 variant)

A T G A T G A T G A T G A T G C C A T A

C C T G C A G A C C C C G G T T C C C T
G G T G C C A G T G G C A G A G G A G T
C

Table 12A:Human IgH JH2 Variant Occurrences

HUMAN POPULATION												
2H AV EPYT												
0.004		NA12812				NA19351		NA18502; NA19172				NA19922
0.002						NA19443	HG01073					
0.002				HG01345				NA20758				
0.002			HG00584						NA19248			
0.001						NA19324						
0.001					HG00383							
0.001										NA19185		
0.001												

Table 12B:Non-Synonymous Human IgH JH2 Variants

						2H	
X/X NOD_LATRAP 46509300000TS NE						EPYT TNA RAV	1
Y GNI DOC_SUO MY NONVSL NOW46509300000TS NE						1	1
WGNI DOC_SUO MY NONVSL NOW46509300000TS NE						1	1

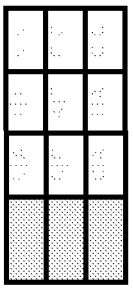
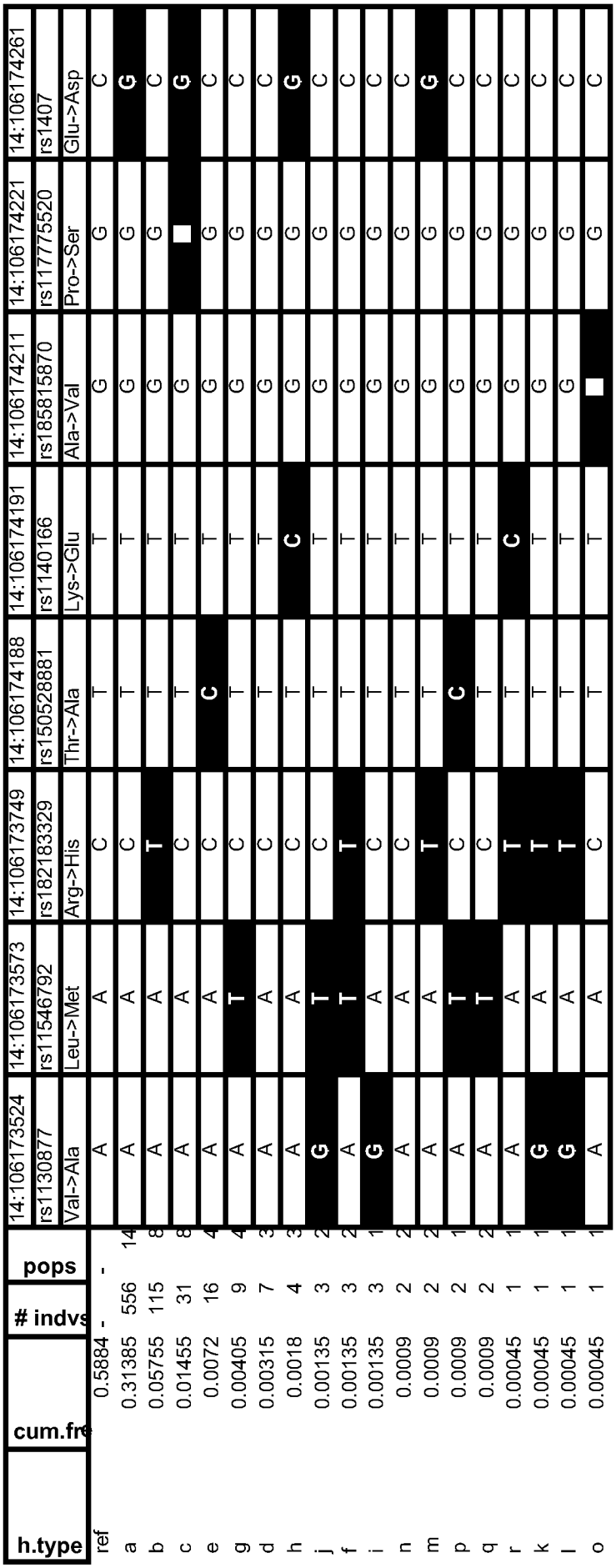


Table 13.1 - IGHA1



h.type	cum.fre	# indvs	pops	14:106173524	14:106173573	14:106173749	14:106174188	14:106174191	14:106174211	14:106174221	14:106174261
				rs1130877	rs11546792	rs182183329	rs150528881	rs1140166	rs185815870	rs117775520	rs1407
				Val->Ala	Leu->Met	Arg->His	Thr->Ala	Lys->Glu	Ala->Val	Pro->Ser	Glu->Asp
ref	0.5884	-	-	A	A	C	T	T	G	G	C
a	0.31385	556	14	A	A	C	T	T	G	G	G
b	0.05755	115	8	A	A	T	T	T	G	G	C
c	0.01455	31	8	A	A	C	T	T	G	■	G
e	0.0072	16	4	A	A	C	C	T	G	G	C
g	0.00405	9	4	A	T	C	T	T	G	G	C
d	0.00315	7	3	A	A	C	T	T	G	G	C
h	0.0018	4	3	A	A	C	T	C	G	G	G
j	0.00135	3	2	G	T	C	T	T	G	G	C
f	0.00135	3	2	A	T	T	T	T	G	G	C
i	0.00135	3	1	G	A	C	T	T	G	G	C
n	0.0009	2	2	A	A	C	T	T	G	G	C
m	0.0009	2	2	A	A	T	T	T	G	G	G
p	0.0009	2	1	A	T	C	C	T	G	G	C
q	0.0009	2	2	A	T	C	T	T	G	G	C
r	0.00045	1	1	A	A	T	T	C	G	G	C
k	0.00045	1	1	G	A	T	T	T	G	G	C
l	0.00045	1	1	G	A	T	T	T	G	G	C
o	0.00045	1	1	A	A	C	T	T	■	G	C

Table 13.1 - IGHA1

h.type	cum.fr	# indvs	pops	14:106174368 rs188231401 Ser->Thr	14:106174744 rs193173336 Val->Met	14:106174984 rs112706225 Lys->Glu	ENST00000390547
ref	0.5884	-	-	A	C	T	
a	0.31385	556	14	A	C	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.05755	115	8	A	C	T	1,5,7,10,11,12,13,14
c	0.01455	31	8	A	C	T	1,2,5,6,7,8,11,13
e	0.0072	16	4	A	C	T	1,10,12,14
g	0.00405	9	4	A	C	T	4,5,10,14
d	0.00315	7	3	A	C	**C**	1,10,14
h	0.0018	4	3	A	C	T	2,6,7
j	0.00135	3	2	A	C	**C**	10,14
f	0.00135	3	2	A	C	T	1,10
i	0.00135	3	1	A	C	**C**	14
n	0.0009	2	2	A	**T**	T	9,10
m	0.0009	2	2	A	C	T	6,9
p	0.0009	2	1	A	C	T	10
q	0.0009	2	2	A	C	**C**	10,12
r	0.00045	1	1	A	C	T	1
k	0.00045	1	1	A	C	T	10
l	0.00045	1	1	**T**	C	T	10
o	0.00045	1	1	A	C	T	4

Table 13.2 - IGHA2

h.type	cum.freq	# indvs	pops	14:106053293	14:106053409	14:106053441	14:106053461	14:106053491	14:106053930	14:106053975	14:106053978
				rs111788346	rs11628783	rs192043389	rs11970	rs138896525	rs139611644	rs72713340	rs142827225
				Ala->Val	Glu->Asp	Arg->Cys	Tyr->Phe	Arg->Gln	Glu->Lys	Pro->Ser	Gln->Glu
ref	0.2023	-	-	G	C	G	T	C	C	G	G
a	0.35655	528	14	■	G	G	■	C	C	G	G
b	0.1647	331	14	G	C	G	T	C	C	G	G
c	0.07815	133	3	G	C	G	T	T	C	■	G
g	0.03195	65	7	G	G	G	T	C	C	G	G
f	0.02645	58	13	G	C	G	■	C	C	G	G
i	0.0228	45	7	G	C	G	■	C	C	G	G
d	0.0227	48	11	G	G	G	■	C	C	G	G
j	0.01865	41	14	G	G	G	T	C	C	G	G
e	0.01465	31	10	■	C	G	■	C	C	G	G
k	0.01045	23	11	■	C	G	T	C	C	G	G
h	0.00955	20	9	■	G	G	T	C	C	G	G
m	0.00955	21	11	G	C	G	T	C	C	G	G
l	0.0041	9	3	G	C	G	T	T	C	G	G
p	0.0036	8	5	G	G	G	T	C	C	G	G
o	0.00315	7	3	■	G	G	■	C	C	■	G
n	0.0027	6	4	G	C	G	■	C	C	■	G
q	0.0018	4	2	G	C	G	T	T	C	G	G
r	0.0018	4	4	■	C	G	T	C	C	G	G
ai	0.0018	4	1	G	G	G	T	T	T	G	G
ag	0.00135	3	3	G	C	G	T	C	C	■	G
s	0.00135	3	1	G	C	G	■	C	C	G	G
y	0.0009	2	2	■	G	G	T	C	C	G	G
v	0.0009	2	2	■	G	G	■	C	C	G	G
ab	0.0009	2	2	G	C	G	T	C	C	G	G
ak	0.0009	2	2	G	C	G	T	C	C	G	G
ad	0.0009	2	2	G	C	G	T	C	T	G	G
ac	0.0009	2	2	G	G	G	■	C	C	G	G
ah	0.00045	1	1	G	C	G	T	C	C	■	G
u	0.00045	1	1	G	C	G	■	C	C	G	G

Table 13.2 - IGHA2

aj	0.00045	1	1	G	G	G	T	T	C	□	G
aa	0.00045	1	1	G	C	G	□	C	T	G	G
t	0.00045	1	1	□	C	G	□	C	C	□	G
z	0.00045	1	1	□	C	G	T	T	C	□	G
w	0.00045	1	1	G	G	G	□	C	C	□	G
af	0.00045	1	1	G	C	□	T	C	C	G	G
x	0.00045	1	1	□	G	G	□	C	C	G	C
ae	0.00045	1	1	□	G	G	T	C	C	□	G

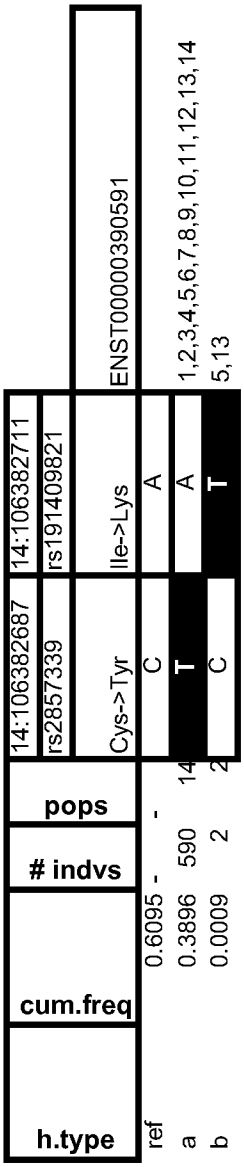
Table 13.2 - IGHA2

h.type	cum.freq	# indvs	pops	14:106054166 rs182368844 Arg->Gln	14:106054428 rs9972103 Arg->ProArg->Pro	14:106054456 rs61984162 Ser->Pro	ENST00000390539
ref	0.2023	-	-	C	C	A	
a	0.35655	528	14	C	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.1647	331	14	C	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.07815	133	3	C	C	A	3,4,9
g	0.03195	65	7	C	C	A	1,3,5,10,11,12,14
f	0.02645	58	13	C	G	G	1,2,3,4,5,6,7,9,10,11,12,13,14
i	0.0228	45	7	C	C	A	1,4,5,9,10,12,14
d	0.0227	48	11	C	G	G	2,3,4,5,6,7,9,11,12,13,14
j	0.01865	41	14	C	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
e	0.01465	31	10	C	G	G	2,3,5,6,7,8,9,11,12,13
k	0.01045	23	11	C	G	G	2,3,5,6,7,8,9,10,11,12,13
h	0.00955	20	9	C	G	G	1,3,5,6,7,9,11,13,14
m	0.00955	21	11	C	C	G	1,2,3,4,7,9,10,11,12,13,14
l	0.0041	9	3	C	C	A	3,4,9
p	0.0036	8	5	C	C	G	6,7,10,11,14
o	0.00315	7	3	C	G	G	2,7,13
n	0.0027	6	4	C	G	G	2,5,6,13
q	0.0018	4	2	C	G	G	4,9
r	0.0018	4	4	C	C	A	2,5,11,13
ai	0.0018	4	1	C	C	A	14
ag	0.00135	3	3	C	G	G	1,5,7
s	0.00135	3	1	C	C	G	1
y	0.0009	2	2	C	C	A	10,14
v	0.0009	2	2	C	C	G	1,9
ab	0.0009	2	2	C	G	A	4,10
ak	0.0009	2	2	T	C	A	10,12
ad	0.0009	2	2	C	C	A	7,10
ac	0.0009	2	2	C	C	A	1,9
ah	0.00045	1	1	C	C	A	5
u	0.00045	1	1	C	G	A	5

Table 13.2 - IGHA2

aj	0.00045	1	1	C	C	A	3
aa	0.00045	1	1	C	C	A	5
t	0.00045	1	1	C	G	G	11
z	0.00045	1	1	T	G	G	3
w	0.00045	1	1	C	G	G	13
af	0.00045	1	1	C	C	A	14
x	0.00045	1	1	C	G	G	13
ae	0.00045	1	1	C	G	G	2

Table 13.3 - IGHD2-2



h.type	cum.freq	# indvs	pops	14:106382687 rs2857339 Cys->Tyr	14:106382711 rs191409821 Ile->Lys	ENST00000390591
ref	0.6095	-	-	C	A	
a	0.3896	590	14	T	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0009	2	2	C	T	5,13

Table 13.4 - IGHD2-8

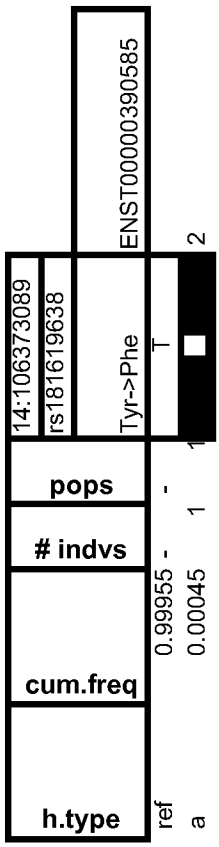


Table 13.5 - IGHD3-10

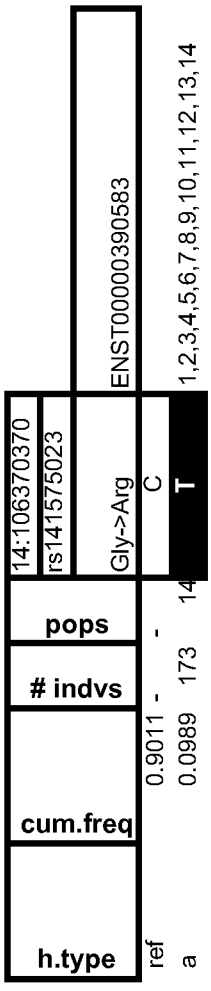


Table 13.6 - IGHD3-16

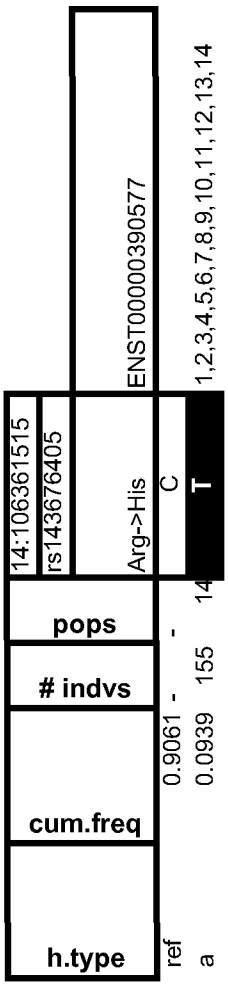


Table 13.7 - IGHD3-3

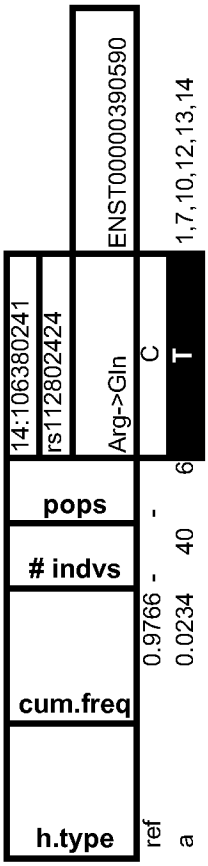


Table 13.8 - IGHD4-23

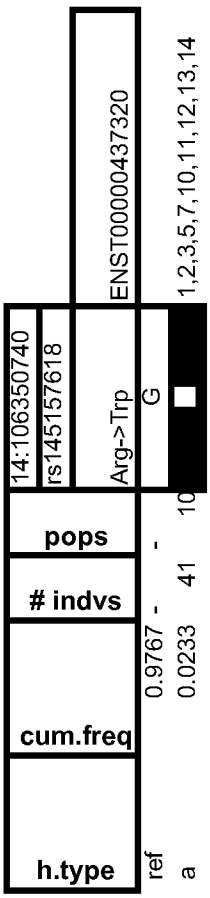


Table 13.9 - IGHD4-4

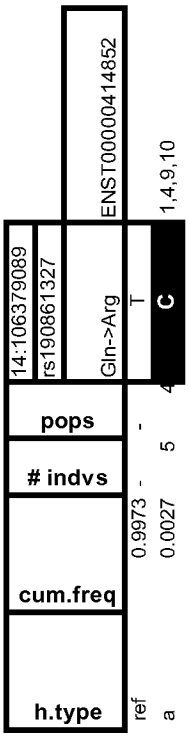


Table 13.10 - IGHD6-13

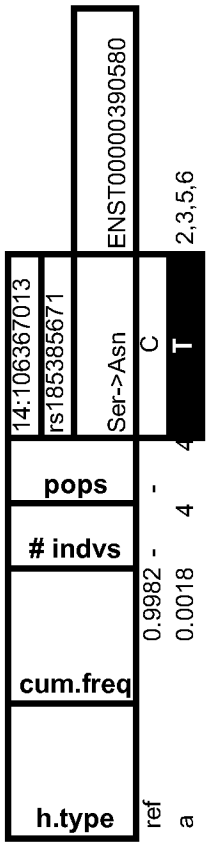


Table 13.11 - IGHD7-27

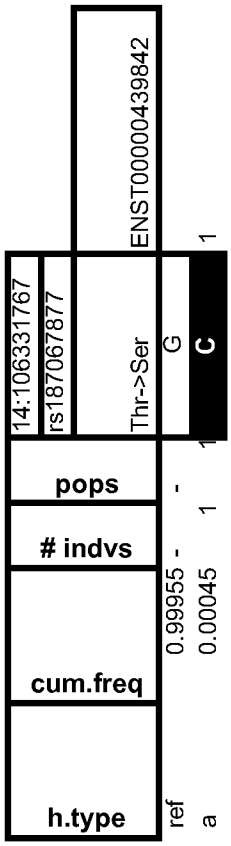
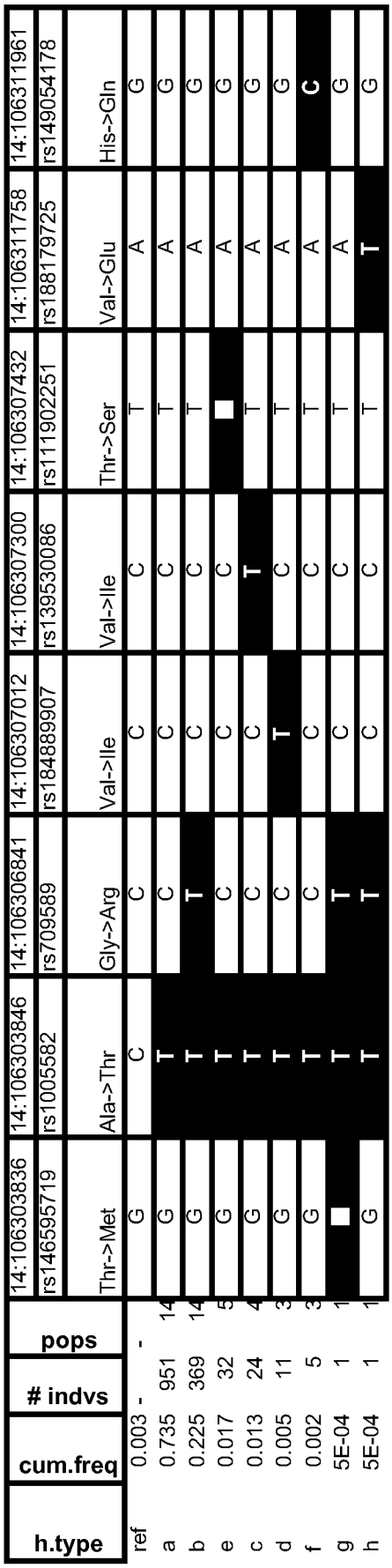


Table 13.12 - IGHD



h.type	cum.freq	# indvs	pops	14:106303836	14:106303846	14:106306841	14:106307012	14:106307300	14:106307432	14:106311758	14:106311961
				rs146595719	rs1005582	rs709589	rs184889907	rs139530086	rs111902251	rs188179725	rs149054178
				Thr->Met	Ala->Thr	Gly->Arg	Val->Ile	Val->Ile	Thr->Ser	Val->Glu	His->Gln
ref	0.003	-	-	G	C	C	C	C	T	A	G
a	0.735	951	14	G	T	C	C	C	T	A	G
b	0.225	369	14	G	T	T	C	C	T	A	G
e	0.017	32	5	G	T	C	C	C	■	A	G
c	0.013	24	4	G	T	C	C	T	T	A	G
d	0.005	11	3	G	T	C	T	C	T	A	G
f	0.002	5	3	G	T	C	C	C	T	A	C
g	5E-04	1	1	■	T	T	C	C	T	A	G
h	5E-04	1	1	G	T	T	C	C	T	T	G

Table 13.12 - IGHD

h.type	cum.freq	# indvs	pops	ENST00000390556
ref	0.003	-	-	
a	0.735	951	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.225	369	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
e	0.017	32	5	1,5,10,12,14
c	0.013	24	4	1,5,11,12
d	0.005	11	3	3,4,9
f	0.002	5	3	2,6,7
g	5E-04	1	1	13
h	5E-04	1	1	6

Table 13.13 - IGHEP1

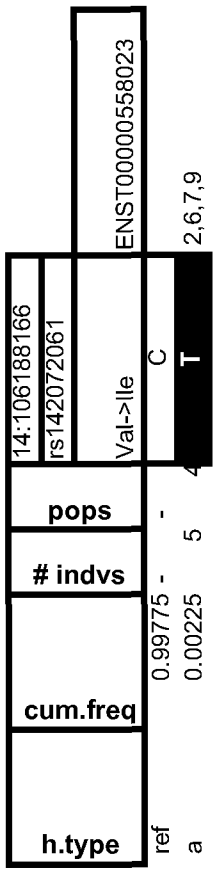


Table 13.14 - IGHE

h.type	cum.freq	# indvs	pops	14:106066524 rs189365488 Arg->His	14:106066668 rs148738013 Arg->His	14:106066690 rs61675149 Trp->Gly	14:106066701 rs57724657 Ala->Val	14:106067096 rs182957708 Pro->Leu	14:106067118 rs150823665 Ala->Thr	14:106067815 rs144154566 Thr->Ile	ENST00000390541
ref	0.87655	-	-	C	C	A	G	G	C	G	
a	0.08375	144	7	C	C	**C**	G	G	C	G	3,4,5,9,10,11,14
b	0.03205	66	5	C	C	**C**	■	G	C	G	1,5,10,12,14
d	0.00225	5	2	C	C	**C**	■	G	**T**	G	10,14
c	0.00135	3	2	**T**	C	A	G	G	C	G	1,10
h	0.00135	3	1	C	C	A	G	G	**T**	G	14
g	0.0009	2	2	C	**T**	A	G	G	C	G	7,13
e	0.0009	2	2	C	C	**C**	G	■	C	G	3,4
f	0.0009	2	2	C	C	A	G	G	C	■	6,13

Table 13.15 - IGHG1

h.type	cum.freq	# indvs	pops	14:106204113	14:106204131	14:106208326	14:106208411		
				rs117518546	rs138262373	rs1043109	rs184594692		
				Gly->Arg	Asp->Asn	Leu->Val	Asp->Glu		
						Leu->Val	Asp->Glu		
ref	0.89185	-	-	C	C	G	G	ENST00000390542	ENST00000390548	ENST00000390549	
a	0.09825	158	4	**T**	C	G	G	3,4,9,13			
b	0.0063	14	6	C	**T**	G	G	2,5,6,7,11,13			
d	0.0018	4	2	C	C	**C**	G	3,9			
c	0.0018	4	1	C	C	G	**T**	14			

Table 13.16 - IGHG2

h.type	cum.freq	# indvs	pops	14:106109830 rs185308159 Pro->Thr	14:106110017 rs182705016 Lys->Glu	14:106110056 rs113678609 Val->Leu	14:106110137 rs8009156 Val->Met	14:106110209 rs190914111 Glu->Lys	14:106110914 rs11627594 Pro->Thr	14:106110926 rs139934867 Val->Met	ENST00000390545
ref	0.6966	-	-	G	T	C	C	C	G	C	
a	0.2868	499	12	G	T	C	**T**	C	**T**	C	1,2,3,4,5,6,7,8,9,11,12,13
b	0.0045	9	1	**T**	T	C	C	C	G	**T**	10
f	0.0036	8	3	G	T	C	C	C	**T**	C	7,11,13
h	0.0027	6	6	G	T	**G**	C	C	G	C	1,6,7,9,10,13
d	0.0023	5	3	G	T	C	C	**T**	G	C	3,4,9
e	0.0018	4	3	G	T	C	**T**	C	G	C	2,4,5
c	0.0009	2	2	**T**	T	C	C	C	G	C	1,10
g	0.0005	1	1	G	**C**	C	C	C	G	C	10
i	0.0005	1	1	G	T	C	C	**T**	**T**	C	4

Table 13.17 - IGHG3

h.type	cum.fre	# indivs	pops	14:106235611	14:106235663	14:106235692	14:106235729	14:106235742	14:106235783	14:106236000	14:106236128
				rs1051112	rs190174000	rs187517394	rs139413052	rs149653267	rs113169458	rs141959627	rs12890621
				Phe->Tyr	Gln->Glu	Lys->Arg	Met->Val	Asn->Lys	Val->Met	Thr->Ala	Tyr->Phe
ref	0.2495	-	-	A	G	T	T	G	C	T	T
a	0.2813	479	14	**T**	G	T	T	G	C	T	T
b	0.19955	345	11	A	G	T	T	G	C	T	**N**
j	0.02345	50	8	A	G	T	T	**C**	C	T	T
d	0.02095	44	6	**T**	G	T	**C**	**C**	**T**	T	T
c	0.0209	46	10	**T**	G	T	T	G	C	T	**N**
i	0.019	41	10	**T**	G	T	T	G	C	T	T
g	0.01735	38	10	**T**	G	T	**C**	**C**	C	T	T
f	0.01725	36	7	**T**	G	T	T	G	C	**C**	T
h	0.0164	36	11	**T**	G	T	T	**C**	C	T	T
e	0.015	33	9	A	G	T	T	G	C	T	T
k	0.01315	27	5	A	**C**	**C**	**C**	G	C	T	T
t	0.0126	28	6	**T**	G	T	T	**C**	C	T	T
q	0.0126	27	3	A	G	T	**C**	G	C	T	T
v	0.0072	16	8	**T**	G	T	**C**	G	C	T	T
o	0.00675	15	7	A	G	T	T	**C**	C	T	**N**
r	0.0063	14	5	**T**	G	T	**C**	**C**	C	T	T
m	0.00585	13	5	A	**C**	T	T	**C**	C	T	T
n	0.00495	11	4	**T**	G	T	T	**C**	**T**	T	T
am	0.0045	10	5	**T**	G	T	**C**	**C**	**T**	T	T
l	0.0036	8	6	**T**	G	T	T	**C**	C	T	**N**
p	0.0036	8	2	A	**C**	T	T	G	C	T	T
u	0.0036	8	4	**T**	G	T	**C**	**C**	C	T	**N**
x	0.00315	7	3	**T**	**C**	T	**C**	**C**	C	T	**N**
ae	0.00225	5	3	**T**	**C**	**C**	**C**	G	C	T	T
ai	0.00225	5	2	**T**	G	T	T	**C**	**T**	T	T
aj	0.0018	4	2	A	G	T	**C**	**C**	C	T	T
w	0.0018	4	1	**T**	G	T	T	G	C	T	T
bd	0.00135	3	1	**T**	G	T	**C**	**C**	C	T	T
al	0.00135	3	2	A	G	T	**C**	**C**	**T**	T	T

Table 13.17 - IGHG3

s	0.00135	3	2	T	C	T	T	G	C	T	■
ab	0.00135	3	2	A	G	T	T	G	C	T	T
ak	0.00135	3	2	A	G	T	T	C	T	T	T
ag	0.0009	2	2	T	G	T	C	G	C	T	T
y	0.0009	2	2	A	G	T	T	G	C	T	■
az	0.0009	2	2	T	G	T	T	G	T	T	■
aa	0.0009	2	2	A	G	T	C	C	C	T	■
ap	0.0009	2	2	A	C	T	C	G	C	T	T
av	0.0009	2	1	T	G	T	T	C	C	T	T
af	0.0009	2	1	T	C	C	C	C	C	T	T
at	0.0009	2	2	T	G	C	C	G	C	T	T
aw	0.0009	2	2	T	G	C	T	C	C	C	T
ay	0.0009	2	1	A	G	T	C	G	C	T	■
ar	0.0009	2	2	A	G	C	C	G	C	T	T
an	0.0009	2	1	T	G	T	T	C	C	T	T
ah	0.00045	1	1	A	G	T	T	G	C	T	T
ba	0.00045	1	1	A	G	T	T	G	C	C	■
ax	0.00045	1	1	A	C	T	C	C	C	T	T
bb	0.00045	1	1	T	G	C	C	C	C	T	T
aq	0.00045	1	1	A	G	T	T	G	C	C	T
be	0.00045	1	1	T	G	C	T	G	C	T	T
z	0.00045	1	1	A	G	T	T	G	C	T	■
ao	0.00045	1	1	T	C	C	T	C	C	T	T
au	0.00045	1	1	T	G	T	T	G	T	T	T
ad	0.00045	1	1	T	C	T	T	C	C	T	T
bc	0.00045	1	1	A	G	T	T	C	T	T	T
ac	0.00045	1	1	T	C	T	T	C	C	T	■
as	0.00045	1	1	A	G	T	T	G	T	T	T

Table 13.17 - IGHG3

h.type	cum.fre	# indvs	pops	14:106236141	14:106236143	14:106236174	14:106236187	14:106236195	14:106236315	14:106237516
				rs60746425	rs74093865	rs184265224	rs141238286	rs145035200	rs189074626	rs186135943
				Arg->Trp	Pro->Leu	Gly->Ser	Lys->Asn	Gln->Lys	Leu->Phe	Leu->Phe
ref	0.2495	-	-	G	G	C	C	G	G	C
a	0.2813	479	14	G	[redacted]	C	C	G	G	C
b	0.19955	345	11	G	G	C	C	G	G	C
j	0.02345	50	8	G	G	C	C	G	G	C
d	0.02095	44	6	[redacted]	G	C	C	G	G	C
c	0.0209	46	10	G	G	C	C	G	G	C
i	0.019	41	10	G	G	C	C	G	G	C
g	0.01735	38	10	G	[redacted]	C	C	G	G	C
f	0.01725	36	7	G	[redacted]	C	C	G	G	C
h	0.0164	36	11	G	[redacted]	C	C	G	G	C
e	0.015	33	9	G	[redacted]	C	C	G	G	C
k	0.01315	27	5	G	G	C	C	G	G	C
t	0.0126	28	6	G	G	C	C	G	G	C
q	0.0126	27	3	G	G	C	C	G	G	C
v	0.0072	16	8	G	[redacted]	C	C	G	G	C
o	0.00675	15	7	G	G	C	C	G	G	C
r	0.0063	14	5	G	G	C	C	G	G	C
m	0.00585	13	5	G	G	C	C	G	G	C
n	0.00495	11	4	[redacted]	G	C	C	G	G	C
am	0.0045	10	5	G	G	C	C	G	G	G
l	0.0036	8	6	G	G	C	C	G	G	C
p	0.0036	8	2	G	G	C	C	G	G	C
u	0.0036	8	4	G	G	C	C	G	G	C
x	0.00315	7	3	G	G	C	C	G	G	C
ae	0.00225	5	3	G	G	C	C	G	G	C
ai	0.00225	5	2	G	G	C	C	G	G	G
aj	0.0018	4	2	G	G	C	C	G	G	C
w	0.0018	4	1	G	[redacted]	C	C	T	G	C
bd	0.00135	3	1	G	[redacted]	C	C	T	G	C
al	0.00135	3	2	G	G	C	C	G	G	G

Table 13.17 - IGHG3

s	0.00135	3	2	G	G	C	C	G	G	C
ab	0.00135	3	2	G	G	C	G	G	G	C
ak	0.00135	3	2	G	G	C	C	G	G	C
ag	0.0009	2	2	G	G	C	C	G	G	C
y	0.0009	2	2	G	G	T	C	G	G	C
az	0.0009	2	2	G	G	C	C	G	G	C
aa	0.0009	2	2	G	G	C	C	G	G	C
ap	0.0009	2	2	G	G	C	C	G	G	C
av	0.0009	2	1	G	■	C	C	T	G	C
af	0.0009	2	1	G	G	C	C	G	G	C
at	0.0009	2	2	G	G	C	C	G	G	C
aw	0.0009	2	2	G	■	C	C	G	G	C
ay	0.0009	2	1	G	G	C	C	G	G	C
ar	0.0009	2	2	G	G	C	C	G	G	C
an	0.0009	2	1	■	G	C	C	G	G	C
ah	0.00045	1	1	G	G	C	C	G	■	C
ba	0.00045	1	1	G	G	C	C	G	G	C
ax	0.00045	1	1	G	G	C	C	G	G	C
bb	0.00045	1	1	G	G	C	C	G	G	C
aq	0.00045	1	1	G	G	C	C	G	G	C
be	0.00045	1	1	■	G	C	C	G	G	C
z	0.00045	1	1	G	G	C	C	G	G	G
ao	0.00045	1	1	G	G	C	C	G	G	C
au	0.00045	1	1	■	G	C	C	G	G	C
ad	0.00045	1	1	G	G	C	C	G	G	C
bc	0.00045	1	1	G	G	C	C	G	G	G
ac	0.00045	1	1	G	G	C	C	G	G	C
as	0.00045	1	1	■	G	C	C	G	G	C

Table 13.17 - IGHG3

h.type	cum.freq	# indvs	pops	ENST00000390551
ref	0.2495	-	-	
a	0.2813	479	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.19955	345	11	1,2,3,4,5,6,7,8,11,12,13
j	0.02345	50	8	1,2,3,4,5,10,12,14
d	0.02095	44	6	1,3,4,8,9,11
c	0.0209	46	10	1,2,3,5,6,7,8,11,12,13
i	0.019	41	10	1,2,3,4,7,8,10,12,13,14
g	0.01735	38	10	2,3,4,5,6,7,9,11,12,13
f	0.01725	36	7	1,2,5,6,7,11,13
h	0.0164	36	11	2,3,4,5,6,7,8,9,11,12,13
e	0.015	33	9	2,4,5,6,7,9,11,12,13
k	0.01315	27	5	1,5,10,12,14
t	0.0126	28	6	1,2,5,10,13,14
q	0.0126	27	3	1,10,14
v	0.0072	16	8	2,3,5,6,7,9,11,12
o	0.00675	15	7	2,5,7,8,11,12,13
r	0.0063	14	5	1,4,10,12,14
m	0.00585	13	5	1,5,10,12,14
n	0.00495	11	4	3,4,9,11
am	0.0045	10	5	1,5,10,12,14
l	0.0036	8	6	2,6,7,8,12,13
p	0.0036	8	2	1,10
u	0.0036	8	4	1,2,6,11
x	0.00315	7	3	2,7,13
ae	0.00225	5	3	1,10,14
ai	0.00225	5	2	1,14
aj	0.0018	4	2	10,14
w	0.0018	4	1	4
bd	0.00135	3	1	9
al	0.00135	3	2	1,14

Table 13.17 - IGHG3

s	0.00135	3	2 7,11
ab	0.00135	3	2 1,14
ak	0.00135	3	2 10,14
ag	0.0009	2	2 1,10
y	0.0009	2	2 2,7
az	0.0009	2	2 2,13
aa	0.0009	2	2 2,6
ap	0.0009	2	2 10,14
av	0.0009	2	1 9
af	0.0009	2	1 14
at	0.0009	2	2 10,14
aw	0.0009	2	2 5,9
ay	0.0009	2	1 11
ar	0.0009	2	2 5,14
an	0.0009	2	1 9
ah	0.00045	1	1 10
ba	0.00045	1	1 13
ax	0.00045	1	1 10
bb	0.00045	1	1 1
aq	0.00045	1	1 5
be	0.00045	1	1 7
z	0.00045	1	1 7
ao	0.00045	1	1 14
au	0.00045	1	1 5
ad	0.00045	1	1 1
bc	0.00045	1	1 1
ac	0.00045	1	1 13
as	0.00045	1	1 4

Table 13.18 - IGHG4

h.type	cum.freq	# indvs	pops	14:106090959	14:106091083	14:106091281	14:106091329	14:106091332	14:106091341	14:106091347	14:106091450
				rs190355823	rs146589796	rs112773794	rs8015545	rs138186657	rs143666134	rs145943426	rs138777601
				Ser->Ala	Met->Ile	Asn->His	Leu->Val	Val->Ile	Val->Ile	Val->Ile	Gln->His
ref	0.77355	-	-	A	C	T	G	C	C	C	C
a	0.19825	389	12	A	C	T	**C**	C	C	C	C
b	0.01005	22	4	A	C	**G**	G	**T**	C	C	C
c	0.0055	12	4	A	**T**	T	G	C	C	C	C
d	0.0041	8	4	A	C	**G**	G	**T**	C	C	C
f	0.0027	6	3	A	C	T	G	C	**T**	C	C
e	0.00225	4	2	A	C	T	G	C	C	C	**G**
g	0.0018	4	2	A	C	T	G	C	C	C	C
j	0.00045	1	1	A	C	**G**	G	C	C	C	C
k	0.00045	1	1	**C**	C	T	G	C	C	C	C
h	0.00045	1	1	A	C	T	G	C	C	**T**	C
i	0.00045	1	1	A	C	T	**C**	C	C	C	**G**

Table 13.18 - IGHG4

h.type	cum.freq	# indvs	pops	14:106091714 rs139754373 Ser->ProSer->Pro	14:106092363 rs183713766 Cys->Tyr	ENST00000390543
ref	0.77355	-	-	A	C	
a	0.19825	389	12	A	C	1,2,3,4,5,6,7,8,9,11,12,13
b	0.01005	22	4	**G**	C	1,10,12,14
c	0.0055	12	4	A	C	1,5,10,14
d	0.0041	8	4	A	C	1,5,10,14
f	0.0027	6	3	A	C	1,10,14
e	0.00225	4	2	A	C	10,11
g	0.0018	4	2	A	**T**	1,14
j	0.00045	1	1	A	C	1
k	0.00045	1	1	A	C	1
h	0.00045	1	1	A	C	10
i	0.00045	1	1	A	C	4

Table 13.19 - IGHJ1

h.type	cum.fre	# indvs	pops	14:106331628	rs181379404	ENST00000390565
						Thr->Ser
ref	0.9982	-	-			G
a	0.0018	3	3			C
						2,3,13

Table 13.20 - IGHJ2

h.type	cum.fre	# indvs	pops	14:106331453 rs181475274 Tyr->Phe	14:106331455 rs185806009 Trp->Cys	ENST00000390564
ref	0.9936	-	-	T	C	
a	0.00415	9	6	T	G	1,3,4,7,10,13
b	0.00225	4	3	■	C	1,10,14

Table 13.21 - IGHJ3

h.type	cum.freq	# indvs	pops	14:106330799 rs190137342 Ser->Leu	14:106330814 rs181332275 Met->Thr	14:106330815 rs185836312 Met->Val	14:106330833 rs190698203 Ile->Phe	14:106330841 rs183030857 Ala->Val	ENST00000463911
ref	0.95755	-	-	G	A	T	T	G	
a	0.01645	34	12	G	A	T	T		1,2,3,4,5,6,7,9,10,12,13,14
b	0.0114	23	10	G	A	**C**	T	G	2,3,4,6,7,9,10,11,13,14
c	0.0064	14	11	G	A	T		G	1,2,5,6,7,8,9,10,11,13,14
d	0.0046	9	7	G	**G**	T	T	G	5,7,9,10,11,13,14
e	0.00315	6	5		A	T	T	G	1,2,3,5,10
f	0.00045	1	1	G	A	**C**	T		5

Table 13.22 - IGHJ4

h.type	cum.freq	# indvs	pops	14:106330436 rs145372686 Thr->Ser	14:106330442 rs183645858 Leu->Arg	14:106330460 rs150237274 Tyr->Phe	14:106330467 rs189135736 Phe->Ile	ENST00000461719
ref	0.9758	-	-	G	A	T	A	
a	0.01785	39	11	G	A	■	A	2,3,4,6,7,9,10,11,12,13,14
b	0.0032	7	5	G	A	T	**T**	1,2,9,11,13
c	0.0018	4	4	G	**C**	T	A	3,4,13,14
d	0.00135	3	3	**C**	A	T	A	2,4,10

Table 13.23 - IGHJ5

h.type	cum.freq	# indvs	pops	14:106330027 rs147551384 Ser->Ala	14:106330035 rs141950820 Thr->Ile	14:106330067 rs78872461 Trp->Cys	ENST00000488476
ref	0.9497	-	-	A	G	C	
a	0.0357	78	13	A	G	**G**	1,2,3,4,5,6,7,9,10,11,12,13,14
b	0.011	24	10	A	■	C	1,2,3,5,6,7,8,9,10,11
c	0.0027	5	4	**C**	G	C	2,3,7,13
d	0.0009	2	2	A	■	**G**	4,14

Table 13.24 - IGHJ6

h.type	cum.freq	# indvs	pops	14:106329414 rs111729691 Ser->Pro	14:106329417 rs142191790 Val->Ile	14:106329434 rs140411866 Lys->Arg	14:106329435 rs1950943 Lys->Gln	ENST00000390560
ref	0.3083	-	-	A	C	T	T	
a	0.6845	945	14	A	C	T	**G**	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.00315	7	2	A	**T**	T	T	10,14
e	0.0009	2	2	A	C	**C**	**G**	4,13
h	0.0009	2	2	**G**	C	T	**G**	2,6
f	0.0009	2	2	A	C	**C**	T	3,10
d	0.00045	1	1	A	**T**	T	**G**	9
g	0.00045	1	1	**G**	C	**C**	**G**	4
c	0.00045	1	1	**G**	C	T	T	10

Table 13.25 - IGHM

h.type	cum.freq	# indvs	pops	14:106320732	14:106321313	14:106321330	14:106321590	14:106321663	14:106322093	
				rs146456238	rs186069870	rs45471499	rs12365	rs1059216	rs113762053	
				Asp->Asn	Ala->Thr	Thr->Ile	Gly->Val	Ser->Gly	Asp->Glu	ENST00000390559
ref	0.0249	-	-	C	C	G	C	T	G	
a	0.7527	974	14	C	C	G	■	T	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.15245	287	14	C	C	G	■	**C**	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.03795	74	9	C	C	G	C	**C**	**C**	1,2,5,7,10,11,12,13,14
c	0.02705	54	3	C	C	■	■	**C**	G	3,4,9
e	0.0018	4	2	C	C	G	C	**C**	G	10,14
g	0.00135	3	3	C	**T**	G	C	T	G	5,12,14
f	0.0009	2	1	C	C	G	C	T	**C**	11
h	0.00045	1	1	C	**T**	G	■	T	G	5
i	0.00045	1	1	**T**	C	G	C	**C**	**C**	10

Table 13.26 - IGHV1-18

h.type	cum.freq	# indvs	pops	14:106641643 rs181586608 Thr->Ala	14:106641660 rs146311654 Gly->Asp	14:106641668 rs184544516 Lys->Asn	14:106641688 rs114910155 Asn->Asp	14:106641769 rs139305030 Thr->Pro	14:106641801 rs188509150 Lys->Arg	14:106641832 rs181324183 Ala->Pro	14:106641981 rs184535212 Ser->Asn
ref	0.97595	-	-	T	C	C	T	T	T	C	C
a	0.01185	23	6	T	C	C	**C**	T	T	C	C
c	0.0045	7	4	T	**T**	C	T	T	T	C	C
b	0.00275	6	5	T	C	**G**	T	T	T	C	C
d	0.00135	3	3	T	C	C	T	T	**C**	C	C
e	0.00135	3	3	T	C	C	T	**G**	T	C	C
f	0.0009	2	2	T	C	C	T	T	T	C	**T**
g	0.00045	1	1	**C**	**T**	C	T	T	T	C	C
h	0.00045	1	1	T	C	C	T	T	T	**G**	C
i	0.00045	1	1	**C**	C	C	T	T	T	C	C

Table 13.26 - IGHV1-18

h.type	cum.freq	# indvs	pops	ENST00000390605
ref	0.97595	-	-	
a	0.01185	23	6	1,6,9,10,11,14
c	0.0045	7	4	1,10,11,14
b	0.00275	6	5	6,10,11,13,14
d	0.00135	3	3	4,11,14
e	0.00135	3	3	2,6,11
f	0.0009	2	2	8,9
g	0.00045	1	1	13
h	0.00045	1	1	10
i	0.00045	1	1	10

Table 13.27 - IGHV1-24

h.type	cum.freq	# indvs	pops	14:106733154 rs141629050 Tyr->Phe	14:106733158 rs150510876 Tyr->His	14:106733166 rs139845206 Thr->Met	14:106733192 rs141655646 Glu->Asp	14:106733263 rs183704154 Ile->Phe	14:106733269 rs114334184 Glu->Lys	14:106733274 rs187836940 Asp->Gly	14:106733336 rs139283249 Met->Ile
ref	0.97175	-	-	T	A	G	C	T	C	T	C
a	0.01185	24	4	T	A	G	C	T	C	T	C
b	0.0082	18	4	T	A	G	C	T	C	T	C
c	0.00415	9	3	T	A	G	**G**	T	C	T	C
g	0.0009	2	1	T	A	G	C	T	C	T	**T**
d	0.00045	1	1	T	**G**	■	C	T	C	T	C
j	0.00045	1	1	T	A	G	C	T	C	T	C
k	0.00045	1	1	T	A	G	C	T	C	**C**	C
e	0.00045	1	1	T	A	G	C	■	C	T	C
h	0.00045	1	1	T	A	G	C	T	**T**	T	C
f	0.00045	1	1	T	A	G	C	T	C	T	C
i	0.00045	1	1	■	A	G	C	T	C	T	C

Table 13.27 - IGHV1-24

h.type	cum.freq	# indvs	pops	14:106733337 rs80012805 Met->Arg	14:106733341 rs186604864 Ser->Pro	14:106733359 rs184694183 Tyr->Asn	14:106733446 rs147700118 Thr->Ala	14:106733570 rs192824042 Thr->Ala	ENST00000390610
ref	0.97175	-	-	A	A	A	T	T	
a	0.01185	24	4	**C**	A	A	T	T	1,10,11,14
b	0.0082	18	4	A	A	A	**C**	T	1,10,12,14
c	0.00415	9	3	A	A	A	T	T	1,10,14
g	0.0009	2	1	A	A	A	T	T	2
d	0.00045	1	1	A	A	A	T	T	14
j	0.00045	1	1	A	A	A	T	**C**	5
k	0.00045	1	1	A	**G**	**T**	**C**	T	5
e	0.00045	1	1	A	**G**	A	T	T	5
h	0.00045	1	1	A	A	A	T	T	14
f	0.00045	1	1	A	A	**T**	T	T	13
i	0.00045	1	1	A	A	A	T	T	3

Table 13.28 - IGHV1-2

h.type	cum.freq	# indvs	pops	14:106452690	14:106452735	14:106452766	14:106452784	14:106452793	14:106452799	14:106452817	14:106452868
				rs12588974	rs188093056	rs112806369	rs180944534	rs186314384	rs189402450	rs1065059	rs192043188
				Ala->Val	Ser->Asn	Arg->Trp	Ala->Ser	Thr->Ser	Gly->Ser	Trp->Arg	Tyr->Asn
ref	0.43185	-	-	G	C	T	C	T	C	A	A
a	0.3268	511	14	G	C	■	C	T	C	A	A
b	0.1304	228	14	■	C	T	C	T	C	**G**	A
c	0.0797	158	14	G	C	T	C	T	C	**G**	A
d	0.01135	25	10	■	C	T	C	T	C	A	A
e	0.0073	16	6	G	C	T	C	T	C	A	A
g	0.0018	4	3	G	**T**	T	C	T	C	A	A
h	0.00135	3	2	■	C	■	C	T	C	A	A
f	0.00135	3	3	G	C	■	C	T	C	A	A
i	0.0009	2	2	■	C	T	C	T	C	**G**	A
m	0.0009	2	2	G	C	T	C	T	C	**G**	A
l	0.0009	2	2	G	C	T	C	T	C	A	A
j	0.00045	1	1	G	C	T	■	T	C	A	A
u	0.00045	1	1	G	**T**	T	C	T	C	A	A
k	0.00045	1	1	G	C	T	C	T	C	**G**	A
t	0.00045	1	1	G	C	T	C	T	C	A	**T**
v	0.00045	1	1	G	C	T	C	T	C	A	A
s	0.00045	1	1	G	**T**	T	C	■	C	A	A
q	0.00045	1	1	G	C	T	C	T	**T**	A	A
w	0.00045	1	1	G	C	■	C	T	C	A	A
r	0.00045	1	1	G	C	T	C	T	C	A	A
n	0.00045	1	1	G	C	T	C	T	C	A	A
p	0.00045	1	1	G	**T**	T	C	T	C	**G**	A
o	0.00045	1	1	G	C	T	C	T	C	A	A

Table 13.28 - IGHV1-2

h.type	cum.freq	# indvs	pops	14:106452873 rs1065058 Gly->Asp	14:106452882 rs11553010 Thr->Ser	14:106452897 rs185487229 Lys->Met	14:106452903 rs189604828 Ser->Cys	14:106452956 rs182256840 Gln->His	14:106452972 rs186783545 Ala->Gly	ENST00000390594
ref	0.43185	-	-	C	G	T	G	C	G	
a	0.3268	511	14	C	G	T	G	C	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.1304	228	14	C	G	T	G	C	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0797	158	14	C	G	T	G	C	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.01135	25	10	C	G	T	G	C	G	1,2,3,4,5,9,10,11,13,14
e	0.0073	16	6	**T**	G	T	G	C	G	1,2,5,10,12,14
g	0.0018	4	3	C	G	T	G	C	G	6,12,14
h	0.00135	3	2	C	G	T	G	C	G	3,4
f	0.00135	3	3	**T**	G	T	G	C	G	2,6,7
i	0.0009	2	2	**T**	G	T	G	C	G	11,13
m	0.0009	2	2	**T**	G	T	G	C	G	3,6
l	0.0009	2	2	C	**C**	T	G	C	G	9,10
j	0.00045	1	1	C	G	■	G	C	G	2
u	0.00045	1	1	**T**	G	T	G	C	G	2
k	0.00045	1	1	C	**C**	T	G	C	G	2
t	0.00045	1	1	C	G	T	G	C	G	2
v	0.00045	1	1	**T**	G	■	G	C	G	10
s	0.00045	1	1	C	G	T	G	C	G	2
q	0.00045	1	1	C	G	T	G	C	G	10
w	0.00045	1	1	C	G	T	**C**	C	G	11
r	0.00045	1	1	C	G	T	G	C	**C**	2
n	0.00045	1	1	**T**	G	T	G	**G**	G	1
p	0.00045	1	1	C	G	T	G	C	G	1
o	0.00045	1	1	C	G	T	G	**G**	G	10

Table 13.29 - IGHV1-3

h.type	cum.freq	# indvs	pops	14:106471253 rs190814253 Cys->Tyr	14:106471268 rs1143505 Met->Thr	14:106471280 rs141136074 Arg->Thr	14:106471296 rs113948114 Glu->Gln	14:106471310 rs191691305 Ser->Asn	14:106471353 rs34069216 Glu->Lys	14:106471361 rs187612074 Tyr->Cys	14:106471365 rs191777242 Lys->Gln
ref	0.34155	-	-	C	A	C	C	C	C	T	T
a	0.3564	555	14	C	G	C	C	C	T	T	T
b	0.0724	155	13	T	A	G	C	C	C	T	T
d	0.0439	97	13	T	A	G	C	C	T	T	T
c	0.02865	63	12	C	G	C	C	C	T	T	T
h	0.019	42	12	C	A	C	C	C	T	T	T
i	0.0113	25	11	C	A	C	C	C	T	T	T
g	0.0105	23	8	C	G	C	G	C	T	T	T
f	0.00995	22	9	C	A	C	C	C	T	T	T
j	0.0077	16	8	C	A	C	C	C	C	T	T
e	0.0073	16	4	T	A	C	C	C	C	T	T
n	0.0072	16	10	C	G	C	C	C	C	T	T
t	0.00675	15	8	T	A	G	C	C	T	T	T
r	0.00585	13	9	C	G	C	C	C	T	T	T
m	0.00585	13	6	T	A	G	C	C	T	T	T
l	0.00585	13	8	C	A	C	C	C	C	T	T
o	0.00495	11	5	C	G	C	C	C	T	T	T
p	0.0045	10	4	T	A	G	C	C	C	T	T
ab	0.00405	9	6	C	A	C	C	C	T	T	T
k	0.0036	8	5	C	G	C	C	C	T	T	T
x	0.00315	7	6	C	A	C	G	C	C	T	T
ah	0.0027	6	5	C	A	C	G	C	T	T	T
u	0.00225	5	2	C	G	C	C	C	T	T	T
v	0.00225	5	2	T	A	G	C	C	T	T	T
ap	0.00225	4	3	C	G	C	C	C	T	T	T
ac	0.00225	5	4	T	A	G	C	C	C	T	T
y	0.0018	4	3	C	A	G	C	C	C	T	T
ar	0.0018	4	3	C	G	C	C	C	T	T	T
ai	0.0018	4	3	C	A	C	G	C	T	T	T

Table 13.29 - IGHV1-3

aa	0.00135	3	1	C	G	C	C	C	T	T	T
ak	0.00135	3	3	C	A	C	C	C	C	T	T
w	0.00135	3	1	C	G	C	C	C	C	T	T
at	0.00135	3	2	T	A	G	C	C	T	T	T
ag	0.0009	2	2	C	A	G	C	C	T	T	T
aj	0.0009	2	2	T	G	C	C	C	T	T	T
aq	0.0009	2	1	T	A	G	C	C	C	T	T
s	0.0009	2	2	T	A	G	G	C	C	T	T
q	0.0009	2	1	C	G	C	G	C	T	T	T
z	0.0009	2	1	C	G	C	C	C	C	T	T
ad	0.0009	2	2	C	G	C	C	C	C	T	T
aw	0.0009	2	2	C	A	C	G	C	T	T	T
ae	0.0009	2	2	C	G	C	G	C	C	T	T
an	0.0009	2	2	T	A	C	C	C	T	T	T
bd	0.00045	1	1	C	G	C	C	C	C	T	T
ba	0.00045	1	1	C	G	C	C	C	T	T	T
ax	0.00045	1	1	C	A	C	C	C	C	T	G
bb	0.00045	1	1	C	G	C	C	C	T	T	G
az	0.00045	1	1	C	G	C	G	C	T	T	T
al	0.00045	1	1	T	A	G	C	C	C	T	T
be	0.00045	1	1	C	G	C	C	C	T	T	T
bf	0.00045	1	1	C	A	C	G	C	T	T	T
av	0.00045	1	1	C	A	C	C	C	C	C	G
ao	0.00045	1	1	T	A	G	C	C	C	T	T
af	0.00045	1	1	C	A	C	C	C	C	T	T
au	0.00045	1	1	C	G	C	C	C	T	T	T
bh	0.00045	1	1	T	A	C	C	C	T	T	T
am	0.00045	1	1	C	A	C	C	T	C	T	T
bc	0.00045	1	1	C	G	C	C	C	T	T	T
bg	0.00045	1	1	T	A	G	C	C	C	T	G
ay	0.00045	1	1	C	G	C	C	T	T	T	T
as	0.00045	1	1	T	A	G	C	C	C	T	G

Table 13.29 - IGHV1-3

h.type	cum.freq	# indvs	pops	14:106471383 rs184320740 Ala->Thr	14:106471388 rs34874585 Ser->Ile	14:106471421 rs74091631 Ala->Asp	14:106471503 rs186640498 Lys->Gln	14:106471508 rs192340916 Val->Glu	14:106471547 rs1143496 Val->Ala	ENST00000390595
ref	0.34155	-	-	C	C	G	T	A	A	
a	0.3564	555	14	C	[redacted]	G	T	A	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0724	155	13	C	C	G	T	A	A	1,2,3,4,5,6,7,9,10,11,12,13,14
d	0.0439	97	13	C	[redacted]	G	T	A	G	1,2,3,4,5,6,7,9,10,11,12,13,14
c	0.02865	63	12	C	[redacted]	G	T	A	A	1,2,3,4,5,6,7,9,11,12,13,14
h	0.019	42	12	C	C	G	T	A	A	1,2,3,4,5,6,7,9,10,11,13,14
i	0.0113	25	11	C	[redacted]	G	T	A	G	1,2,3,5,6,7,10,11,12,13,14
g	0.0105	23	8	C	[redacted]	G	T	A	G	2,3,4,6,7,9,11,13
f	0.00995	22	9	C	[redacted]	G	T	A	A	2,3,5,7,8,9,11,13,14
j	0.0077	16	8	C	[redacted]	G	T	A	A	1,4,5,6,7,10,12,14
e	0.0073	16	4	C	C	G	T	A	A	1,10,12,14
n	0.0072	16	10	C	C	G	T	A	A	2,4,5,6,7,9,10,12,13,14
t	0.00675	15	8	C	C	G	T	A	A	3,4,6,9,10,11,13,14
r	0.00585	13	9	C	C	G	T	A	A	3,4,5,6,7,10,12,13,14
m	0.00585	13	6	C	[redacted]	G	T	A	A	1,2,10,11,13,14
l	0.00585	13	8	C	C	G	T	A	G	1,3,5,6,7,9,10,13
o	0.00495	11	5	C	C	G	T	A	G	2,3,6,7,13
p	0.0045	10	4	C	[redacted]	G	T	A	A	1,9,10,14
ab	0.00405	9	6	C	C	G	T	A	G	4,6,8,9,10,13
k	0.0036	8	5	C	[redacted]	G	T	T	G	2,6,7,12,13
x	0.00315	7	6	C	C	G	T	A	A	3,4,7,9,10,11
ah	0.0027	6	5	C	[redacted]	G	T	A	G	3,5,6,9,13
u	0.00225	5	2	C	[redacted]	T	T	A	G	1,14
v	0.00225	5	2	C	[redacted]	T	T	A	G	1,14
ap	0.00225	4	3	C	[redacted]	G	G	A	G	3,4,9
ac	0.00225	5	4	C	[redacted]	G	T	A	G	1,10,12,14
y	0.0018	4	3	C	C	G	T	A	A	1,3,12
ar	0.0018	4	3	T	[redacted]	G	T	A	G	1,10,13
ai	0.0018	4	3	C	C	G	T	A	G	2,4,7

Table 13.29 - IGHV1-3

aa	0.00135	3	1	C	■	T	T	A	A	14
ak	0.00135	3	3	C	■	G	T	A	G	1,9,10
w	0.00135	3	1	C	C	G	T	A	A	9
at	0.00135	3	2	C	■	G	T	A	G	6,10
ag	0.0009	2	2	C	■	G	T	A	A	10,14
aj	0.0009	2	2	T	C	G	T	A	G	3,4
aq	0.0009	2	1	C	C	G	T	A	A	10
s	0.0009	2	2	C	C	G	T	A	A	7,10
q	0.0009	2	1	C	C	G	T	A	A	2
z	0.0009	2	1	T	C	G	T	A	A	5
ad	0.0009	2	2	C	■	G	T	A	G	5,9
aw	0.0009	2	2	C	C	G	T	A	A	7,13
ae	0.0009	2	2	C	C	G	T	A	A	5,10
an	0.0009	2	2	C	■	G	T	A	G	12,14
bd	0.00045	1	1	C	C	T	T	A	A	10
ba	0.00045	1	1	C	C	G	G	A	G	3
ax	0.00045	1	1	C	■	G	T	A	G	5
bb	0.00045	1	1	C	■	G	T	A	A	8
az	0.00045	1	1	C	C	G	T	A	G	4
al	0.00045	1	1	C	C	G	T	A	G	10
be	0.00045	1	1	T	C	G	T	A	A	14
bf	0.00045	1	1	C	■	G	T	A	A	2
av	0.00045	1	1	C	C	G	T	A	A	13
ao	0.00045	1	1	C	■	T	T	A	G	1
af	0.00045	1	1	T	C	G	T	A	A	13
au	0.00045	1	1	C	C	T	T	A	A	14
bh	0.00045	1	1	C	■	G	T	A	A	14
am	0.00045	1	1	C	C	G	T	A	A	10
bc	0.00045	1	1	T	■	G	T	A	A	10
bg	0.00045	1	1	C	■	G	T	A	A	10
ay	0.00045	1	1	C	■	G	T	A	G	4
as	0.00045	1	1	C	■	G	T	A	G	10

Table 13.30 - IGHV1-45

h.type	cum.freq	# indvs	pops	14:106962947 rs186738824 Met->Thr	14:106962984 rs145326338 Met->Leu	14:106963101 rs7141669 Gly->Arg	14:106963167 rs182373995 Val->Phe	14:106963326 rs12884037 Ala->Val	ENST00000390621
ref	0.75745	-	-	A	T	C	C	G	
a	0.1286	251	12	A	T	**T**	C	G	1,2,3,4,5,6,7,9,10,11,12,13
b	0.1126	220	12	A	T	C	C	■	1,2,3,5,6,7,8,10,11,12,13,14
d	0.00045	1	1	**G**	T	C	C	G	1
e	0.00045	1	1	A	**G**	C	C	G	9
c	0.00045	1	1	A	T	C	■	G	3

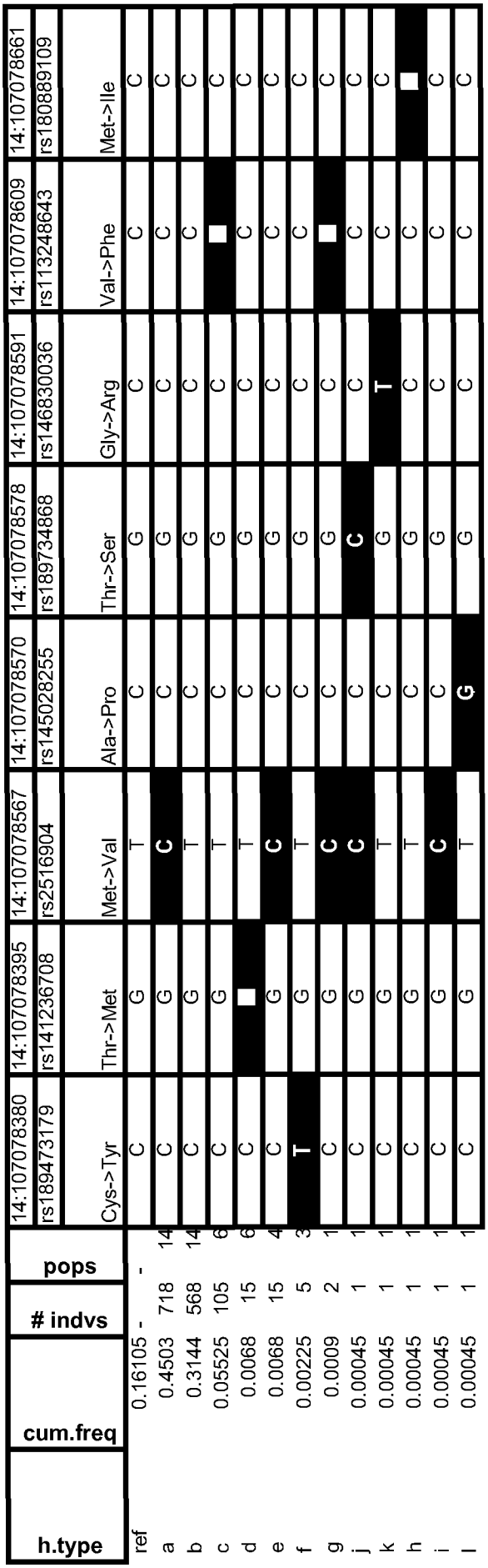
Table 13.31 - IGHV1-46

h.type	cum.fre	# indvs	pops	14:106967122	14:106967151	14:106967155	14:106967171	14:106967173	14:106967177	14:106967241	14:106967249
				rs147211698	rs55801711	rs149338091	rs181189514	rs185595166	rs190309173	rs144704015	rs182132309
				Thr->Met	Phe->Leu	Lys->Arg	Thr->Ala	Ser->Thr	Gly->Ser	Met->Ile	Tyr->Asn
ref	0.9778	-	-	G	G	T	T	C	C	C	A
a	0.0082	16	6	G	**C**	T	T	C	C	C	A
b	0.0032	7	3	■	G	T	T	C	C	C	A
c	0.00225	5	3	G	G	T	T	C	C	C	A
d	0.0018	4	4	G	G	T	T	C	C	**T**	A
k	0.0009	2	2	G	G	**C**	T	C	C	C	A
e	0.0009	2	2	G	G	T	T	**G**	C	C	A
n	0.0009	2	2	G	G	T	T	C	C	C	A
j	0.00045	1	1	G	G	T	T	C	C	C	A
g	0.00045	1	1	G	G	T	T	**G**	C	C	**T**
h	0.00045	1	1	G	G	T	T	C	C	C	A
f	0.00045	1	1	G	G	T	T	C	C	C	A
i	0.00045	1	1	G	G	T	T	C	C	C	**T**
m	0.00045	1	1	G	G	**C**	T	C	**T**	C	A
l	0.00045	1	1	G	G	T	**C**	C	C	C	A
p	0.00045	1	1	G	G	T	T	C	C	C	A
o	0.00045	1	1	G	G	T	T	**G**	C	**T**	A

Table 13.31 - IGHV1-46

h.type	cum.freq	# indivs	# pops	14:106967264 rs187613260 Tyr->Asp	14:106967273 rs192285778 Ala->Ser	14:106967275 rs61995748 Lys->Met	14:106967285 rs61747196 Val->Ile	14:106967306 rs188566927 Lys->Gln	14:106967451 rs191694446 Ala->Thr	14:106967453 rs143810901 Leu->Pro	ENST00000390622
ref	0.9778	-	-	A	C	T	C	T	C	A	
a	0.0082	16	6	A	C	T	C	T	C	A	1,5,7,8,12,13
b	0.0032	7	3	A	C	T	C	T	C	A	3,10,14
c	0.00225	5	3	A	C	[redacted]	C	T	C	A	1,10,14
d	0.0018	4	4	A	C	T	C	T	C	A	2,10,11,14
k	0.0009	2	2	A	C	T	C	T	C	A	1,10
e	0.0009	2	2	A	C	T	C	T	C	A	6,10
n	0.0009	2	2	A	C	T	T	T	C	A	4,6
j	0.00045	1	1	A	[redacted]	T	C	T	C	A	1
g	0.00045	1	1	A	C	T	C	T	C	A	5
h	0.00045	1	1	A	C	T	T	G	C	A	5
f	0.00045	1	1	A	C	T	T	T	T	A	10
i	0.00045	1	1	A	C	T	C	T	C	A	9
m	0.00045	1	1	A	C	T	C	T	C	A	3
l	0.00045	1	1	A	C	T	C	T	C	A	3
p	0.00045	1	1	A	C	T	C	G	C	A	4
o	0.00045	1	1	C	C	T	C	T	C	G	10

Table 13.32 - IGHV1-58



h.type	cum.freq	# indvs	pops	14:107078380 rs189473179 Cys->Tyr	14:107078395 rs141236708 Thr->Met	14:107078567 rs2516904 Met->Val	14:107078570 rs145028255 Ala->Pro	14:107078578 rs189734868 Thr->Ser	14:107078591 rs146830036 Gly->Arg	14:107078609 rs113248643 Val->Phe	14:107078661 rs180889109 Met->Ile
ref	0.16105	-	-	C	G	T	C	G	C	C	C
a	0.4503	718	14	C	G	**C**	C	G	C	C	C
b	0.3144	568	14	C	G	T	C	G	C	C	C
c	0.05525	105	6	C	G	T	C	G	C	■	C
d	0.0068	15	6	C	■	T	C	G	C	C	C
e	0.0068	15	4	C	G	**C**	C	G	C	C	C
f	0.00225	5	3	**T**	G	T	C	G	C	C	C
g	0.0009	2	1	C	G	**C**	C	G	C	■	C
j	0.00045	1	1	C	G	**C**	C	**C**	C	C	C
k	0.00045	1	1	C	G	T	C	G	**T**	C	C
h	0.00045	1	1	C	G	T	C	G	C	C	■
i	0.00045	1	1	C	G	**C**	C	G	C	C	C
l	0.00045	1	1	C	G	T	**G**	G	C	C	C

Table 13.32 - IGHV1-58

h.type	cum.freq	# indvs	pops	14:107078674 rs186300171 Ala->Val	14:107078775 rs148981028 Gly->Arg	14:107078790 rs1858692 Val->Ile	ENST00000390628
ref	0.16105	-	-	G	C	C	
a	0.4503	718	14	G	C	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.3144	568	14	G	C	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.05525	105	6	G	C	T	1,5,10,12,13,14
d	0.0068	15	6	G	C	C	3,4,5,9,11,12
e	0.0068	15	4	G	T	T	1,10,13,14
f	0.00225	5	3	G	C	T	3,4,9
g	0.0009	2	1	G	C	T	10
j	0.00045	1	1	G	C	T	10
k	0.00045	1	1	G	C	C	13
h	0.00045	1	1	■	C	C	10
i	0.00045	1	1	■	C	T	1
l	0.00045	1	1	G	C	C	13

Table 13.33 - IGHV1-69

h.type	cum.freq	# indvs	pops	14:107169948 rs190613018 Val->Leu	14:107169992 rs183124973 Thr->Arg	14:107169995 rs1064565 Ser->Thr	14:107170004 rs191670994 Lys->Arg	14:107170011 rs74089691 Ala->Thr	14:107170014 rs11557979 Thr->Ser	14:107170019 rs192735419 Thr->Lys	14:107170025 rs185522094 Arg->Thr
ref	0.49795	-	-	C	G	C	T	C	T	G	C
a	0.1689	323	14	C	G	C	T	C	T	G	C
b	0.1543	312	14	C	G	C	T	C	T	G	C
c	0.0561	118	13	C	G	C	T	C	T	G	C
g	0.0348	76	7	C	G	C	T	**T**	T	G	C
e	0.014	31	4	C	G	C	T	**T**	T	G	C
l	0.0108	24	6	C	G	C	T	C	T	G	C
d	0.00995	22	3	C	G	C	T	C	T	G	C
h	0.00765	17	10	C	G	C	T	C	T	G	C
f	0.00685	15	8	C	G	C	T	C	T	G	C
i	0.00675	15	6	C	G	C	T	C	T	G	C
m	0.0027	6	3	C	G	C	T	C	T	G	C
j	0.00225	5	3	C	G	C	T	C	T	G	C
u	0.00225	5	4	C	G	C	T	C	T	G	C
s	0.0018	4	3	C	G	C	T	C	T	G	C
k	0.00135	3	1	C	G	C	T	**T**	T	G	C
ac	0.00135	3	3	C	G	C	T	C	T	G	C
ah	0.0009	2	2	C	G	**G**	T	C	T	G	C
t	0.0009	2	2	C	G	C	T	C	T	G	C
ao	0.0009	2	1	C	G	C	T	**T**	T	G	C
af	0.0009	2	2	C	G	**G**	T	C	T	G	C
ad	0.0009	2	2	C	G	C	T	C	T	G	C
n	0.0009	2	2	C	G	C	T	C	T	G	C
an	0.0009	2	2	C	G	C	T	**T**	T	G	C
o	0.0009	2	2	C	G	C	T	C	T	G	C
ba	0.00045	1	1	C	G	C	T	C	T	G	C
ax	0.00045	1	1	C	G	C	T	**T**	T	G	C
ag	0.00045	1	1	C	G	C	T	C	T	G	C
y	0.00045	1	1	C	G	C	T	C	T	G	C

Table 13.33 - IGHV1-69

az	0.00045	1	1	C	G	C	T	T	T	G	C
aj	0.00045	1	1	C	G	C	T	C	T	G	G
aa	0.00045	1	1	C	G	C	T	C	T	G	C
aq	0.00045	1	1	C	G	C	T	C	T	G	C
v	0.00045	1	1	C	G	C	T	C	T	G	C
al	0.00045	1	1	C	G	C	T	C	T	G	C
ab	0.00045	1	1	C	G	C	T	C	T	G	C
ak	0.00045	1	1	G	G	C	T	C	T	G	C
q	0.00045	1	1	C	G	C	T	C	T	G	C
ap	0.00045	1	1	C	G	C	C	C	T	T	C
z	0.00045	1	1	C	G	C	T	C	T	G	C
w	0.00045	1	1	C	G	C	T	C	T	G	C
av	0.00045	1	1	C	C	C	T	C	T	G	C
r	0.00045	1	1	C	G	C	T	C	T	G	C
x	0.00045	1	1	C	G	C	T	C	T	G	C
au	0.00045	1	1	G	G	C	T	C	T	G	C
am	0.00045	1	1	C	G	C	T	T	T	G	C
at	0.00045	1	1	C	G	G	T	C	T	G	C
aw	0.00045	1	1	C	G	C	T	C	■	G	C
ay	0.00045	1	1	C	G	C	T	C	T	G	C
ar	0.00045	1	1	C	G	C	T	C	T	G	C
p	0.00045	1	1	C	G	C	T	C	T	G	C
as	0.00045	1	1	C	G	C	T	C	T	G	C
ae	0.00045	1	1	C	G	C	T	C	T	G	C
ai	0.00045	1	1	C	G	G	T	C	T	G	C

Table 13.33 - IGHV1-69

h.type	cum.freq	# indvs	pops	14:107170028 rs11557995 Gly->Ala	14:107170034 rs181467093 Phe->Ser	14:107170055 rs8009570 Thr->Ile	14:107170056 rs184214708 Thr->Ala	14:107170062 rs55891010 Phe->Leu	14:107170077 rs11845244 Gly->Arg	14:107170091 rs189019973 Leu->Pro	14:107170121 rs1064564 Ser->Asn
ref	0.49795	-	-	C	A	G	T	A	C	A	C
a	0.1689	323	14	C	A	■	T	**G**	**T**	A	C
b	0.1543	312	14	C	A	■	T	**G**	**T**	A	C
c	0.0561	118	13	C	A	■	T	**G**	C	A	C
g	0.0348	76	7	C	A	G	T	A	C	A	C
e	0.014	31	4	C	A	G	T	A	**T**	A	C
l	0.0108	24	6	C	A	G	T	A	**T**	A	C
d	0.00995	22	3	C	A	G	T	A	**T**	A	C
h	0.00765	17	10	C	A	■	T	A	C	A	C
f	0.00685	15	8	C	A	■	T	A	**T**	A	C
i	0.00675	15	6	C	A	■	T	A	**T**	A	C
m	0.0027	6	3	C	A	G	T	A	C	A	C
j	0.00225	5	3	C	A	G	T	**G**	**T**	A	C
u	0.00225	5	4	C	A	■	T	**G**	C	A	C
s	0.0018	4	3	C	A	G	T	A	C	A	**T**
k	0.00135	3	1	C	A	■	T	**G**	**T**	A	C
ac	0.00135	3	3	C	A	G	T	A	C	A	C
ah	0.0009	2	2	C	A	G	T	A	C	A	C
t	0.0009	2	2	C	A	G	T	A	C	A	C
ao	0.0009	2	1	C	A	■	T	**G**	C	A	C
af	0.0009	2	2	C	A	■	T	**G**	C	A	C
ad	0.0009	2	2	C	A	G	T	A	C	A	C
n	0.0009	2	2	C	A	■	T	**G**	**T**	A	**T**
an	0.0009	2	2	C	A	■	T	**G**	**T**	A	C
o	0.0009	2	2	C	A	■	T	**G**	**T**	A	C
ba	0.00045	1	1	C	A	G	T	A	C	A	C
ax	0.00045	1	1	C	A	G	T	A	C	A	C
ag	0.00045	1	1	C	A	■	T	**G**	**T**	A	C
y	0.00045	1	1	C	A	■	T	**G**	**T**	**G**	C

Table 13.33 - IGHV1-69

az	0.00045	1	1	C	A	G	T	A	T	A	C
aj	0.00045	1	1	C	A	G	C	A	C	A	T
aa	0.00045	1	1	C	A		T	A	C	A	C
aq	0.00045	1	1	C	A	G	C	A	C	A	C
v	0.00045	1	1	C	A		T	G	T	A	C
al	0.00045	1	1	C	A		T	G	T	A	C
ab	0.00045	1	1	C	A		T	G	C	A	T
ak	0.00045	1	1	C	A	G	T	A	C	A	C
q	0.00045	1	1	C	A	G	T	A	C	A	C
ap	0.00045	1	1	C	A	G	T	A	C	A	C
z	0.00045	1	1	C	A	G	T	G	C	A	C
w	0.00045	1	1	C	A		T	G	T	A	T
av	0.00045	1	1	C	A		T	G	T	A	C
r	0.00045	1	1	C	A		T	G	T	A	C
x	0.00045	1	1	C	A		T	G	T	A	C
au	0.00045	1	1	C	A		T	G	T	A	C
am	0.00045	1	1	C	A	G	T	A	C	A	C
at	0.00045	1	1	C	A		T	G	T	A	C
aw	0.00045	1	1	C	G		T	G	T	A	T
ay	0.00045	1	1	C	A	G	C	A	C	A	T
ar	0.00045	1	1	C	A		T	G	T	A	C
p	0.00045	1	1	C	A		T	G	T	A	C
as	0.00045	1	1	G	A		T	G	C	A	C
ae	0.00045	1	1	C	A	G	T	A	C	A	C
ai	0.00045	1	1	C	A		T	G	T	A	C

Table 13.33 - IGHV1-69

h.type	cum.freq	# indvs	pops	14:107170127 rs11557984 Ala->Val	14:107170128 rs11558005 Ala->Thr	14:107170130 rs140392997 Tyr->Phe	14:107170133 rs138194290 Ser->Asn	14:107170139 rs11558018 Phe->Ser	14:107170143 rs186595111 Thr->Pro	14:107170145 rs189122962 Gly->Asp	14:107170173 rs150196675 Val->Leu
ref	0.49795	-	-	G	C	T	C	A	T	C	C
a	0.1689	323	14	G	C	T	C	A	T	C	C
b	0.1543	312	14	G	T	T	C	A	T	C	C
c	0.0561	118	13	G	C	T	C	A	T	C	C
g	0.0348	76	7	G	C	T	C	A	T	C	C
e	0.014	31	4	G	C	T	C	A	T	C	C
l	0.0108	24	6	G	C	T	C	A	T	C	C
d	0.00995	22	3	G	T	T	C	A	T	C	C
h	0.00765	17	10	G	C	T	C	A	T	C	C
f	0.00685	15	8	G	T	T	C	A	T	C	C
i	0.00675	15	6	G	C	T	C	A	T	C	C
m	0.0027	6	3	G	C	T	C	A	T	C	C
j	0.00225	5	3	G	C	T	C	A	T	C	C
u	0.00225	5	4	G	T	T	C	A	T	C	C
s	0.0018	4	3	G	C	T	C	A	T	C	C
k	0.00135	3	1	G	T	T	C	A	T	C	C
ac	0.00135	3	3	G	T	T	C	A	T	C	C
ah	0.0009	2	2	G	C	T	C	A	T	C	C
t	0.0009	2	2	■	C	T	C	A	T	C	C
ao	0.0009	2	1	G	C	T	C	A	T	C	C
af	0.0009	2	2	G	C	T	C	A	T	C	C
ad	0.0009	2	2	G	C	T	T	A	T	C	C
n	0.0009	2	2	G	C	T	C	A	T	C	C
an	0.0009	2	2	G	C	T	C	A	T	C	C
o	0.0009	2	2	■	T	T	C	A	T	C	C
ba	0.00045	1	1	G	C	T	C	A	T	C	C
ax	0.00045	1	1	G	C	T	C	A	T	C	G
ag	0.00045	1	1	G	T	■	T	A	T	C	C
y	0.00045	1	1	G	C	T	C	A	T	C	C

Table 13.33 - IGHV1-69

az	0.00045	1	1	G	C	T	T	A	T	C	C
aj	0.00045	1	1	G	C	T	C	A	T	C	C
aa	0.00045	1	1	G	T	T	C	A	T	C	C
aq	0.00045	1	1	G	C	T	C	A	T	C	C
v	0.00045	1	1	G	C	T	C	A	G	C	C
al	0.00045	1	1	G	C	T	C	A	T	C	C
ab	0.00045	1	1	G	C	T	T	A	T	C	C
ak	0.00045	1	1	G	C	T	C	A	T	C	C
q	0.00045	1	1	G	C	T	C	G	T	C	C
ap	0.00045	1	1	G	C	T	C	A	T	C	C
z	0.00045	1	1	G	C	T	C	A	T	C	C
w	0.00045	1	1	G	T	T	C	A	T	C	C
av	0.00045	1	1	G	C	T	C	A	T	C	C
r	0.00045	1	1		C	T	C	A	T	C	C
x	0.00045	1	1	G	C	T	T	A	T	C	C
au	0.00045	1	1	G	C	T	C	A	T	C	C
am	0.00045	1	1	G	T	T	C	A	T	C	C
at	0.00045	1	1	G	C	T	C	A	T	C	C
aw	0.00045	1	1	G	C	T	C	A	T	C	C
ay	0.00045	1	1	G	C	T	C	A	T	C	C
ar	0.00045	1	1	G	C	T	C	A	T	C	C
p	0.00045	1	1	G	T	T	T	A	T	C	C
as	0.00045	1	1	G	C	T	C	A	T	C	C
ae	0.00045	1	1	G	C	T	C	A	T	T	C
ai	0.00045	1	1	G	T	T	C	A	T	C	C

Table 13.33 - IGHV1-69

h.type	cum.freq	# indvs	dops	14:107170185 rs181067247 Pro->Thr	14:107170326 . ->Ala	14:107170343 rs190658595 Phe->Val	ENST00000390633
ref	0.49795	-	-	G	AGCT	A	
a	0.1689	323	14	G	AGCT	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.1543	312	14	G	AGCT	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0561	118	13	G	AGCT	A	1,2,3,4,5,6,7,9,10,11,12,13,14
g	0.0348	76	7	G	AGCT	A	1,5,10,11,12,13,14
e	0.014	31	4	G	AGCT	A	1,10,12,14
l	0.0108	24	6	G	AGCT	A	1,3,8,10,12,14
d	0.00995	22	3	G	AGCT	A	1,10,14
h	0.00765	17	10	G	AGCT	A	1,2,3,4,6,7,10,11,12,13
f	0.00685	15	8	G	AGCT	A	1,3,5,6,9,10,11,14
i	0.00675	15	6	G	AGCT	A	3,4,5,9,10,12
m	0.0027	6	3	G	■	A	10,12,14
j	0.00225	5	3	G	AGCT	A	5,10,14
u	0.00225	5	4	G	AGCT	A	1,2,8,14
s	0.0018	4	3	G	AGCT	A	6,8,10
k	0.00135	3	1	G	AGCT	A	14
ac	0.00135	3	3	G	AGCT	A	2,10,14
ah	0.0009	2	2	G	AGCT	A	11,13
t	0.0009	2	2	G	AGCT	A	2,9
ao	0.0009	2	1	G	AGCT	A	10
af	0.0009	2	2	G	AGCT	A	10,14
ad	0.0009	2	2	G	AGCT	A	2,7
n	0.0009	2	2	G	AGCT	A	2,3
an	0.0009	2	2	G	AGCT	A	10,14
o	0.0009	2	2	G	AGCT	A	3,10
ba	0.00045	1	1	T	AGCT	A	13
ax	0.00045	1	1	G	AGCT	A	10
ag	0.00045	1	1	G	AGCT	A	13
y	0.00045	1	1	G	AGCT	A	10

Table 13.33 - IGHV1-69

az	0.00045	1	1	G	AGCT	A	1	1	10
aj	0.00045	1	1	G	AGCT	A	1	1	7
aa	0.00045	1	1	G	AGCT	A	1	1	2
aq	0.00045	1	1	G	AGCT	A	1	1	9
v	0.00045	1	1	G	AGCT	A	1	1	11
al	0.00045	1	1	G	■	A	1	1	10
ab	0.00045	1	1	G	AGCT	A	1	1	14
ak	0.00045	1	1	G	AGCT	A	1	1	4
q	0.00045	1	1	G	AGCT	A	1	1	4
ap	0.00045	1	1	G	AGCT	A	1	1	13
z	0.00045	1	1	G	AGCT	A	1	1	14
w	0.00045	1	1	G	AGCT	A	1	1	14
av	0.00045	1	1	G	AGCT	A	1	1	4
r	0.00045	1	1	G	AGCT	A	1	1	8
x	0.00045	1	1	G	AGCT	A	1	1	2
au	0.00045	1	1	G	AGCT	A	1	1	2
am	0.00045	1	1	G	AGCT	A	1	1	1
at	0.00045	1	1	G	AGCT	A	1	1	3
aw	0.00045	1	1	G	AGCT	A	1	1	2
ay	0.00045	1	1	G	AGCT	A	1	1	10
ar	0.00045	1	1	G	AGCT	C	1	1	3
p	0.00045	1	1	G	AGCT	A	1	1	13
as	0.00045	1	1	G	AGCT	A	1	1	12
ae	0.00045	1	1	G	AGCT	A	1	1	13
ai	0.00045	1	1	G	AGCT	A	1	1	11

Table 13.34 - IGHV1-8

h.type	cum.fre	# indvs	pops	14:106539147	rs148477632	14:106539163	rs73371428	14:106539171	rs187649961	14:106539278	rs184294529	14:106539284	rs151035485	14:106539291	rs187522932	14:106539476	rs189071216	14:106539497	rs180817521
				Ile->Val		Met->Ile		Val->Ile		Tyr->Phe		Thr->Ile		Thr->Pro		Ala->Thr		Ile->Phe	
ref	0.8844	-	-	T		C		C		T		G		T		C		T	
a	0.1016	160	9	T		■		C		T		G		T		C		T	
c	0.00405	8	2	**C**		C		C		T		G		T		C		T	
b	0.0032	6	6	T		C		C		T		G		T		C		■	
d	0.0018	4	2	T		■		C		T		■		T		C		T	
e	0.00135	3	2	T		■		C		T		G		T		C		■	
f	0.00135	3	2	T		C		**T**		T		G		T		C		T	
h	0.0009	2	2	T		C		C		■		G		T		C		T	
j	0.00045	1	1	T		C		C		T		G		**G**		C		T	
g	0.00045	1	1	T		■		C		T		G		T		**T**		T	
i	0.00045	1	1	T		C		C		T		■		T		C		T	

Table 13.34 - IGHV1-8

h.type	cum.fre	# indvs	pops	ENST00000390599
ref	0.8844	-	-	
a	0.1016	160	9	1,5,7,8,10,11,12,13,14
c	0.00405	8	2	10,13
b	0.0032	6	6	2,5,6,11,12,13
d	0.0018	4	2	1,14
e	0.00135	3	2	2,11
f	0.00135	3	2	7,11
h	0.0009	2	2	3,5
j	0.00045	1	1	2
g	0.00045	1	1	14
i	0.00045	1	1	13

Table 13.35 - IGHV2-26

h.type	cum.freq	# indvs	pops	14:106757655	14:106757706	14:106757798	14:106757831	14:106757865	14:106757880	14:106758074	
				rs12586893	rs185534639	rs181476099	rs111777969	rs185965403	rs146488407	rs186573875	
				Arg->Trp	Leu->Phe	Ala->Glu	Arg->His	Leu->Val	Val->Ile	Thr->Ala	ENST00000390611
ref	0.9245	-	-	G	G	G	C	G	C	T	
a	0.06135	126	13	■	G	G	C	G	C	T	1,3,4,5,6,7,8,9,10,11,1 2,13,14
b	0.0087	15	5	G	G	G	**T**	G	C	T	1,6,10,11,14
c	0.0032	7	2	G	G	**T**	C	**C**	C	T	10,14
f	0.0009	2	2	G	G	G	C	G	**T**	T	13,14
d	0.00045	1	1	G	■	G	C	G	C	T	1
g	0.00045	1	1	G	G	G	C	**C**	C	T	5
e	0.00045	1	1	G	G	G	C	G	C	**C**	10

Table 13.36 - IGHV2-5

h.type	cum.fre	# indvs	pops	14:106494234	14:106494245	14:106494248	14:106494251	14:106494269	14:106494325	14:106494346	14:106494404
				rs1065554	rs1065552	rs181774040	rs141264800	rs12895651	rs150364725	rs185589849	rs190103562
				Ser->Arg	Ser->Thr	Pro->Thr	Ser->Gly	Asn->Asp	Gly->Ala	Ser->Thr	Leu->Val
ref	0.3179	-	-	G	A	G	T	T	C	C	G
a	0.6657	872	14	G	A	G	T	**C**	C	C	G
b	0.01055	18	3	G	A	G	**C**	**C**	C	C	G
d	0.0018	4	2	G	A	**T**	T	**C**	C	C	G
c	0.0009	2	2	G	A	G	T	**C**	**G**	C	G
j	0.00045	1	1	G	**T**	**T**	T	**C**	C	C	G
k	0.00045	1	1	G	A	G	T	T	C	**G**	G
g	0.00045	1	1	**C**	A	G	T	**C**	**G**	C	G
e	0.00045	1	1	G	A	G	T	**C**	C	**G**	G
h	0.00045	1	1	**C**	A	**T**	T	**C**	C	C	G
f	0.00045	1	1	G	A	G	T	**C**	C	C	**C**
i	0.00045	1	1	**C**	A	G	T	**C**	C	C	G

ENST00000390597
1,2,3,4,5,6,7,8,9,10,11,1
2,13,14
1,10,14
10,14
4,13
6
1
9
3
5
10
12

Table 13.37 - IGHV2-70

h.type	cum.freq	# indvs	pops	14:107178820 rs139383686 Ile->Met	14:107178837 rs144976346 Tyr->Asp	14:107178849 rs185317654 Asp->Tyr	14:107178875 rs188731942 Leu->Pro	14:107178878 rs192620545 Val->Ala	14:107178890 rs138836314 Lys->Arg	14:107178899 rs112655724 Asp->Ala	14:107178911 rs188243654 Thr->Ile	14:107178938 rs17113976 Tyr->Cys
ref	0.36095	-	-	T	A	C	A	A	T	T	G	T
a	0.31	584	14	T	A	C	A	A	T	T	G	T
b	0.16595	353	14	T	A	C	A	A	T	T	G	T
d	0.06035	132	13	T	A	C	A	A	T	**G**	G	**C**
c	0.04295	94	11	T	A	C	A	A	T	T	G	T
e	0.0319	70	6	T	A	C	A	A	T	T	G	**C**
f	0.00405	9	5	T	A	C	A	A	T	T	G	**C**
l	0.00225	5	4	T	A	C	**G**	A	T	T	G	T
y	0.00135	3	3	T	A	C	A	A	T	T	G	T
z	0.00135	3	3	**C**	A	C	A	A	T	T	G	T
af	0.00135	3	2	T	A	C	A	A	T	**G**	G	**C**
h	0.00135	3	3	T	A	C	A	A	T	**G**	G	**C**
g	0.0009	2	2	T	A	C	A	A	T	T	G	T
s	0.0009	2	2	T	A	C	A	A	T	T	G	T
ab	0.0009	2	2	T	A	C	A	A	**C**	T	G	T
ah	0.00045	1	1	**C**	**C**	C	A	A	T	T	G	T
ag	0.00045	1	1	T	A	C	A	A	T	**G**	G	**C**
j	0.00045	1	1	T	**C**	C	A	A	T	T	G	T
u	0.00045	1	1	T	A	C	A	A	T	T	G	**C**
k	0.00045	1	1	T	A	C	A	A	T	T	G	T
aj	0.00045	1	1	**C**	A	C	A	A	T	T	G	T
aa	0.00045	1	1	T	A	C	A	A	T	T	G	**C**
t	0.00045	1	1	T	A	C	A	A	T	T	G	**C**
aq	0.00045	1	1	T	A	C	A	A	T	**G**	G	**C**
v	0.00045	1	1	**C**	A	C	A	A	T	T	G	T
al	0.00045	1	1	T	A	C	A	A	**C**	T	G	T
ak	0.00045	1	1	T	A	C	A	A	T	T	G	**C**
q	0.00045	1	1	T	A	C	A	A	T	T	G	T
ap	0.00045	1	1	T	A	C	A	A	T	**G**	G	**C**

Table 13.37 - IGHV2-70

w	0.00045	1	T	C	C	A	A	T	T	G	T	
ao	0.00045	1	T	A	C	A	A	T	G	G	C	
r	0.00045	1	T	A	C	A	A	T	G	G	T	
x	0.00045	1	C	A	C	A	A	T	T	■	T	
ad	0.00045	1	T	A	C	A	G	T	G	G	C	
am	0.00045	1	T	A	C	A	A	C	T	G	C	
i	0.00045	1	T	A	C	A	A	T	T	G	T	
n	0.00045	1	T	A	■	A	A	T	T	G	T	
ac	0.00045	1	C	A	C	A	A	T	G	G	C	
m	0.00045	1	T	A	C	A	A	T	T	■	T	
ar	0.00045	1	T	A	C	A	A	T	T	G	T	
p	0.00045	1	T	A	C	A	A	T	T	G	T	
ae	0.00045	1	T	A	C	A	A	T	T	■	T	
ai	0.00045	1	T	A	C	A	A	T	T	G	C	
an	0.00045	1	T	A	C	A	A	T	T	G	T	
o	0.00045	1	T	C	C	A	A	T	G	G	C	

Table 13.37 - IGHV2-70

h.type	cum.freq	# indvs	pops	14:107178959 rs184814603 Asp->Gly	14:107178965 rs2073669 Leu->Arg	14:107178976 rs148856487 Glu->Asp	14:107179098 rs191880827 Gly->Ser	14:107179103 rs185635448 Glu->Val	14:107179106 rs10144703 Arg->Lys	14:107179115 rs182292946 Val->Gly	14:107179123 rs187704106 Leu->Phe	14:107179243 rs61734101 Thr->Met
ref	0.36095	-	-	T	A	C	C	T	C	A	T	G
a	0.31	584	14	T	C	C	C	T	C	A	T	G
b	0.16595	353	14	T	C	C	C	T	T	A	T	G
d	0.06035	132	13	T	C	C	C	T	T	A	T	■
c	0.04295	94	11	T	A	C	C	T	T	A	T	G
e	0.0319	70	6	T	C	C	C	T	T	A	T	■
f	0.00405	9	5	T	C	C	C	T	T	A	T	G
l	0.00225	5	4	T	C	C	C	T	C	A	T	G
y	0.00135	3	3	**C**	C	C	C	T	C	A	T	G
z	0.00135	3	3	T	C	C	C	T	T	A	T	G
af	0.00135	3	2	T	C	C	C	T	T	A	T	G
h	0.00135	3	3	T	C	C	C	T	T	A	T	■
g	0.0009	2	2	T	A	**G**	C	T	C	A	T	G
s	0.0009	2	2	T	C	C	**T**	T	C	A	T	G
ab	0.0009	2	2	T	A	C	C	T	C	A	T	G
ah	0.00045	1	1	T	C	C	C	T	C	A	T	G
ag	0.00045	1	1	**C**	C	C	C	T	**T**	A	T	■
j	0.00045	1	1	T	C	C	C	T	C	A	T	G
u	0.00045	1	1	T	A	C	C	T	C	A	T	G
k	0.00045	1	1	T	A	**G**	C	T	**T**	A	T	G
aj	0.00045	1	1	T	C	C	C	T	C	A	T	G
aa	0.00045	1	1	T	C	C	C	T	**T**	A	T	■
t	0.00045	1	1	T	A	C	C	T	**T**	A	T	G
aq	0.00045	1	1	T	C	C	C	■	**T**	A	T	■
v	0.00045	1	1	T	A	C	C	T	C	A	T	G
al	0.00045	1	1	T	C	C	C	T	C	A	T	G
ak	0.00045	1	1	T	C	C	C	T	**T**	A	■	■
q	0.00045	1	1	T	C	C	C	T	C	A	■	G
ap	0.00045	1	1	T	C	C	C	T	C	A	T	■

Table 13.37 - IGHV2-70

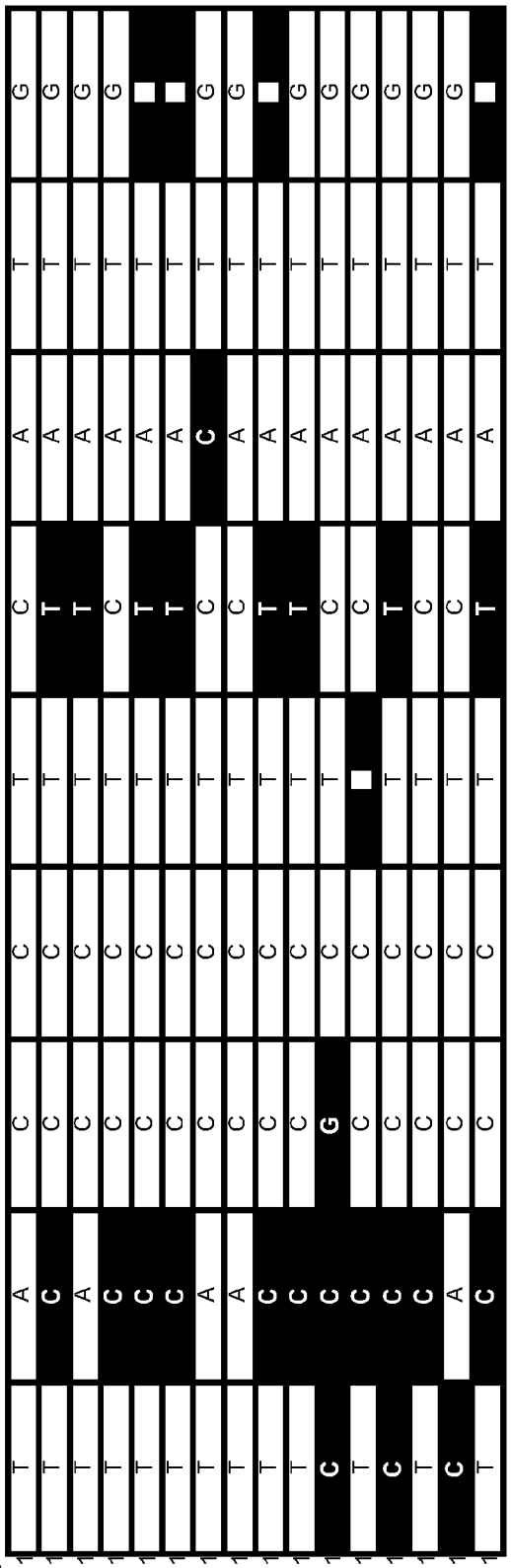


Table 13.37 - IGHV2-70

h.type	cum.freq	# indvs	pops	14:107179254	rs61734098	ENST00000390634
				Ile->Met		
ref	0.36095	-	-	T		
a	0.31	584	14	T		1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.16595	353	14	T		1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.06035	132	13	C		1,2,3,4,5,6,7,9,10,11,12,13,14
c	0.04295	94	11	T		1,2,5,6,7,8,10,11,12,13,14
e	0.0319	70	6	C		1,5,10,12,13,14
f	0.00405	9	5	T		1,5,9,10,14
l	0.00225	5	4	T		2,3,4,10
y	0.00135	3	3	T		7,9,12
z	0.00135	3	3	T		10,13,14
af	0.00135	3	2	T		6,10
h	0.00135	3	3	T		2,10,13
g	0.0009	2	2	T		4,6
s	0.0009	2	2	T		5,6
ab	0.0009	2	2	T		2,5
ah	0.00045	1	1	T		2
ag	0.00045	1	1	C		10
j	0.00045	1	1	T		10
u	0.00045	1	1	T		10
k	0.00045	1	1	T		13
aj	0.00045	1	1	T		10
aa	0.00045	1	1	T		14
t	0.00045	1	1	T		10
aq	0.00045	1	1	C		4
v	0.00045	1	1	T		13
al	0.00045	1	1	T		14
ak	0.00045	1	1	C		10
q	0.00045	1	1	T		1
ap	0.00045	1	1	C		10

Table 13.37 - IGHV2-70

w	0.00045	1	1	T	13									
ao	0.00045	1	1	C	1									
r	0.00045	1	1	T	2									
x	0.00045	1	1	T	4									
ad	0.00045	1	1	C	12									
am	0.00045	1	1	C	1									
i	0.00045	1	1	T	4									
n	0.00045	1	1	T	5									
ac	0.00045	1	1	T	13									
m	0.00045	1	1	T	12									
ar	0.00045	1	1	T	11									
p	0.00045	1	1	T	4									
ae	0.00045	1	1	T	13									
ai	0.00045	1	1	T	9									
an	0.00045	1	1	T	13									
o	0.00045	1	1	C	14									

Table 13.38 - IGHV3-11

h.type	cum.freq	# indvs	pops	14:106573345 rs186167416 Ala->Val	14:106573637 rs191020138 Lys->ArgLys->Arg	14:106573662 rs74207678 Trp->Arg	ENST00000390601
ref	0.9817	-	-	G	T	A	
a	0.01695	36	3	G	T	**G**	3,4,9
b	0.0009	2	2	■	T	A	3,5
c	0.00045	1	1	G	**C**	A	3

Table 13.39 - IGHV3-13

h.type	cum.freq	# indvs	pops	14:106586159	rs192397984	14:106586165	rs140926945	14:106586168	rs184992069	14:106586231	rs10151633	14:106586259	rs11625174	14:106586267	rs189564282	14:106586354	rs10131161	14:106586389	rs181545751
				G	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
ref	0.69725	-	-	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
a	0.15245	276	13	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	G	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
b	0.1011	167	9	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	T	Arg->Gln	T	Thr->Pro	G	Ala->Val	C	Ser->Cys	C	Gln->His
c	0.02925	59	10	G	Thr->Met	T	Gly->Glu	G	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
d	0.0105	19	4	■	Thr->Met	C	Gly->Glu	G	Ala->Gly	T	Arg->Gln	T	Thr->Pro	G	Ala->Val	C	Ser->Cys	C	Gln->His
e	0.00225	5	3	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	C	Ser->Cys	C	Gln->His
j	0.0018	4	2	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	T	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
g	0.00135	3	2	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	■	Gln->His
f	0.0009	2	2	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	T	Arg->Gln	T	Thr->Pro	■	Ala->Val	C	Ser->Cys	C	Gln->His
k	0.00045	1	1	G	Thr->Met	T	Gly->Glu	G	Ala->Gly	C	Arg->Gln	G	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
h	0.00045	1	1	■	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
i	0.00045	1	1	■	Thr->Met	C	Gly->Glu	G	Ala->Gly	T	Arg->Gln	T	Thr->Pro	G	Ala->Val	C	Ser->Cys	■	Gln->His
n	0.00045	1	1	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	G	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
m	0.00045	1	1	G	Thr->Met	C	Gly->Glu	G	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
l	0.00045	1	1	G	Thr->Met	C	Gly->Glu	C	Ala->Gly	C	Arg->Gln	T	Thr->Pro	G	Ala->Val	G	Ser->Cys	C	Gln->His
o	0.00045	1	1	G	Thr->Met	T	Gly->Glu	G	Ala->Gly	T	Arg->Gln	T	Thr->Pro	G	Ala->Val	C	Ser->Cys	C	Gln->His

Table 13.39 - IGHV3-13

h.type	cum.freq	# indvs	pops	14:106586557 rs112340466 Val->Ile	14:106586563 rs137975618 Phe->Leu	ENST00000390602
ref	0.69725	-	-	C	A	
a	0.15245	276	13	C	A	1,2,3,4,5,6,7,8,9,10,11,12,13
b	0.1011	167	9	C	A	1,5,7,8,10,11,12,13,14
c	0.02925	59	10	C	A	1,2,5,6,7,10,11,12,13,14
d	0.0105	19	4	C	A	1,10,12,14
e	0.00225	5	3	C	A	1,10,14
j	0.0018	4	2	C	A	1,10
g	0.00135	3	2	C	A	6,14
f	0.0009	2	2	C	A	10,12
k	0.00045	1	1	C	A	2
h	0.00045	1	1	C	A	1
i	0.00045	1	1	C	A	14
n	0.00045	1	1	C	**G**	11
m	0.00045	1	1	**T**	**G**	14
l	0.00045	1	1	C	A	3
o	0.00045	1	1	C	A	10

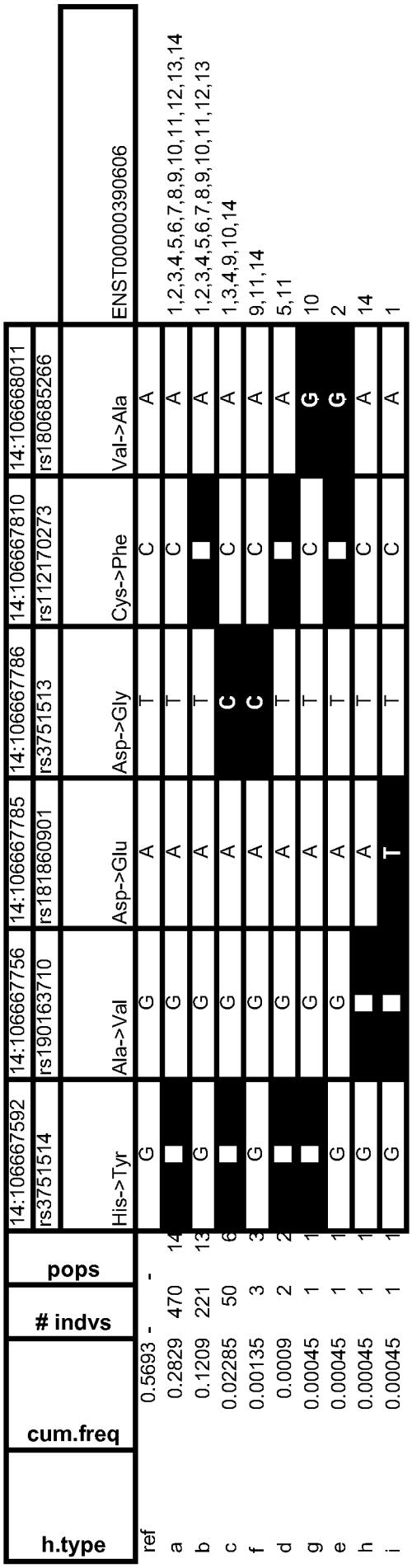
Table 13.40 - IGHV3-15

h.type	cum.freq	# indvs	pops	14:106610377 rs182550955 Asn->Thr	14:106610467 rs186087290 Gly->Ala	14:106610509 rs147758697 Ser->Asn	14:106610570 rs184104714 Gly->Arg	14:106610736 rs183000727 Ile->Phe	ENST00000390603
ref	0.91495	-	-	T	C	C	C	T	
a	0.0572	104	11	T	C	**T**	C	T	1,2,3,5,6,7,10,11,12,13,14
b	0.0265	56	5	T	C	C	**T**	T	1,3,4,9,13
d	0.00045	1	1	T	C	C	C	■	6
e	0.00045	1	1	**G**	C	C	C	T	6
c	0.00045	1	1	T	**G**	C	C	T	9

Table 13.41 - IGHV3-16

h.type	cum.fre	# indvs	pops	14:106622052 rs12323556 Glu->Lys	14:106622073 rs142354696 Lys->Glu	14:106622076 rs12433709 Arg->Cys	14:106622132 rs12323559 Arg->Ile	ENST00000390604
ref	0.8039	-	-	C	T	G	C	
a	0.11075	200	13	C	T	G	C	1,2,3,4,5,6,7,8,9,11,12,13,14
b	0.07085	122	8	**T**	T	G	**C**	1,5,7,10,11,12,13,14
c	0.01	21	4	C	**C**	G	**C**	1,10,11,14
d	0.00315	7	3	**T**	T	G	C	10,11,14
e	0.00135	3	2	C	**C**	G	C	10,12

Table 13.42 - IGHV3-20



h.type	cum.freq	# indvs	pops	14:106667592 rs3751514 His->Tyr	14:106667756 rs190163710 Ala->Val	14:106667785 rs181860901 Asp->Glu	14:106667786 rs3751513 Asp->Gly	14:106667810 rs112170273 Cys->Phe	14:106668011 rs180685266 Val->Ala	ENST00000390606
ref	0.5693	-	-	G	G	A	T	C	A	
a	0.2829	470	14	■	G	A	T	C	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.1209	221	13	G	G	A	T	■	A	1,2,3,4,5,6,7,8,9,10,11,12,13
c	0.02285	50	6	■	G	A	**C**	C	A	1,3,4,9,10,14
f	0.00135	3	3	G	G	A	**C**	C	A	9,11,14
d	0.0009	2	2	■	G	A	T	■	A	5,11
g	0.00045	1	1	■	G	A	T	C	**G**	10
e	0.00045	1	1	G	G	A	T	■	**G**	2
h	0.00045	1	1	G	■	A	T	C	A	14
i	0.00045	1	1	G	■	**T**	T	C	A	1

Table 13.43 - IGHV3-21

h.type	cum.freq	# indvs	pops	14:106691762	14:106691807	14:106691824	14:106691833	14:106691849	14:106691865	14:106691867	14:106691879
				rs183314182	rs188058888	rs143336904	rs192148312	rs184550642	rs192949902	rs184209619	rs180974163
				Thr->Ala	Ser->Gly	Val->Ala	Leu->Pro	Ala->Thr	Met->Ile	Met->Leu	Ser->Gly
ref	0.9892	-	-	T	T	A	A	C	C	T	T
a	0.00225	5	5	T	**C**	A	A	C	C	T	T
d	0.0018	4	3	T	T	A	A	C	C	T	T
c	0.0009	2	2	T	T	**G**	A	C	C	T	T
b	0.0009	1	1	T	T	A	A	C	C	T	T
h	0.0009	2	2	**C**	T	A	A	C	C	T	T
j	0.00045	1	1	T	T	A	**G**	C	C	T	T
k	0.00045	1	1	T	T	A	A	C	C	T	T
g	0.00045	1	1	T	**C**	A	A	C	C	T	T
e	0.00045	1	1	T	T	A	A	C	C	■	T
f	0.00045	1	1	T	T	A	A	C	**G**	■	**C**
i	0.00045	1	1	T	T	A	A	**T**	C	T	T
n	0.00045	1	1	T	**C**	A	A	**T**	C	T	T
m	0.00045	1	1	T	T	A	A	C	C	T	T
l	0.00045	1	1	T	T	A	A	C	C	T	T

Table 13.43 - IGHV3-21

h.type	cum.freq	# indvs	pops	14:106691884 rs184687518 Thr->Ile	14:106691896 rs187020320 Ala->Gly	14:106691899 rs191067794 Ala->Val	14:106691909 rs181216108 Leu->Phe	ENST00000390607
ref	0.9892	-	-	G	G	G	G	
a	0.00225	5	5	G	G	G	G	7,9,11,13,14
d	0.0018	4	3	■	G	G	G	1,5,14
c	0.0009	2	2	G	G	G	G	10,14
b	0.0009	1	1	■	**C**	G	G	1
h	0.0009	2	2	G	G	G	G	1,13
j	0.00045	1	1	■	G	G	G	2
k	0.00045	1	1	G	G	■	G	1
g	0.00045	1	1	■	G	G	G	5
e	0.00045	1	1	G	G	G	G	2
f	0.00045	1	1	G	**C**	■	■	4
i	0.00045	1	1	G	G	G	G	14
n	0.00045	1	1	G	G	G	G	3
m	0.00045	1	1	G	**C**	■	G	11
l	0.00045	1	1	G	**C**	G	G	1

Table 13.44 - IGHV3-23

h.type	cum.freq	# indvs	pops	14:106725320	14:106725328	14:106725329	14:106725346	14:106725357	14:106725397	14:106725400	14:106725433
				rs189958807	rs1064091	rs1064090	rs1055799	rs143624256	rs61752504	rs183319819	rs187384491
				Tyr->His	Gly->Asp	Gly->Ser	Ala->Val	Glu->Asp	Ala->Gly	Tyr->Ser	Ser->Phe
ref	0.83265	-	-	A	C	C	G	C	G	T	G
a	0.09825	190	12	A	C	C	G	C	G	T	G
b	0.02695	58	9	A	C	T	■	C	G	T	G
c	0.01365	29	9	A	C	T	G	C	G	T	G
d	0.00865	18	4	A	C	C	G	C	C	T	G
e	0.00455	10	4	A	C	T	■	C	G	T	G
f	0.00225	5	3	A	C	C	G	C	C	T	G
k	0.00225	5	5	A	T	C	G	C	G	T	G
g	0.00225	5	3	A	C	C	■	C	G	T	G
n	0.0018	4	3	A	C	C	G	C	C	T	G
j	0.0009	2	1	A	C	C	■	C	G	T	G
h	0.0009	2	2	A	C	C	G	C	G	G	G
u	0.00045	1	1	A	T	C	G	C	G	T	G
t	0.00045	1	1	A	T	C	G	G	G	T	G
v	0.00045	1	1	A	C	C	G	C	G	T	G
s	0.00045	1	1	A	T	C	G	C	C	T	G
q	0.00045	1	1	A	C	C	G	C	G	T	G
r	0.00045	1	1	A	C	T	G	C	G	T	G
i	0.00045	1	1	A	C	C	G	C	G	T	G
m	0.00045	1	1	A	C	C	G	C	G	T	G
l	0.00045	1	1	G	C	C	G	C	G	T	G
p	0.00045	1	1	A	T	C	G	C	G	G	G
o	0.00045	1	1	A	C	C	G	C	G	T	■

Table 13.44 - IGHV3-23

h.type	cum.freq	# indvs	pops	14:106725446 rs61750837 Ser->Thr	14:106725482 rs56069819 Leu->Val	14:106725490 rs191754191 Val->Gly	14:106725493 rs182117479 Glu->Gly	14:106725621 rs188250527 Ala->Pro	14:106725630 rs145913230 Phe->Val	14:106725634 rs181072648 Trp->Cys	ENST00000390609
ref	0.83265	-	-	A	A	A	T	C	A	C	
a	0.09825	190	12	A	**C**	A	T	C	A	C	1,2,3,5,6,7,8,10,11,12,13,14
b	0.02695	58	9	A	A	A	T	C	A	C	1,3,4,5,9,10,11,12,14
c	0.01365	29	9	A	A	A	T	C	A	C	1,3,4,5,6,9,10,11,14
d	0.00865	18	4	**T**	**C**	A	T	C	A	C	1,5,10,**14**
e	0.00455	10	4	A	**C**	A	T	C	A	C	1,10,11,**14**
f	0.00225	5	3	**T**	A	A	T	C	A	C	1,7,10
k	0.00225	5	5	A	A	A	T	C	A	C	1,5,9,10,11
g	0.00225	5	3	A	A	A	T	C	A	C	1,5,11
n	0.0018	4	3	A	A	A	T	C	A	C	2,9,14
j	0.0009	2	1	A	**C**	A	T	C	A	C	**14**
h	0.0009	2	2	A	A	A	T	C	A	C	6,10
u	0.00045	1	1	A	**C**	A	T	C	A	C	**14**
t	0.00045	1	1	A	A	A	T	C	A	C	13
v	0.00045	1	1	A	A	A	**C**	**G**	A	C	8
s	0.00045	1	1	**T**	**C**	A	T	C	A	C	10
q	0.00045	1	1	A	A	A	T	C	A	**G**	**14**
r	0.00045	1	1	A	**C**	A	T	C	A	C	**14**
i	0.00045	1	1	A	A	A	T	C	**C**	C	9
m	0.00045	1	1	A	A	**C**	T	C	A	C	6
l	0.00045	1	1	A	A	A	T	C	A	C	10
p	0.00045	1	1	A	A	A	T	C	A	C	2
o	0.00045	1	1	A	A	A	T	C	A	C	2

Table 13.45 - IGHV3-30

h.type	cum.freq	# indvs	pops	14:106791117 rs138563161 Ala->Val	14:106791127 rs190763974 Lys->Glu	14:106791150 rs183005459 Val->Ala	14:106791154 rs187117835 Ala->Thr	14:106791216 rs192185012 Thr->Ser	ENST00000390613
ref	0.97985	-	-	G	T	A	C	G	
a	0.0147	28	7	■	T	A	C	G	2,3,4,5,9,10,11
b	0.0023	5	4	G	T	A	C	**C**	1,2,9,13
d	0.00135	3	2	G	T	**G**	C	G	10,13
c	0.00135	3	3	G	**C**	A	C	G	3,9,11
e	0.00045	1	1	G	T	A	**T**	G	13

Table 13.46 - IGHV3-33

h.type	cum.fre	# indvs	pops	14:106815855 rs183940317 Asp->Gly	14:106815858 rs187379689 Tyr->Ser	14:106815862 rs112373679 Trp->Arg	14:106815868 rs111734923 Val->Phe	14:106815924 rs192341521 Ser->Asn	14:106815933 rs187304239 Thr->Ile	14:106815948 rs192234669 Ala->Val	14:106815970 rs138488134 Arg->Gly	ENST00000390615
ref	0.82905	-	-	T	T	A	C	C	G	G	T	
a	0.145	243	12	T	T	**G**	**■**	C	G	G	**C**	1,2,4,5,6,7,8,10,11,12,13,14
b	0.00965	21	10	T	T	**G**	**■**	C	G	G	T	1,2,5,6,7,10,11,12,13,14
c	0.0082	18	6	T	T	A	C	C	G	G	**C**	1,2,7,10,12,14
e	0.0027	6	3	**C**	T	A	C	C	G	G	T	1,10,14
d	0.0018	4	3	T	T	A	C	C	**■**	G	T	2,13,14
h	0.00135	3	3	T	T	A	C	**T**	G	G	T	1,6,10
f	0.0009	2	2	T	**G**	A	C	C	G	G	T	5,8
j	0.00045	1	1	**C**	T	A	C	**T**	G	**■**	T	10
g	0.00045	1	1	**C**	T	**G**	**■**	C	G	G	**C**	10
i	0.00045	1	1	T	T	**G**	**■**	**T**	G	G	**C**	1

Table 13.47 - IGHV3-35

h.type	cum.freq	# indvs	pops	14:106845417 rs112759551 Arg->Gln	14:106845487 rs74514616 Gly->Arg	14:106845753 rs145022218 Val->Leu	ENST00000390617
ref	0.78085	-	-	C	C	C	
a	0.11765	227	14	**T**	C	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.09695	190	10	C	**G**	C	1,3,4,5,6,7,9,10,12,14
c	0.0041	9	5	C	C	**G**	1,2,5,7,8
d	0.00045	1	1	**T**	**G**	C	14

Table 13.48 - IGHV3-38

h.type	cum.freq	# indvs	pops	14:106866449 rs78850219 Asn->Ser	14:106866451 rs144366955 Asn->Lys	14:106866664 rs150283006 Leu->Phe	14:106866828 rs112981911 Leu->Pro	ENST00000390618
ref	0.8889	-	-	T	G	C	A	
a	0.0723	158	14	**C**	G	C	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.02515	52	6	**C**	**C**	C	A	1,5,10,11,12,14
c	0.00725	16	5	T	G	**G**	A	1,10,11,12,14
d	0.0064	13	3	T	G	C	**G**	2,7,13

Table 13.49 - IGHV3-43

h.type	cum.freq	# indvs	pops	14:106926211 rs116899367 Ala->Thr	14:106926223 rs2467912 Thr->Ala	14:106926231 rs182842897 Ser->Asn	14:106926323 rs186760979 Asp->Glu	14:106926328 rs61999676 Trp->Gly	14:106926388 rs2467910 Thr->Ala	14:106926435 rs182650845 Ser->Phe	14:106926454 rs188914311 Val->Met
ref	0.57745	-	-	C	T	C	A	A	T	G	C
a	0.2428	512	14	C	**C**	C	A	A	**C**	G	C
b	0.08365	164	13	C	T	C	A	**C**	**C**	G	C
c	0.0302	66	10	C	T	C	A	**C**	T	G	C
e	0.0182	40	11	C	T	C	A	A	**C**	G	C
d	0.01455	30	4	**T**	T	C	A	A	T	G	C
g	0.00995	22	7	C	T	C	A	**C**	T	G	C
h	0.0064	13	5	C	**C**	C	A	A	**C**	G	C
f	0.0055	12	5	C	T	C	A	A	T	■	C
i	0.00365	8	3	C	**C**	C	A	A	**C**	■	C
j	0.00225	5	4	C	T	C	A	**C**	**C**	G	C
k	0.00135	3	2	C	T	**T**	A	A	T	G	C
m	0.00135	3	3	C	**C**	C	A	A	T	G	C
p	0.0009	2	2	C	**C**	C	A	A	**C**	G	**T**
q	0.00045	1	1	C	T	C	**T**	A	T	G	C
n	0.00045	1	1	C	T	C	A	A	**C**	■	C
l	0.00045	1	1	C	T	C	A	A	**C**	G	C
o	0.00045	1	1	C	**C**	C	A	A	**C**	G	C

Table 13.49 - IGHV3-43

h.type	cum.freq	# indvs	pops	14:106926456 rs79008247 Val->Gly	14:106926481 rs111739001 Val->Met	ENST00000434710
ref	0.57745	-	-	A	C	
a	0.2428	512	14	A	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.08365	164	13	**C**	C	1,2,3,4,5,6,7,9,10,11,12,13,14
c	0.0302	66	10	**C**	C	1,3,4,5,6,7,9,11,12,13
e	0.0182	40	11	A	C	1,3,4,5,6,8,10,11,12,13,14
d	0.01455	30	4	A	C	2,5,6,7
g	0.00995	22	7	A	C	2,3,4,6,9,10,11
h	0.0064	13	5	A	**T**	1,2,6,7,13
f	0.0055	12	5	A	C	1,7,10,12,14
i	0.00365	8	3	A	C	1,10,14
j	0.00225	5	4	A	C	3,4,9,10
k	0.00135	3	2	A	C	10,14
m	0.00135	3	3	A	C	2,7,13
p	0.0009	2	2	A	C	1,10
q	0.00045	1	1	A	C	13
n	0.00045	1	1	A	C	10
l	0.00045	1	1	**C**	C	10
o	0.00045	1	1	**C**	C	9

Table 13.50 - IGHV3-48

h.type	cum.freq	# indvs	pops	14:106993845 rs7148408 Asp->Ala	14:106993938 . ->Thr	14:106993945 rs7148607 Ser->Gly	14:106993996 rs183130922 Arg->Cys	14:106994231 rs192884486 Ile->Val	14:106994237 rs184894692 Val->Leu	14:106994262 rs188639638 Glu->Asp	ENST00000390624
ref	0.3919	-	-	T	GTAC	T	G	T	C	C	
a	0.3355	589	14	**G**	GTAC	T	G	T	C	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.20455	404	14	**G**	GTAC	**C**	G	T	C	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.027	59	11	**G**	**G**	T	G	T	C	C	1,2,4,5,6,7,10,11,12,13,14
d	0.01595	35	11	T	GTAC	**C**	G	T	C	C	1,2,3,4,5,6,7,8,9,10,13
f	0.0128	28	9	**G**	**G**	**C**	G	T	C	C	1,3,4,6,7,9,10,11,14
e	0.0105	23	10	T	**G**	T	G	T	C	C	2,3,4,5,6,7,10,11,13,14
j	0.00045	1	1	T	GTAC	T	**■**	T	C	C	11
g	0.00045	1	1	T	GTAC	T	G	T	**G**	C	5
h	0.00045	1	1	**G**	GTAC	T	G	**C**	C	C	2
i	0.00045	1	1	**G**	GTAC	T	G	T	C	**G**	1

Table 13.51 - IGHV3-49

h.type	cum.freq	# indvs	pops	14:107012994	14:107013065	14:107013072	14:107013093	14:107013129	14:107013146	14:107013171	14:107013201
				rs140678750	rs140027079	rs191590980	rs187070171	rs2073674	rs184357967	rs193139111	rs2073673
				Tyr->His	Gly->Glu	Tyr->Asp	Gly->Ser	Phe->Val	Asp->Ala	Thr->Ala	Gln->Lys
ref	0.322	-	-	A	C	A	C	A	T	T	G
a	0.38705	663	14	A	C	A	C	**C**	T	T	G
b	0.28555	499	13	A	C	A	C	A	T	T	**T**
c	0.0027	6	3	**G**	C	A	C	A	T	T	G
d	0.00045	1	1	A	C	**C**	**T**	**C**	**G**	T	G
g	0.00045	1	1	A	C	A	C	A	T	**C**	G
e	0.00045	1	1	**G**	C	A	C	**C**	T	T	G
h	0.00045	1	1	A	C	A	C	**C**	T	**C**	G
f	0.00045	1	1	A	C	A	C	**C**	T	T	G
i	0.00045	1	1	A	**T**	A	C	A	T	T	G

Table 13.51 - IGHV3-49

h.type	cum.freq	# indvs	pops	14:107013209 rs184512071 Gly->Asp	14:107013210 rs189168917 Gly->Ser	ENST00000390625
ref	0.322	-	-	C	C	
a	0.38705	663	14	C	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.28555	499	13	C	C	1,2,3,4,5,6,7,8,9,10,11,12,13
c	0.0027	6	3	C	C	7,11,13
d	0.00045	1	1	C	C	2
g	0.00045	1	1	C	C	1
e	0.00045	1	1	C	C	4
h	0.00045	1	1	C	C	1
f	0.00045	1	1	T	T	2
i	0.00045	1	1	C	C	10

Table 13.52 - IGHV3-53

h.type	cum.freq	# indvs	pops	14:107048717	14:107048749	14:107048795	14:107048796	14:107048929	14:107048944	14:107049080	
				rs141949273	rs11627238	.	rs182065680	rs35608607	rs35850880	rs2731152	ENST00000390627
				Met->Ile	Asp->His	Ser->	Ser->Asn	Ile->Val	Ser->Thr	Leu->Ser	
ref	0.4053	-	-	C	C	G	C	T	A	A	
a	0.2535	483	14	C	**G**	G	C	**C**	A	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.2264	399	13	C	C	G	C	T	**T**	**G**	1,2,3,4,5,6,7,8,10,11,12,13,14
c	0.0453	99	13	C	C	G	C	**C**	A	A	1,2,3,4,5,6,7,9,10,11,12,13,14
d	0.0247	54	6	**T**	C	G	C	T	A	A	1,5,10,12,13,14
e	0.015	33	9	C	C	G	C	**C**	**T**	**G**	1,2,5,7,8,10,11,12,14
f	0.0092	20	6	C	C	-	C	T	A	A	3,4,9,10,11,12
g	0.0046	10	5	C	C	G	C	**C**	A	**G**	7,10,11,13,14
k	0.0041	9	8	C	**G**	G	C	T	A	A	3,6,7,9,11,12,13,14
h	0.0036	8	2	**T**	C	G	C	**C**	A	A	10,14
j	0.0023	5	4	C	**G**	-	C	**C**	A	A	1,3,4,9
i	0.0018	4	4	C	C	G	C	T	A	**G**	8,11,13,14
n	0.0014	3	2	C	C	-	C	T	**T**	**G**	2,10
m	0.0014	3	3	C	C	-	C	**C**	A	A	5,9,10
q	0.0005	1	1	C	C	G	C	**C**	**T**	A	14
l	0.0005	1	1	C	**G**	G	**T**	**C**	A	A	3
p	0.0005	1	1	C	**G**	-	C	T	A	A	4
o	0.0005	1	1	**T**	C	G	C	**C**	**T**	**G**	14

Table 13.53 - IGHV3-64

h.type	cum.freq	# indvs	pops	14:107113770	14:107113851	14:107113877	14:107113902	14:107114009	14:107114176	14:107114189	14:107114193
				rs187347757	rs2072045	rs190575084	rs138428007	rs11846079	rs182438044	rs139161493	rs2073670
				Glu->Gln	Asp->Asn	Ser->Asn	Leu->Val	Glu->Gly	Trp->Gly	Glu->Asp	Met->Thr
ref	0.4221	-	-	C	C	C	G	T	A	C	A
a	0.47965	793	14	C	**T**	C	G	**C**	A	C	**G**
b	0.06225	131	13	C	C	C	G	**C**	A	C	A
c	0.02785	61	14	C	C	C	G	**C**	A	C	**G**
d	0.00275	6	4	C	**T**	C	G	T	A	C	**G**
e	0.0018	4	2	C	**T**	C	G	**C**	**C**	C	**G**
f	0.00135	3	1	C	**T**	C	**C**	**C**	A	C	**G**
h	0.0009	2	1	C	**T**	C	G	**C**	A	C	A
j	0.00045	1	1	C	C	C	G	T	A	**G**	A
g	0.00045	1	1	**G**	C	C	G	T	A	C	A
i	0.00045	1	1	C	**T**	**T**	G	T	A	C	A

Table 13.53 - IGHV3-64

h.type	cum.freq	# indvs	pops	ENST00000454421
ref	0.4221	-	-	
a	0.47965	793	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.06225	131	13	1,2,3,4,5,6,7,9,10,11,12,13,14
c	0.02785	61	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.00275	6	4	2,4,13,14
e	0.0018	4	2	1,10
f	0.00135	3	1	10
h	0.0009	2	1	12
j	0.00045	1	1	2
g	0.00045	1	1	3
i	0.00045	1	1	8

Table 13.54 - IGHV3-66

h.type	cum.freq	# indvs	pops	14:107131089	rs77593173	14:107131164	rs6423677	14:107131223	rs183864069	14:107131290	rs149638514	ENST00000390632
				Tyr->His		Cys->Gly		Met->Lys		Ile->Val		
ref	0.43755	-	-	A		A		A		T		
a	0.27865	502	14	A		C		A		C		1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.12575	252	14	A		A		A		C		1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0649	140	13	G		A		A		C		1,2,3,4,5,6,7,9,10,11,12,13,14
d	0.03975	87	13	G		C		A		C		1,2,3,4,5,6,7,9,10,11,12,13,14
e	0.02745	60	14	G		A		A		T		1,2,3,4,5,6,7,8,9,10,11,12,13,14
f	0.021	46	14	A		C		A		T		1,2,3,4,5,6,7,8,9,10,11,12,13,14
g	0.0027	6	3	G		C		A		T		10,13,14
h	0.0009	2	2	G		C		T		C		3,4
j	0.00045	1	1	A		A		T		T		5
k	0.00045	1	1	A		C		T		C		4
i	0.00045	1	1	A		A		T		C		3

Table 13.55 - IGHV3-72

h.type	cum.freq	# indvs	pops	14:107198966	14:107199062	14:107199136	14:107199164	14:107199222	14:107199355	
				rs193196315	rs184752334	rs139291045	rs191373077	rs188380722	rs192371094	
				Lys->Arg	Ser->Asn	His->Gln	Ala->Gly	Leu->Val	Ile->Leu	ENST00000433072
ref	0.9973	-	-	T	C	G	G	G	T	
a	0.00045	1	1	**C**	**T**	G	G	G	T	5
d	0.00045	1	1	T	C	**C**	**C**	G	T	2
e	0.00045	1	1	**C**	C	G	G	G	T	11
c	0.00045	1	1	T	C	G	G	**C**	T	2
b	0.00045	1	1	T	C	G	**C**	G	T	1
f	0.00045	1	1	T	C	G	G	G	**G**	10

Table 13.56 - IGHV3-73

h.type	cum.freq	# indvs	dops	14:107211000 rs142927366 Lys->Glu	14:107211135 rs184224767 Ala->Pro	14:107211176 rs61750833 Lys->Arg	14:107211199 rs188470998 Leu->Phe	14:107211371 rs61752554 Trp->Cys	ENST00000390636
ref	0.91625	-	-	T	C	T	C	C	
a	0.05265	96	6	T	C	T	C	**G**	1,5,10,12,13,14
b	0.0293	60	5	T	C	**C**	C	C	1,7,9,10,14
c	0.00135	3	3	T	C	T	**■**	C	1,10,14
d	0.00045	1	1	**C**	**G**	T	C	C	2

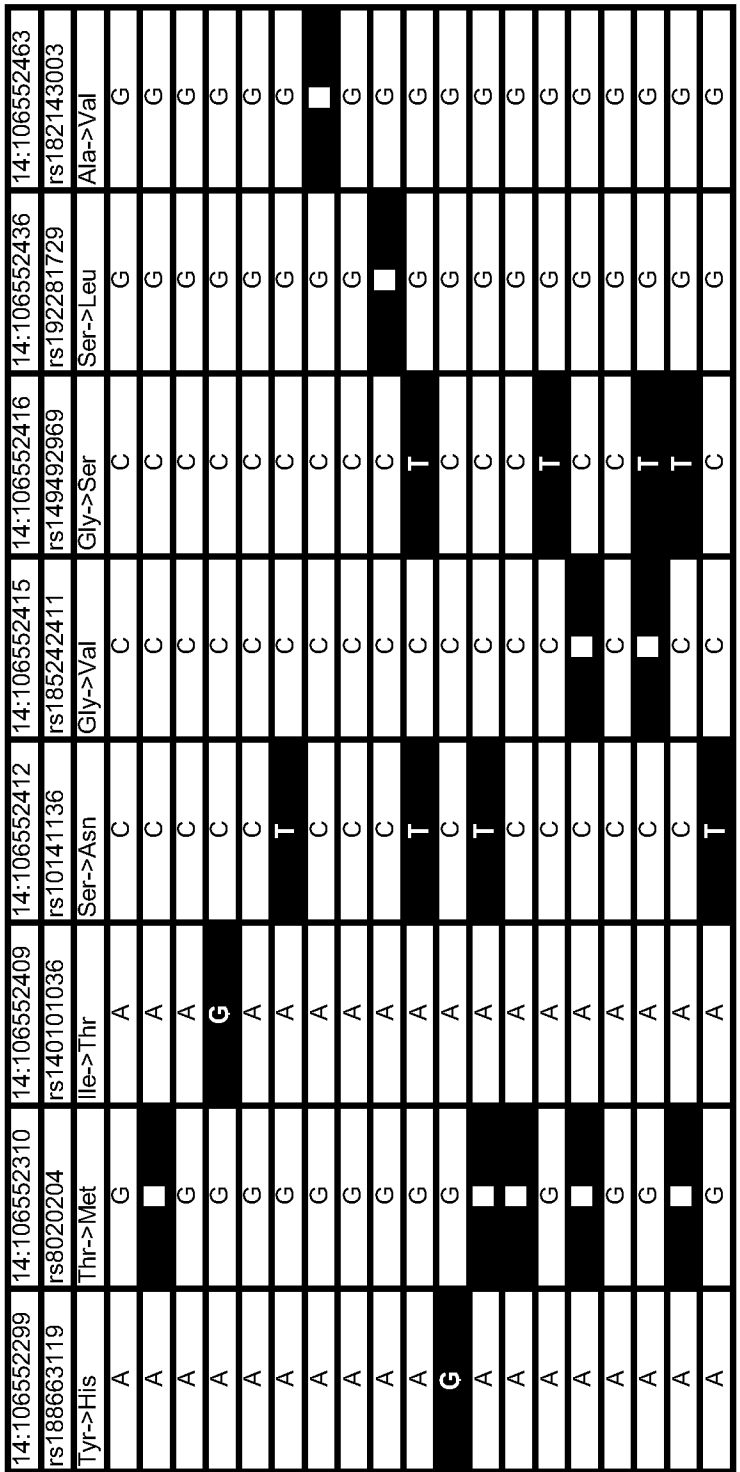
Table 13.57 - IGHV3-74

h.type	cum.freq	# indvs	pops	14:107218737 rs184901617 Thr->Met	14:107218809 rs183213717 Asp->Gly	14:107218812 rs140462183 Ser->Thr	14:107218858 rs191764284 Arg->Cys	14:107218878 rs183293744 Ser->Asn	14:107218908 rs149597490 Ser->Cys	ENST00000424969
ref	0.98635	-	-	G	T	C	G	C	G	
a	0.0091	20	6	G	T	**G**	G	C	G	1,9,10,12,13,14
b	0.0023	5	4	■	T	C	G	C	G	2,3,5,11
c	0.0009	2	2	G	T	C	G	**T**	G	10,14
d	0.00045	1	1	G	T	C	■	C	G	1
e	0.00045	1	1	G	**C**	C	G	C	G	11
f	0.00045	1	1	G	T	C	G	C	**C**	10

Table 13.58 - IGHV3-7

h.type	cum.fre	# indvs	pops	14:106518545	14:106518602	14:106518686	14:106518840	
				rs185279906	rs185538237	rs180765266	rs190666315	
				Asn->Ser	Ser->Asn	Gln->Arg	Leu->Arg	ENST00000390598
ref	0.9808	-	-	T	C	T	A	
a	0.016	29	5	T	C	**C**	A	1,5,10,11,14
b	0.0023	5	4	T	**T**	T	A	5,9,13,14
d	0.00045	1	1	**C**	C	T	A	4
c	0.00045	1	1	T	C	T	**C**	13

Table 13.59 - IGHV3-9



h.type	cum.fr	# indvs	pops	14:106552299	rs188663119	14:106552310	rs8020204	14:106552409	rs140101036	14:106552412	rs10141136	14:106552415	rs185242411	14:106552416	rs149492969	14:106552436	rs192281729	14:106552463	rs182143003
					Tyr->His		Thr->Met		Ile->Thr		Ser->Asn		Gly->Val		Gly->Ser		Ser->Leu		Ala->Val
ref	0.84305	-	-	A		G		A		C		C		C		G		G	
a	0.10975	167	9	A		G		A		C		C		C		G		G	
b	0.0228	46	7	A		G		A		C		C		C		G		G	
c	0.0064	12	5	A		G		G		C		C		C		G		G	
e	0.00405	6	4	A		G		A		C		C		C		G		G	
d	0.00315	7	7	A		G		A		T		C		C		G		G	
f	0.00225	5	4	A		G		A		C		C		C		G		■	
g	0.0018	4	3	A		G		A		C		C		C		G		G	
h	0.00135	2	1	A		G		A		C		C		C		■		G	
r	0.0009	1	1	A		G		A		T		C		T		G		G	
o	0.0009	2	1	G		G		A		C		C		C		G		G	
j	0.00045	1	1	A		■		A		T		C		C		G		G	
k	0.00045	1	1	A		■		A		C		C		C		G		G	
q	0.00045	1	1	A		G		A		C		C		T		G		G	
i	0.00045	1	1	A		■		A		C		■		C		G		G	
n	0.00045	1	1	A		G		A		C		C		C		G		G	
m	0.00045	1	1	A		G		A		C		■		T		G		G	
l	0.00045	1	1	A		■		A		C		C		T		G		G	
p	0.00045	1	1	A		G		A		T		C		C		G		G	

Table 13.59 - IGHV3-9

h.type	cum.freq	# indvs	pops	14:106552464 rs186818749 Ala->Pro	14:106552472 rs191640600 Val->Ala	14:106552491 rs182569652 Asp->Asn	14:106552493 rs141844186 Asp->Gly	14:106552713 rs116484141 Ser->Asn	ENST00000390600
ref	0.84305	-	-	C	A	C	T	C	
a	0.10975	167	9	C	A	C	T	C	1,5,7,8,10,11,12,13,14
b	0.0228	46	7	C	A	C	**C**	C	3,4,6,7,9,11,13
c	0.0064	12	5	C	A	C	T	C	1,7,10,12,14
e	0.00405	6	4	C	A	**T**	T	C	1,3,12,14
d	0.00315	7	7	C	A	C	T	C	1,2,3,5,6,8,10
f	0.00225	5	4	C	A	C	T	C	6,7,9,13
g	0.0018	4	3	**G**	A	C	T	C	2,10,14
h	0.00135	2	1	C	A	C	T	C	1
r	0.0009	1	1	C	A	C	T	C	2
o	0.0009	2	1	C	A	C	T	C	13
j	0.00045	1	1	C	A	C	T	C	10
k	0.00045	1	1	**G**	A	C	T	C	10
q	0.00045	1	1	C	A	C	T	**T**	2
i	0.00045	1	1	C	A	C	T	C	14
n	0.00045	1	1	C	**G**	C	T	C	6
m	0.00045	1	1	C	A	C	T	C	11
l	0.00045	1	1	C	A	C	T	C	14
p	0.00045	1	1	C	A	C	**C**	C	3

Table 13.60 - IGHV4-28

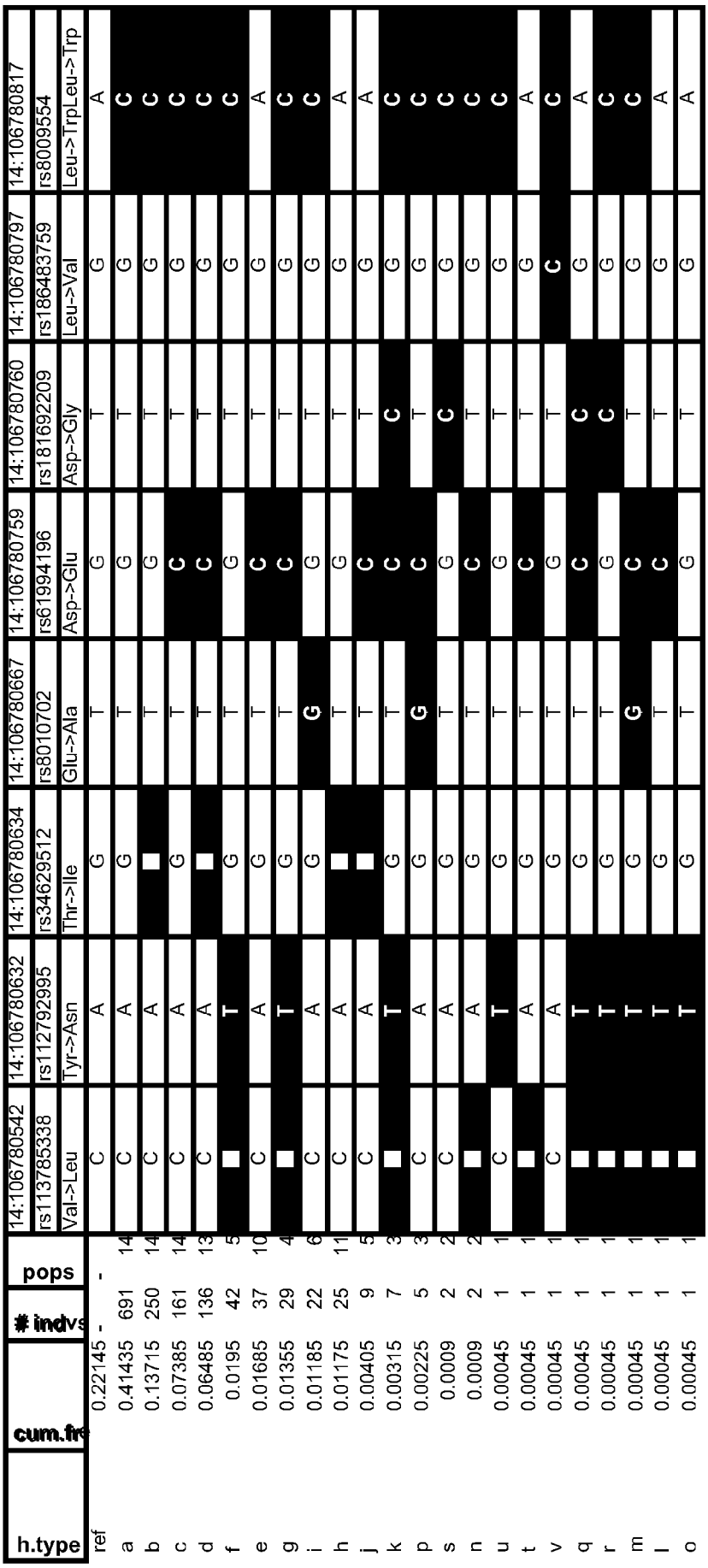


Table 13.60 - IGHV4-28

h.type	cum.fre	# indvs	pops	ENST00000390612
ref	0.22145	-	-	
a	0.41435	691	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.13715	250	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.07385	161	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.06485	136	13	1,2,3,5,6,7,8,9,10,11,12,13,14
f	0.0195	42	5	1,10,11,12,14
e	0.01685	37	10	1,2,5,6,7,10,11,12,13,14
g	0.01355	29	4	1,5,10,14
i	0.01185	22	6	1,8,10,11,13,14
h	0.01175	25	11	1,2,3,5,7,8,9,10,11,13,14
j	0.00405	9	5	2,6,7,10,13
k	0.00315	7	3	1,10,14
p	0.00225	5	3	1,10,14
s	0.0009	2	2	1,11
n	0.0009	2	2	10,14
u	0.00045	1	1	10
t	0.00045	1	1	11
v	0.00045	1	1	13
q	0.00045	1	1	11
r	0.00045	1	1	1
m	0.00045	1	1	10
l	0.00045	1	1	10
o	0.00045	1	1	14

Table 13.61 - IGHV4-31

h.type	cum.fr	# indvs	pops	14:106805289	14:106805303	14:106805321	14:106805345	14:106805358	14:106805381	14:106805388	14:106805408
				rs188101179	rs145562667	rs192498675	rs185638910	rs190512954	rs77489245	rs149858616	rs61995642
				Val->Leu	Arg->Leu	Asn->Ser	Tyr->Ser	Ile->Leu	His->Pro	Arg->Cys	Gly->Asp
ref	0.01175	-	-	C	C	T	T	T	T	G	C
a	0.5475	886	14	C	C	T	T	T	T	G	C
b	0.2035	420	14	C	C	T	T	T	G	G	T
c	0.1554	303	14	C	C	T	T	T	G	G	C
d	0.04565	83	11	C	■	T	T	T	T	G	C
e	0.00865	18	11	C	C	T	T	T	T	G	T
f	0.00865	19	7	C	C	T	T	T	G	G	T
i	0.0054	10	3	C	C	T	T	T	T	■	C
g	0.00315	7	4	C	■	T	T	T	T	G	C
j	0.00225	5	3	C	C	T	T	T	G	G	C
h	0.00135	3	2	C	C	T	T	T	T	■	C
t	0.0009	2	2	C	C	T	G	T	T	G	C
m	0.0009	2	2	C	■	T	T	T	G	G	C
l	0.0009	2	2	C	■	T	T	T	G	G	T
u	0.00045	1	1	C	C	T	T	G	T	G	C
k	0.00045	1	1	C	■	C	T	T	T	G	C
v	0.00045	1	1	C	C	T	G	T	G	G	T
s	0.00045	1	1	C	C	T	T	T	G	G	T
q	0.00045	1	1	G	■	T	T	T	T	G	C
r	0.00045	1	1	C	C	T	T	T	G	G	C
n	0.00045	1	1	C	■	T	G	T	T	G	C
p	0.00045	1	1	C	C	C	T	T	G	G	C
o	0.00045	1	1	G	C	T	T	T	T	G	C

Table 13.61 - IGHV4-31

h.type	cum.fre	# indvs	pops	14:106805411	14:106805421	14:106805497	14:106805508	
				rs185736587	rs188982921	rs181045179	rs4462488	
				Gly->Asp	Ile->Phe	Gln->His	Pro->Ser	ENST00000438142
ref	0.01175	-	-	C	T	C	G	
a	0.5475	886	14	C	T	C		1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.2035	420	14	C	T	C		1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.1554	303	14	C	T	C		1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.04565	83	11	C	T	C		1,2,5,6,7,8,10,11,12,13,14
e	0.00865	18	11	C	T	C		1,2,3,4,6,9,10,11,12,13,14
f	0.00865	19	7	C	T	C	G	1,3,6,10,11,12,14
i	0.0054	10	3	**T**	■	C	■	1,10,14
g	0.00315	7	4	C	T	C	G	1,2,5,7
j	0.00225	5	3	C	T	C	G	2,10,14
h	0.00135	3	2	C	T	C	■	1,10
t	0.0009	2	2	C	T	C	■	4,14
m	0.0009	2	2	C	T	C	■	1,10
l	0.0009	2	2	C	T	C	■	1,10
u	0.00045	1	1	C	T	C	■	4
k	0.00045	1	1	C	T	C	■	13
v	0.00045	1	1	C	T	C	■	1
s	0.00045	1	1	C	T	**G**	■	5
q	0.00045	1	1	C	T	C	■	5
r	0.00045	1	1	C	T	**G**	■	2
n	0.00045	1	1	C	T	C	■	7
p	0.00045	1	1	C	T	C	■	10
o	0.00045	1	1	C	T	C	■	5

Table 13.62 - IGHV4-34

h.type	cum.freq	# indvs	pops	14:106829649 rs1832223225 Ser->Phe	14:106829717 rs11546808 Ser->Arg	14:106829767 rs188280613 Pro->Ser	14:106829775 rs191592421 Ile->Ser	14:106829793 rs11546811 Gly->Ala	14:106829817 rs1064183 Ala->Asp	14:106829874 rs181039074 Leu->Pro	ENST00000390616
ref	0.9955	-	-	G	G	G	A	C	G	A	
d	0.0009	2	2	G	G	G	A	**G**	G	A	2,14
g	0.0009	2	2	G	**T**	G	A	C	G	A	11,14
b	0.0009	2	2	G	G	G	A	C	G	**G**	3,10
a	0.00045	1	1	G	G	G	**C**	C	G	A	6
e	0.00045	1	1	G	G	■	A	C	G	A	5
c	0.00045	1	1	■	G	G	A	C	G	A	7
f	0.00045	1	1	G	G	G	A	C	**T**	A	2

Table 13.63 - IGHV4-39

h.type	cum.freq	# indvs	pops	14:106877626	14:106877648	14:106877705	14:106877730	14:106877753	14:106877755	14:106877761	14:106877831
				rs188639181	rs191739990	rs187288552	rs191860623	rs35281264	rs186830319	rs115523671	rs190205925
				Cys->Tyr	Ala->Thr	Ile->Leu	Asn->Lys	Tyr->His	Tyr->Ser	Ser->Asn	Ile->Phe
ref	0.6445	-	-	C	C	T	G	A	T	C	T
a	0.10295	220	14	C	C	T	G	**G**	T	C	T
b	0.0918	197	14	C	C	T	G	**G**	T	C	T
d	0.04425	93	13	C	C	T	G	A	T	C	T
c	0.02555	56	10	C	C	T	G	**G**	T	C	T
e	0.01595	35	12	C	C	T	G	A	T	**T**	T
f	0.0132	29	10	C	C	T	G	**G**	T	C	T
g	0.0105	23	9	C	C	T	G	**G**	T	C	T
h	0.00595	13	8	C	C	T	G	**G**	T	C	T
i	0.0059	13	7	C	C	T	G	A	T	C	T
j	0.0055	12	8	C	C	T	G	A	T	C	T
m	0.00415	9	5	C	C	T	G	**G**	T	C	T
k	0.00365	8	5	C	C	T	G	**G**	T	C	T
l	0.00275	6	6	C	C	T	G	**G**	T	**T**	T
y	0.0018	4	3	C	C	T	G	A	T	C	T
n	0.0018	4	3	C	C	T	G	**G**	T	C	T
o	0.0018	4	3	C	C	T	G	**G**	T	C	T
u	0.00135	3	3	C	C	T	G	A	T	C	T
t	0.00135	3	2	C	C	T	G	**G**	T	C	T
q	0.00135	3	3	C	C	T	G	A	T	C	T
aa	0.0009	2	2	C	C	T	G	**G**	T	C	T
s	0.0009	2	2	C	C	T	G	**G**	T	C	T
x	0.0009	2	2	C	C	T	G	**G**	T	C	T
p	0.0009	2	2	C	C	T	G	A	T	**T**	T
ah	0.00045	1	1	C	C	T	G	A	T	C	T
ag	0.00045	1	1	C	C	T	G	A	T	**T**	T
aj	0.00045	1	1	C	C	T	G	**G**	T	C	T
aq	0.00045	1	1	C	**T**	T	G	A	T	C	T
v	0.00045	1	1	C	C	T	G	**G**	T	C	T
al	0.00045	1	1	C	C	T	**C**	A	T	C	T

Table 13.63 - IGHV4-39

ab	0.00045	1	1	C	C	T	C	A	G	C	T
ak	0.00045	1	1	C	C	T	G	G	T	T	T
ap	0.00045	1	1	C	C	T	G	A	T	T	T
z	0.00045	1	1	C	C	T	G	A	T	T	T
w	0.00045	1	1	C	C	T	G	G	T	C	T
ao	0.00045	1	1	C	C	T	G	A	T	C	T
r	0.00045	1	1	C	C	T	G	A	T	C	T
af	0.00045	1	1	T	C	T	G	A	T	C	T
ad	0.00045	1	1	C	C	T	G	A	T	C	T
am	0.00045	1	1	C	C	T	G	A	T	C	T
at	0.00045	1	1	C	C	T	G	A	G	C	□
ac	0.00045	1	1	C	C	T	G	A	T	C	T
ar	0.00045	1	1	C	T	T	G	A	T	C	T
as	0.00045	1	1	C	C	G	G	A	T	T	T
ae	0.00045	1	1	C	C	T	G	A	T	C	T
ai	0.00045	1	1	C	C	T	G	A	T	C	T
an	0.00045	1	1	C	C	T	G	A	T	C	T

Table 13.63 - IGHV4-39

h.type	cum.freq	# indvs	pops	14:106877836 rs143740707 Gly->Asp	14:106877837 rs145759866 Gly->Cys	14:106877848 rs181859654 Thr->Ser	14:106877877 rs181228803 Lys->Asn	14:106877888 rs190955722 Gly->Arg	14:106877907 rs144078801 Gln->His	14:106877912 rs74093494 Leu->Val	14:106878019 rs140461651 Ala->Thr
ref	0.6445	-	-	C	C	G	C	C	C	G	C
a	0.10295	220	14	C	C	G	C	C	C	**C**	C
b	0.0918	197	14	**T**	**C**	G	C	C	C	**C**	C
d	0.04425	93	13	C	C	G	C	C	C	G	C
c	0.02555	56	10	C	C	G	C	C	C	**C**	C
e	0.01595	35	12	C	C	G	C	C	C	G	C
f	0.0132	29	10	C	C	G	C	C	C	G	C
g	0.0105	23	9	C	C	G	C	C	C	G	C
h	0.00595	13	8	C	C	G	C	C	C	**C**	C
i	0.0059	13	7	C	C	G	C	C	C	**C**	C
j	0.0055	12	8	C	C	G	C	C	C	**C**	C
m	0.00415	9	5	C	C	G	C	C	C	**C**	C
k	0.00365	8	5	**T**	**C**	G	C	C	C	G	C
l	0.00275	6	6	C	C	G	C	C	C	**C**	C
y	0.0018	4	3	**T**	C	G	C	C	C	G	C
n	0.0018	4	3	C	C	G	C	C	C	G	C
o	0.0018	4	3	**T**	**C**	G	C	C	C	**C**	C
u	0.00135	3	3	C	C	G	C	C	C	G	C
t	0.00135	3	2	**T**	**C**	G	C	C	C	G	C
q	0.00135	3	3	C	C	G	C	C	C	G	C
aa	0.0009	2	2	C	C	G	C	C	C	G	C
s	0.0009	2	2	C	**C**	G	C	C	C	**C**	C
x	0.0009	2	2	**T**	C	G	C	C	C	**C**	C
p	0.0009	2	2	C	C	G	C	C	C	**C**	C
ah	0.00045	1	1	C	C	G	C	C	C	G	C
ag	0.00045	1	1	C	C	G	C	C	C	**C**	C
aj	0.00045	1	1	**T**	**C**	G	C	C	C	**C**	C
aq	0.00045	1	1	C	C	G	C	C	C	G	C
v	0.00045	1	1	**T**	C	G	C	C	C	G	C
al	0.00045	1	1	C	C	G	C	**T**	C	G	C

Table 13.63 - IGHV4-39

ab	0.00045	1	1	C	C	G	C	C	C	G	C
ak	0.00045	1	1	C	C	G	C	C	C	G	C
ap	0.00045	1	1	C	C	G	G	C	C	G	C
z	0.00045	1	1	C	C	G	C	C	C	G	C
w	0.00045	1	1	T	C	G	C	C	C	C	C
ao	0.00045	1	1	T	C	G	C	C	C	C	C
r	0.00045	1	1	C	C	G	C	C	□	G	C
af	0.00045	1	1	C	C	G	C	C	C	G	C
ad	0.00045	1	1	T	□	G	C	C	C	C	C
am	0.00045	1	1	C	C	G	C	C	C	G	C
at	0.00045	1	1	C	C	C	C	C	C	G	C
ac	0.00045	1	1	C	C	C	C	C	C	G	C
ar	0.00045	1	1	C	C	G	C	C	C	G	C
as	0.00045	1	1	T	C	G	C	T	C	C	T
ae	0.00045	1	1	T	□	G	C	C	C	G	C
ai	0.00045	1	1	C	C	G	C	C	C	G	C
an	0.00045	1	1	T	□	G	C	C	C	C	C

Table 13.63 - IGHV4-39

h.type	cum.freq	# indvs	pops	14:106878048 rs187579450 His->Arg	14:106878052 rs138364008 Lys->Gln	14:106878056 rs4774113 Lys->Asn	14:106878057 rs190089302 Lys->Thr	14:106878067 rs4774114 Met->Leu	ENST00000390619
ref	0.6445	-	-	T	T	T	T	T	
a	0.10295	220	14	T	T	G	T	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0918	197	14	T	T	G	T	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.04425	93	13	T	T	G	T	G	1,2,3,4,5,6,7,9,10,11,12,13,14
c	0.02555	56	10	T	T	T	T	T	1,2,3,5,6,7,10,11,13,14
e	0.01595	35	12	T	T	T	T	T	1,2,3,4,5,6,9,10,11,12,13,14
f	0.0132	29	10	T	T	T	T	T	1,2,3,4,5,6,7,10,11,14
g	0.0105	23	9	T	T	G	T	G	1,2,4,5,10,11,12,13,14
h	0.00595	13	8	T	T	T	T	G	1,2,6,7,8,10,13,14
i	0.0059	13	7	T	T	G	T	G	1,2,4,8,10,13,14
j	0.0055	12	8	T	T	T	T	T	1,2,3,4,5,9,11,12
m	0.00415	9	5	T	T	G	T	T	2,3,7,10,14
k	0.00365	8	5	T	T	G	T	G	2,3,4,9,13
l	0.00275	6	6	T	T	G	T	G	1,2,5,7,12,13
y	0.0018	4	3	T	T	T	T	T	2,5,13
n	0.0018	4	3	T	T	G	T	T	3,13,14
o	0.0018	4	3	T	T	T	T	T	4,6,13
u	0.00135	3	3	T	T	T	T	G	1,4,14
t	0.00135	3	2	T	T	T	T	T	3,4
q	0.00135	3	3	T	T	G	T	T	2,3,6
aa	0.0009	2	2	T	T	T	T	G	1,6
s	0.0009	2	2	T	T	G	T	G	7,8
x	0.0009	2	2	T	T	G	T	G	3,4
p	0.0009	2	2	T	T	G	T	G	5,14
ah	0.00045	1	1	C	T	T	T	T	13
ag	0.00045	1	1	T	T	T	T	T	11
aj	0.00045	1	1	T	T	T	T	G	3
aq	0.00045	1	1	T	T	G	T	G	2
v	0.00045	1	1	T	T	T	T	T	7
al	0.00045	1	1	T	T	G	T	G	14

Table 13.63 - IGHV4-39

ab	0.00045	1	1	T	T	T	T	T	4
ak	0.00045	1	1	T	T	T	T	T	13
ap	0.00045	1	1	T	T	T	T	T	13
z	0.00045	1	1	T	T	G	T	T	10
w	0.00045	1	1	T	T	G	T	T	10
ao	0.00045	1	1	T	T	T	T	T	12
r	0.00045	1	1	T	T	T	T	T	3
af	0.00045	1	1	T	T	T	T	T	12
ad	0.00045	1	1	T	T	T	T	T	5
am	0.00045	1	1	T	T	T	G	T	13
at	0.00045	1	1	T	T	T	T	T	1
ac	0.00045	1	1	T	T	T	T	T	13
ar	0.00045	1	1	T	T	T	T	T	6
as	0.00045	1	1	T	T	G	T	G	2
ae	0.00045	1	1	T	T	G	T	G	9
ai	0.00045	1	1	T	G	T	T	T	13
an	0.00045	1	1	T	T	G	T	G	2

Table 13.64 - IGHV4-4

h.type	cum.freq	# indvs	dops	14:106478120 rs150123115 Tyr->Cys	14:106478183 rs77628366 Thr->Lys	14:106478194 rs76899215 Met->Ile	14:106478252 rs73365408 Arg->His	14:106478311 . Ser->	14:106478354 rs188216744 Glu->Gly	ENST00000390596
ref	0.34	-	-	T	G	C	C	A	T	
a	0.1136	222	14	T	G	C	C		**C**	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.0817	166	14	T	G	**T**	C		T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
f	0.06675	138	14	T	G	**T**	C		**C**	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.06385	140	13	T	G	**T**	C	A	T	1,2,3,4,5,6,7,9,10,11,12,13,14
c	0.0599	127	14	T	G	C	C		T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
e	0.05215	110	12	T	**T**	**T**	C		**C**	1,2,3,4,5,6,7,9,11,12,13,14
g	0.03255	65	10	**C**	**T**	**T**	C		**C**	1,2,3,4,5,9,10,11,12,14
m	0.0267	58	12	T	**T**	**T**	**T**		**C**	1,2,3,4,5,6,7,9,10,12,13,14
i	0.0262	58	14	T	**T**	**T**	C		T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
j	0.01715	38	11	T	**T**	**T**	C	A	T	1,2,4,5,6,7,9,10,11,12,14
h	0.01535	33	11	T	G	C	C	A	**C**	2,3,4,5,6,7,9,10,11,12,13
k	0.0113	25	10	**C**	G	C	C		**C**	2,3,4,5,6,8,9,11,12,13
o	0.00855	19	9	T	**T**	C	C		**C**	1,2,3,4,5,6,7,8,13
q	0.00765	17	7	**C**	G	C	C	A	T	1,3,9,10,11,12,14
n	0.00765	16	5	T	G	**T**	**T**	A	T	1,5,10,13,14
p	0.00765	16	5	T	G	**T**	**T**		T	1,5,9,10,14
l	0.00685	15	4	**C**	G	C	C	A	**C**	3,4,9,10
s	0.00585	13	4	T	G	C	**T**	A	T	1,10,12,14
aj	0.0054	12	8	T	G	**T**	C	A	**C**	2,3,4,5,9,12,13,14
ah	0.00495	11	7	T	**T**	**T**	**T**		T	1,2,3,9,10,13,14
t	0.0045	10	6	**C**	G	**T**	C		**C**	3,4,6,8,9,11
w	0.00405	9	7	**C**	**T**	**T**	C		T	3,4,7,9,11,13,14
r	0.0036	8	7	T	G	**T**	**T**		**C**	2,4,9,11,12,13,14
v	0.00315	7	5	**C**	G	C	C		T	7,9,11,13,14
ag	0.0027	6	5	T	**T**	**T**	C	A	**C**	2,5,6,10,14
y	0.0027	6	4	**C**	**T**	C	C		**C**	3,4,5,9
u	0.0027	6	3	T	**T**	**T**	**T**	A	T	1,10,14
z	0.0027	6	4	**C**	**T**	**T**	**T**		**C**	2,4,9,10

Table 13.64 - IGHV4-4

ak	0.0018	4	4	T	G	C	T		T	3,4,5,10
ao	0.00135	3	3	C	G	T	C		T	3,9,11
ac	0.00135	3	1	T	G	C	T		C	10
ae	0.00135	3	2	T	T	C	C	A	C	2,4
aa	0.0009	2	2	C	T	T	C	A	T	9,14
af	0.0009	2	2	T	T	C	T		C	6,9
x	0.0009	2	2	C	G	T	C	A	T	9,12
ad	0.0009	2	2	T	T	C	C		T	3,7
ai	0.0009	2	1	T	T	C	C	A	T	7
al	0.00045	1	1	C	G	C	T		C	4
ab	0.00045	1	1	C	T	C	T		C	10
am	0.00045	1	1	C	G	T	T	A	C	9
an	0.00045	1	1	C	G	C	T	A	C	1

Table 13.65 - IGHV4-59

h.type	cum.freq	# indvs	pops	14:107082722	rs1813881142	Ser->Leu	14:107083453	rs112607697	Tyr->His	14:107083457	rs1900037133	Ser->Arg	ENST00000390629	ENST00000455737	Ser->Arg
ref	0.95005	-	-	G			A			A			1,2,5,10,12,14		
a	0.04905	102	6	G			**G**			A			1		
c	0.00045	1	1	G			A			**C**			14		
b	0.00045	1	1	■			A			A					

Table 13.66 - IGHV4-61

h.type	cum.freq	# indvs	pops	14:107095319 rs111864170 Tyr->Ser	14:107095325 rs186969565 Gly->Asp	14:107095326 rs1064309 Gly->Ser	14:107095338 rs2072046 Val->Ile	ENST00000390630
ref	0.0109	-	-	T	C	C	C	
a	0.5701	806	14	T	C	T	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.18115	339	14	T	C	T	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.13535	275	13	G	C	C	T	1,2,3,4,5,6,7,9,10,11,12,13,14
f	0.05855	120	14	T	C	C	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
e	0.0191	42	12	G	C	C	C	2,3,4,5,6,7,9,10,11,12,13,14
d	0.0109	24	10	G	C	T	C	1,2,3,6,7,9,10,11,12,13
g	0.00675	15	4	T	T	C	T	1,9,10,14
h	0.00585	13	7	G	C	T	T	2,3,6,7,9,10,14
j	0.00045	1	1	T	T	C	C	7
k	0.00045	1	1	T	T	T	T	10
i	0.00045	1	1	G	T	C	T	14

Table 13.67 - IGHV5-51

h.type	cum.freq	# indvs	pops	14:107034771	14:107034793	14:107034796	14:107034813	14:107034850	14:107034851	14:107034868	14:107034910
				rs180672762	rs181977500	rs186695491	rs191700780	rs142572296	rs186314656	rs139665108	rs117410356
				Ser->Arg	Ser->Asn	Ile->Ser	Ile->Met	Thr->Ser	Thr->Ala	Tyr->Phe	Arg->His
ref	0.97045	-	-	G	C	A	G	G	T	T	C
a	0.0164	35	3	G	C	A	G	G	T	T	**T**
b	0.0055	11	3	G	C	A	G	G	T	T	C
g	0.00135	3	3	G	**T**	A	G	G	T	T	C
d	0.0009	2	2	G	C	A	G	**C**	T	T	C
k	0.0009	2	2	G	C	A	G	G	T	T	C
c	0.0009	2	2	G	**T**	A	G	G	T	T	C
j	0.00045	1	1	G	C	A	G	G	T	**-**	C
e	0.00045	1	1	G	C	A	**C**	G	T	T	C
h	0.00045	1	1	G	C	A	G	G	T	T	C
f	0.00045	1	1	G	C	A	G	G	T	T	C
i	0.00045	1	1	G	C	A	G	G	T	T	C
n	0.00045	1	1	G	C	**C**	G	G	T	T	C
m	0.00045	1	1	**T**	C	A	**C**	G	**C**	T	C
l	0.00045	1	1	G	C	A	G	G	T	T	C

Table 13.67 - IGHV5-51

h.type	cum.freq	# indvs	pops	14:107034931 rs113154616 Ser->Thr	14:107034940 rs113988349 Ser->Asn	14:107034998 rs183518874 Ala->Pro	14:107035123 rs116300992 Leu->Ile	14:107035141 rs147695427 Ala->Thr	ENST00000390626
ref	0.97045	-	-	C	C	C	G	C	
a	0.0164	35	3	C	C	C	G	C	3,4,9
b	0.0055	11	3	C	C	C	**T**	C	5,10,14
g	0.00135	3	3	C	C	C	G	C	6,7,13
d	0.0009	2	2	C	C	C	G	C	2,13
k	0.0009	2	2	**G**	C	C	G	C	6,10
c	0.0009	2	2	C	**T**	C	G	C	2,10
j	0.00045	1	1	C	C	C	G	C	13
e	0.00045	1	1	C	C	C	G	C	10
h	0.00045	1	1	C	C	C	G	**T**	7
f	0.00045	1	1	**G**	**T**	C	G	C	10
i	0.00045	1	1	C	**T**	C	G	C	10
n	0.00045	1	1	C	C	C	G	C	4
m	0.00045	1	1	**G**	C	C	G	C	4
l	0.00045	1	1	C	C	**G**	G	C	6

Table 13.68 - IGHV6-1

h.type	cum.freq	# indvs	pops	14:106405816 rs186349163 Ser->Asn	14:106405877 rs72715486 Lys->Glu	14:106405910 rs188294175 Val->Ile	ENST00000390593
ref	0.99775	-	-	C	T	C	
a	0.0009	2	2	C	**C**	C	3,8
c	0.0009	2	1	C	T	**T**	4
b	0.00045	1	1	**T**	T	C	8

Table 13.69 - IGHV7-81

h.type	cum.freq	# indvs	pops	14:107282845	14:107282853	14:107282887	14:107283009	14:107283018	14:107283073	
				rs61741324	rs61741323	rs61741308	rs11845591	rs144292230	rs148758470	
				Leu->Met	Thr->Ile	Arg->Trp	Gly->Val	Lys->Met	Val->Leu	ENST00000390639
ref	0.9127	-	-	G	G	G	C	T	C	
a	0.0732	136	6	T	■	G	C	T	C	1,5,10,12,13,14
b	0.01005	21	3	G	G	G	■	T	C	1,10,14
c	0.0018	4	3	G	G	■	C	T	C	1,5,14
e	0.00135	3	1	G	G	G	C	■	C	10
d	0.0009	2	2	G	G	G	C	T	■	1,10

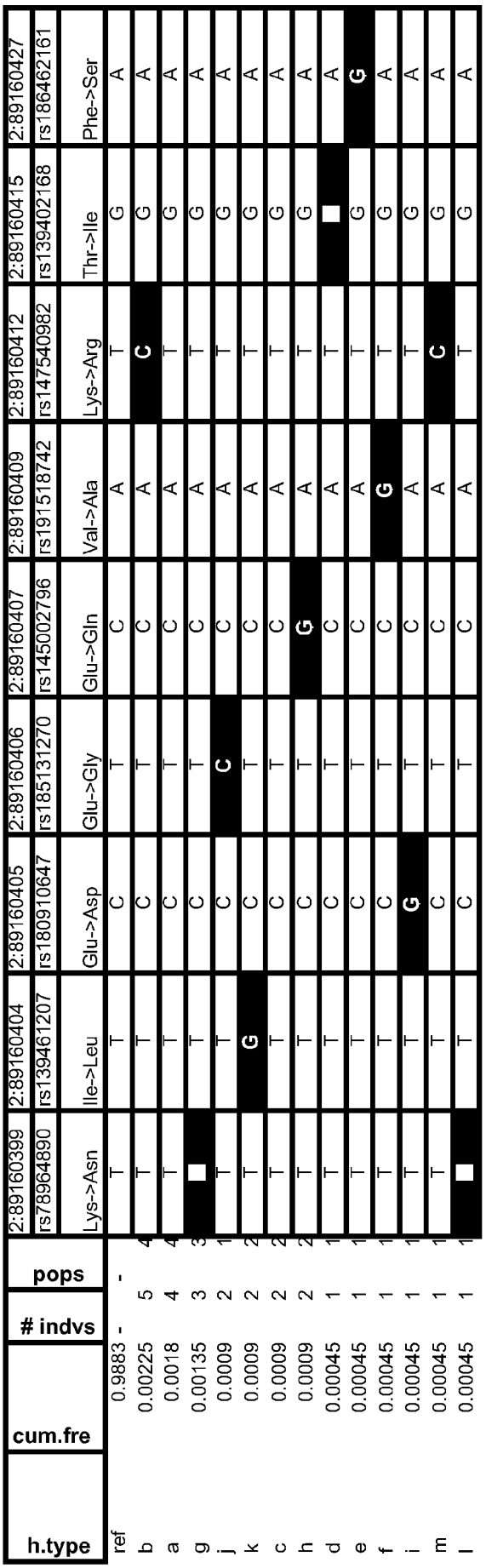
Table 13.70 - IGKJ2

h.type	cum.freq	# indvs	pops	2:89161040 rs76129343 Lys->Glu	2:89161044 rs75187701 Glu->Asp	2:89161045 rs146019185 Glu->Gly	2:89161050 rs181828141 Lys->Asn	2:89161055 rs183845604 Thr->Pro	2:89161067 rs150749817 Phe->Val	2:89161069 rs113458930 Thr->Ser	2:89161072 rs113539064 Tyr->Cys	ENST00000390241
ref	0.921	-	-	T	C	T	C	T	A	G	T	
a	0.0449	86	13	T	C	T	C	T	A	**C**	T	1,2,3,4,5,6,7,9,10,11,12,13,14
b	0.0243	48	10	T	C	T	C	T	A	**C**	**C**	2,4,5,6,7,8,10,11,12,13
d	0.0032	6	3	T	C	T	C	T	A	G	**C**	2,6,7
e	0.0018	4	3	T	C	**C**	C	T	A	G	T	2,10,13
c	0.0018	4	4	T	C	T	**G**	T	A	G	T	1,2,7,9
f	0.0014	3	2	T	C	T	C	T	**C**	G	T	2,10
g	0.0009	2	2	T	■	T	C	T	A	G	T	11,14
h	0.0005	1	1	**C**	C	T	C	T	**C**	G	**C**	9
i	0.0005	1	1	T	C	T	C	**G**	A	G	T	10

Table 13.71 - IGKJ3

h.type	cum.freq	# indvs	pops	2:89160742 rs188533155 Asp->His	2:89160748 rs181878952 Lys->Glu	2:89160762 rs190475170 Phe->Cys	2:89160769 rs139877244 Phe->Leu	ENST00000390240
ref	0.99045	-	-	C	T	A	A	
a	0.00685	15	10	C	T	A	**G**	2,4,5,6,9,10,11,12,13,14
d	0.00135	3	3	C	**C**	A	A	5,11,12
b	0.0009	2	2	**G**	T	A	A	2,3
c	0.00045	1	1	C	T	**C**	A	5

Table 13.72 - IGKJ4



h.type	cum.fre	# indvs	pops	2:89160399	2:89160404	2:89160405	2:89160406	2:89160407	2:89160409	2:89160412	2:89160415	2:89160427
				rs78964890	rs1394612 07	rs1809106 47	rs1851312 70	rs1450027 96	rs1915187 42	rs1475409 82	rs1394021 68	rs1864621 61
				Lys->Asn	Ile->Leu	Glu->Asp	Glu->Gly	Glu->Gln	Val->Ala	Lys->Arg	Thr->Ile	Phe->Ser
ref	0.9883	-	-	T	T	C	T	C	A	T	G	A
b	0.00225	5	4	T	T	C	T	C	A	C	G	A
a	0.0018	4	4	T	T	C	T	C	A	T	G	A
g	0.00135	3	3	■	T	C	T	C	A	T	G	A
j	0.0009	2	1	T	T	C	C	C	A	T	G	A
k	0.0009	2	2	T	G	C	T	C	A	T	G	A
c	0.0009	2	2	T	T	C	T	C	A	T	G	A
h	0.0009	2	2	T	T	C	T	G	A	T	G	A
d	0.00045	1	1	T	T	C	T	C	A	T	■	A
e	0.00045	1	1	T	T	C	T	C	A	T	G	G
f	0.00045	1	1	T	T	C	T	C	G	T	G	A
i	0.00045	1	1	T	T	G	T	C	A	T	G	A
m	0.00045	1	1	T	T	C	T	C	A	C	G	A
l	0.00045	1	1	■	T	C	T	C	A	T	G	A

Table 13.72 - IGKJ4

h.type	cum.fre	# indvs	pops	2:89160428 rs191712435 Phe->Leu	2:89160430 rs187905157 Thr->Ile	ENST00000390239
ref	0.9883	-	-	A	G	
b	0.00225	5	4	A	G	2,5,13,14
a	0.0018	4	4	A		2,5,6,10
g	0.00135	3	3	A	G	3,4,10
j	0.0009	2	1	A	G	10
k	0.0009	2	2	A	G	9,10
c	0.0009	2	2	G	G	2,14
h	0.0009	2	2	A	G	2,10
d	0.00045	1	1	A	G	5
e	0.00045	1	1	A	G	6
f	0.00045	1	1	A	G	14
i	0.00045	1	1	A	G	9
m	0.00045	1	1	G	G	9
l	0.00045	1	1	G	G	13

Table 13.73 - IGKJ5

h.type	cum.fre	# indvs	dops	2:89160082 rs185692630 Lys->Arg	2:89160087 rs142632901 Glu->Asp	2:89160106 rs190985433 Gly->Ala	2:89160109 rs140799895 Phe->Cys	2:89160110 rs114105515 Phe->Leu	2:89160113 rs188703994 Thr->Ser	ENST00000390238
ref	0.9928	-	-	T	C	C	A	A	T	
a	0.00225	5	4	**C**	C	C	A	A	T	4,5,6,14
c	0.00135	3	3	T	C	C	A	**G**	T	9,10,11
b	0.00135	3	3	T	**G**	C	A	A	T	3,5,9
d	0.00045	1	1	T	**G**	C	A	**G**	T	10
g	0.00045	1	1	T	C	C	A	A	**■**	10
e	0.00045	1	1	T	C	**G**	A	A	T	10
h	0.00045	1	1	**C**	**G**	C	A	A	T	14
f	0.00045	1	1	T	C	C	**C**	A	T	3

Table 13.74 - IGKV1-16

h.type	cum.freq	# indvs	pops	2:89399455 rs193208674 Lys->Arg	2:89399488 rs184390186 Ala->Val	2:89399509 rs116145759 Ala->Val	2:89399515 rs2848410 Gly->Glu	2:89399809 rs188494311 Ala->Thr	ENST00000479981
ref	0.97625	-	-	T	G	G	C	C	
a	0.01145	25	8	**C**	G	G	**T**	C	1,4,6,7,8,9,11,13
b	0.00595	13	10	**C**	G	G	C	C	2,4,5,6,7,8,9,10,13,14
c	0.00455	10	7	T	G	G	**T**	C	2,5,6,9,11,13,14
d	0.0009	2	2	T	G	■	C	C	1,14
e	0.00045	1	1	T	G	G	C	**T**	9
f	0.00045	1	1	T	■	G	C	C	7

Table 13.75 - IGKV1-17

h.type	cum.freq	# indvs	pops	2:89416951 rs185657549 Ser->Thr	2:89417000 rs188985676 Pro->Ser	2:89417011 rs191430092 Tyr->Phe	2:89417047 rs188555641 Arg->Gln	2:89417087 rs181054320 Leu->Met	ENST00000490686
ref	0.83295	-	-	C	G	T	C	G	
a	0.08655	170	13	C	G	T	C	**T**	1,2,3,4,5,6,7,9,10,11,12,13,14
b	0.0545	111	12	C	G	▮	C	**T**	1,2,3,4,5,7,9,10,11,12,13,14
c	0.0224	47	11	C	G	▮	C	G	1,2,4,5,6,7,9,10,12,13,14
d	0.00225	4	1	C	▮	T	C	G	10
e	0.0009	2	2	**G**	G	T	C	G	2,10
f	0.00045	1	1	C	G	T	**T**	G	7

Table 13.76 - IGKV1-5

h.type	cum.freq	# indvs	pops	2:89246874 rs11546105 Ser->Asn	2:89246904 rs182050897 Ser->Tyr	2:89246946 rs11546098 Ser->Asn	2:89247012 rs146302149 Ser->Thr	2:89247045 rs191761898 Thr->Ser	2:89247051 rs182340840 Arg->Thr	2:89247106 rs189714283 Cys->Gly	ENST00000496168
ref	0.9823	-	-	C	G	C	C	G	C	A	
a	0.0073	16	8	C	G	C	**G**	G	C	A	1,6,8,9,11,12,13,14
b	0.005	11	8	C	G	**T**	C	G	C	A	1,2,3,6,7,11,12,14
f	0.0018	4	4	**T**	G	C	C	G	C	A	7,10,13,14
c	0.00135	3	3	C	G	**T**	**G**	G	C	A	10,12,13
d	0.0009	2	1	C	G	C	C	G	C	**C**	10
g	0.00045	1	1	C	G	C	C	G	**G**	A	11
e	0.00045	1	1	C	G	C	C	**C**	C	A	11
h	0.00045	1	1	C	**T**	C	C	G	C	A	14

Table 13.77 - IGKV1-6

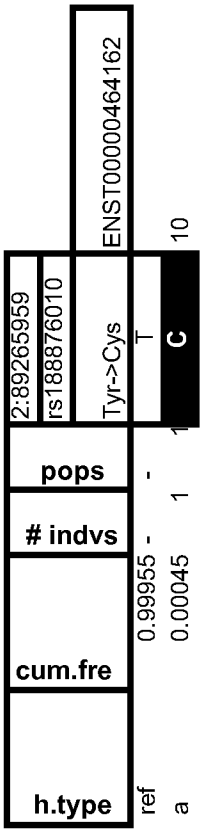


Table 13.78 - IGKV1-8

h.type	cum.freq	# indvs	pops	2:89291968	2:89292080	2:89292130	2:89292149	2:89292182	2:89292188	2:89292201	
				rs187771921	rs190853230	rs183046670	rs186879210	rs59175543	rs192180011	rs35262290	
				Asp->Ala	Lys->Glu	Gly->Asp	Thr->Ser	Phe->Leu	Ser->Pro	Met->Ile	ENST00000495489
ref	0.60715	-	-	T	T	C	T	A	A	C	
a	0.20315	348	14	T	T	C	T	G	A	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.1334	235	14	T	T	C	T	G	A	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.03325	70	14	T	C	C	T	A	A	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
e	0.01035	23	8	T	C	C	T	G	A	C	1,6,7,10,11,12,13,14
d	0.00955	21	8	T	C	C	T	G	A	T	1,3,4,5,7,9,10,14
h	0.0009	2	2	T	T	T	T	G	A	T	9,10
j	0.00045	1	1	G	T	C	T	A	A	C	7
k	0.00045	1	1	T	C	T	T	G	A	C	1
g	0.00045	1	1	T	T	C	■	G	A	C	14
f	0.00045	1	1	T	T	C	T	G	G	T	5
i	0.00045	1	1	T	T	T	T	A	A	C	11

Table 13.79 - IGKV1-9

h.type	cum.freq	# indvs	pops	2:89309554 rs147109427 Glu->Asp	2:89309694 rs185997823 Arg->Trp	2:89309735 rs80322626 Phe->Ser	2:89309762 rs77800356 Asp->Ala	ENST00000493819
ref	0.7266	-	-	T	G	A	T	
a	0.1854	317	14		G	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.02735	57	10	T	G	A	G	1,3,4,5,7,9,10,11,13,14
c	0.0197	43	13		G	A	G	1,2,3,4,5,6,7,9,10,11,12,13,14
d	0.01825	40	13		G	A	T	1,2,3,4,5,7,8,9,10,11,12,13,14
e	0.01095	24	11	T		A	T	2,4,5,6,7,8,9,10,11,12,13
f	0.00545	12	7	T	G	G	G	1,3,4,5,7,10,11
g	0.0018	4	4	T	G	G	T	1,10,11,14
h	0.0018	4	4	T		A	G	3,4,10,11
j	0.00135	3	3		G	G	T	5,12,13
i	0.0009	2	2			G	G	4,14
k	0.00045	1	1			A	G	4

Table 13.80 - IGKV1D-12

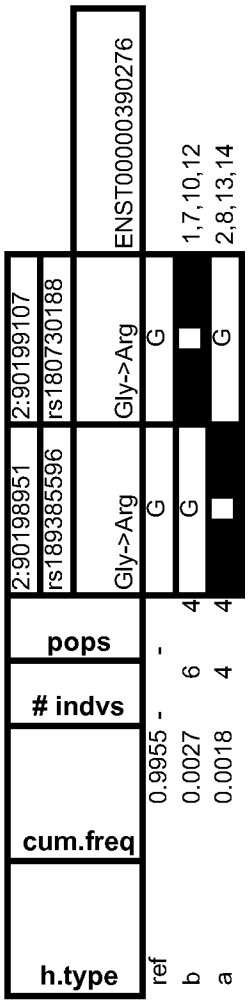


Table 13.81 - IGKV1D-16

h.type	cum.freq	# indvs	pops	2:90139373 rs6760987 Ser->Arg	2:90139377 rs184550323 Gly->Ser	2:90139422 rs189369337 Ala->Thr	ENST00000492446
ref	0.96165	-	-	T	G	G	
a	0.03655	71	13	**G**	G	G	1,2,3,4,5,6,7,8,10,11,12,13,14
c	0.0009	2	1	T	G	■	1
b	0.0009	2	1	T	■	G	10

Table 13.82 - IGKV1D-17

h.type	cum.fre	# indvs	pops	2:90121919	2:90121961	2:90121999	2:90122063	
				rs182773099	rs191806633	rs182889841	rs193008187	ENST00000483379
				Arg->Gln	Gln->Arg	Ala->Ser	Thr->Ile	
ref	0.99315	-	-	G	A	G	C	
a	0.0037	8	4	G	A	**T**	C	1,5,10,14
c	0.00135	3	2	G	**G**	G	C	7,12
b	0.00135	3	2	■	A	G	C	1,10
d	0.00045	1	1	G	A	G	**T**	6

Table 13.83 - IGKV1D-42

h.type	cum.freq	# indvs	pops	2:90229083	2:90229247	2:90229260	2:90229265	2:90229292	2:90229314	2:90229471	2:90229477	2:90229491
				rs193096486	rs139321345	rs842173	rs145538558	rs147724164	rs192596253	rs184070665	rs188365115	rs181576535
				Ala->Thr	Asp->Tyr	Ile->Thr	Ser->Pro	Gly->Arg	Cys->Phe	Ile->Met	Ser->Arg	Asp->Val
ref	0.7584	-	-	G	G	T	T	G	G	C	C	A
a	0.22345	383	14	G	G	**C**	T	G	G	C	C	A
b	0.01005	21	3	G	G	T	**C**	G	G	C	C	A
c	0.0027	6	1	G	G	T	T	■	G	C	C	A
d	0.00135	3	2	■	G	T	T	G	G	C	C	A
e	0.00135	3	1	G	**T**	T	T	G	G	C	C	A
g	0.0009	2	2	G	G	T	T	G	G	**G**	C	A
f	0.0009	2	1	G	G	T	T	G	**T**	C	C	A
h	0.00045	1	1	G	G	**C**	T	G	G	C	■	A
i	0.00045	1	1	G	G	T	T	G	G	C	C	**T**

Table 13.83 - IGKV1D-42

h.type	cum.freq	# indvs	pops	ENST00000390278
ref	0.7584	-	-	
a	0.22345	383	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.01005	21	3	1,10,14
c	0.0027	6	1	9
d	0.00135	3	2	10,14
e	0.00135	3	1	13
g	0.0009	2	2	12,14
f	0.0009	2	1	14
h	0.00045	1	1	13
i	0.00045	1	1	3

Table 13.84 - IGKV1D-43

h.type	cum.freq	# indvs	pops	2:90248957 rs116586875 Leu->Gln	2:90249120 rs188075270 Met->Val	2:90249147 rs184607123 Ala->Thr	2:90249151 rs138753337 Ser->Phe	2:90249237 rs142796670 Ala->Thr	2:90249247 rs192607496 Leu->Arg	2:90249348 rs184789392 Pro->Ser	2:90249372 rs188397844 Cys->Arg	ENST00000468879 ENST00000560950 Cys->Arg	
ref	0.98095	-	-	T	A	G	C	G	T	C	T		
b	0.00685	14	3	T	A	G	**T**	G	T	C	T		1,10,14
a	0.00635	12	4	■	A	G	C	G	T	C	T		1,10,11,14
c	0.0027	5	3	T	A	G	C	■	T	C	T		1,6,14
d	0.0009	2	2	T	A	G	C	G	T	C	**C**		1,14
e	0.0009	2	1	T	A	G	C	G	T	**T**	T		11
g	0.00045	1	1	T	**G**	G	C	G	T	C	T		2
h	0.00045	1	1	T	A	G	C	G	**G**	C	T		2
f	0.00045	1	1	T	A	■	C	G	T	C	T		14

Table 13.85 - IGKV1D-8

h.type	cum.freq	# indvs	pops	2:90259956	2:90259965	2:90260022	2:90260033	2:90260085	2:90260091	2:90260178	
				rs190239413	rs842156	rs17699428	rs183140988	rs186043133	rs190501712	rs138677672	
				Ala->Val	Val->Ala	Thr->Ile	Arg->Trp	Gly->Glu	Ala->Val	Thr->Ile	ENST00000471857
ref	0.7813	-	-	C	T	C	C	G	C	C	
a	0.1872	319	14	C	C	C	C	G	C	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.01785	36	3	C	T	C	C	G	C	T	3,4,9
c	0.0105	22	6	C	T	T	C	G	C	C	2,5,6,7,12,13
d	0.0009	2	1	C	C	C	C	G	T	C	1
f	0.0009	2	2	T	T	C	C	G	C	C	2,12
g	0.00045	1	1	C	T	C	C	-	C	C	1
e	0.00045	1	1	C	T	C	T	G	C	C	1
h	0.00045	1	1	C	T	C	C	G	T	T	9

Table 13.86 - IGKV2-24

h.type	cum.freq	# indvs	pops	2:89475831 rs149900655 Met->Thr	2:89476044 rs183417517 Cys->Phe	ENST00000484817
ref	0.8082	-	-	A	C	
a	0.1708	293	13	G	C	1,2,3,4,5,6,7,9,10,11,12,13,14
b	0.02055	44	13	G	■	1,2,3,4,5,6,7,8,9,10,12,13,14
c	0.00045	1	1	A	■	8

Table 13.87 - IGKV2-30

h.type	cum.fre	# indvs	pops	2:89544389	2:89544392	2:89544473	2:89545042	
				rs141762272	rs182126464	rs147079411	rs186725977	
				Arg->Trp	Asn->Asp	Tyr->His	Leu->Arg	ENST00000468494
ref	0.75005	-	-	G	T	A	A	
a	0.24085	437	13	G	T	**G**	A	1,2,3,4,5,6,7,9,10,11,12,13,14
b	0.0082	18	7	■	T	A	A	2,4,6,7,9,12,13
d	0.00045	1	1	G	T	A	**C**	11
c	0.00045	1	1	G	**C**	A	A	11

Table 13.88 - IGKV2D-24

h.type	cum.freq	# indvs	pops	2:90044207 rs150942262 Phe->Cys	2:90044426 rs190167237 Ala->Val	ENST00000462693
ref	0.9863	-	-	T	C	
a	0.01235	27	12	**G**	C	1,2,3,4,6,7,8,9,11,12,13,14
b	0.00135	3	2	T	**T**	5,10

Table 13.89 - IGKV2D-26

h.type	cum.freq	# indvs	pops	2:90025217 rs187527217 Ile->Asn	2:90025275 rs147187346 Met->Ile	2:90025397 rs192055191 Gly->Glu	2:90025483 rs184819753 Tyr->Asn	ENST00000390268
ref	0.6664	-	-	T	G	G	T	
a	0.325	556	14	T	**C**	G	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.00455	9	3	T	G	G	■	2,11,12
c	0.0018	4	1	■	G	G	T	10
d	0.0009	2	2	T	**C**	G	■	3,4
f	0.0009	2	2	T	G	■	T	1,10
e	0.00045	1	1	■	**C**	G	T	10

Table 13.90 - IGKV2D-29

h.type	cum.freq	# indvs	pops	2:89986777 rs140693346 Ala->Thr	2:89986852 rs191653248 Ser->Pro	2:89986921 rs424406 Pro->Ser	ENST00000491977
ref	0.6969	-	-	G	T	C	
a	0.30175	481	14	G	T	**T**	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0009	2	1	■	T	C	13
c	0.00045	1	1	G	**C**	**T**	12

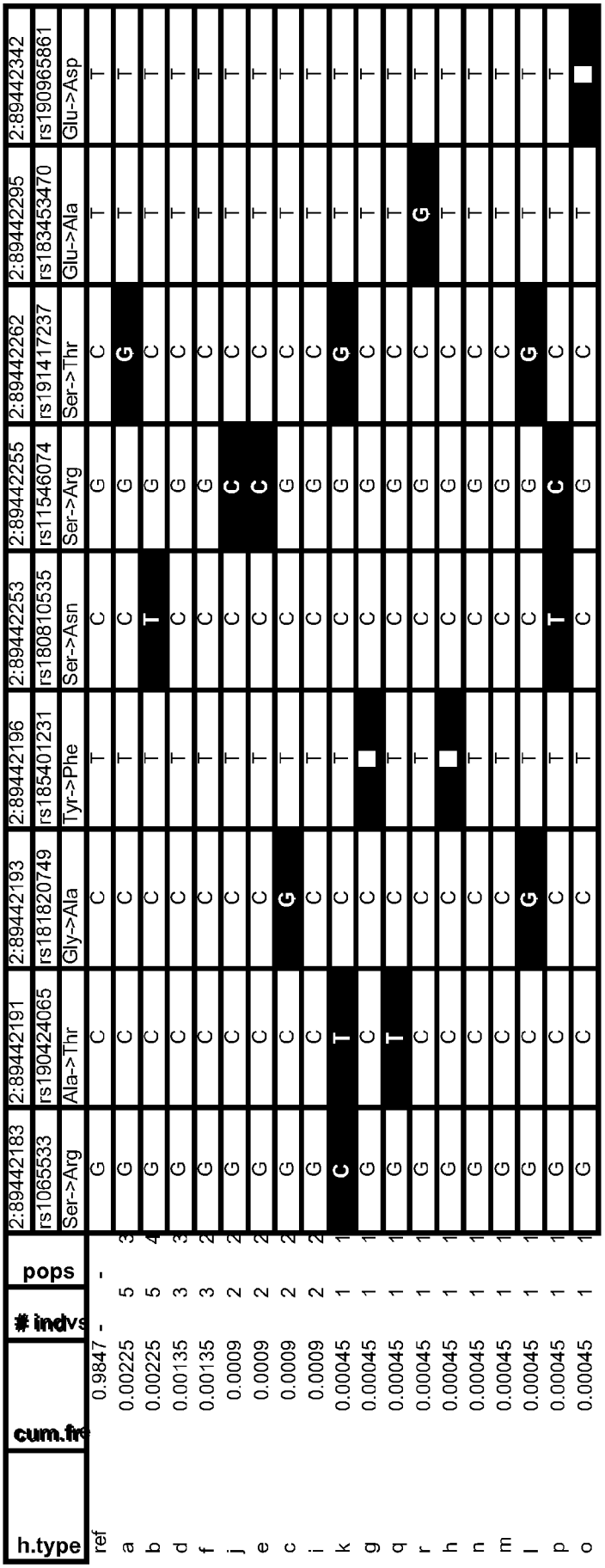
Table 13.91 - IGKV2D-30

h.type	cum.freq	# indvs	pops	2:89975711	2:89976341	2:89976353	2:89976364		
				rs185844838	rs148937425	rs184425514	rs147056147		
				Leu->His	Arg->His	Lys->Thr	Trp->Arg	ENST00000474213	
				->	Arg->His	Lys->Thr	Trp->Arg	ENST00000558962	
ref	0.9689	-	-	T	G	A	T		
a	0.02335	51	14	T	G	A	**C**	1,2,3,4,5,6,7,8,9,10,11,12,13,14	
b	0.00595	13	4	**■**	G	A	T	1,5,10,14	
c	0.00135	3	2	T	**■**	A	T	1,14	
d	0.00045	1	1	T	G	**C**	T	14	

Table 13.92 - IGKV3-11

h.type	cum.fr	# indvs	pops	2:89326669 rs182958807 Pro->His	2:89326681 rs185920904 Arg->His	2:89326754 rs191612627 Ser->Pro	2:89326858 rs182380932 Tyr->Ser	2:89326861 rs191953548 Ser->Thr	2:89326865 rs183365521 Ser->Gly	2:89326872 rs189180740 Gln->His	ENST00000483158
ref	0.94435	-	-	G	C	A	T	C	T	C	
a	0.02835	60	14	**T**	C	**G**	T	C	T	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0096	21	11	G	C	**G**	T	C	T	C	2,3,4,5,6,7,8,9,11,13,14
d	0.0073	14	3	G	**T**	A	T	C	T	C	1,10,14
c	0.00545	12	7	**T**	C	A	T	C	T	C	1,6,7,9,11,13,14
e	0.0018	4	4	G	C	A	T	**G**	T	C	2,5,7,9
g	0.0009	2	1	G	**T**	**G**	T	C	T	C	14
f	0.0009	2	2	G	C	A	T	C	T	**G**	9,13
j	0.00045	1	1	G	C	A	**G**	C	**C**	C	14
h	0.00045	1	1	G	C	A	T	C	**C**	C	13
i	0.00045	1	1	G	C	A	**G**	C	T	C	9

Table 13.93 - IGKV3-20



h.type	cum.fre	# indvs	pops	2:89442183 rs1065533 Ser->Arg	2:89442191 rs190424065 Ala->Thr	2:89442193 rs181820749 Gly->Ala	2:89442196 rs185401231 Tyr->Phe	2:89442253 rs180810535 Ser->Asn	2:89442255 rs11546074 Ser->Arg	2:89442262 rs191417237 Ser->Thr	2:89442295 rs183453470 Glu->Ala	2:89442342 rs190965861 Glu->Asp
ref	0.9847	-	-	G	C	C	T	C	G	C	T	T
a	0.00225	5	3	G	C	C	T	C	G	G	T	T
b	0.00225	5	4	G	C	C	T	T	G	C	T	T
d	0.00135	3	3	G	C	C	T	C	G	C	T	T
f	0.00135	3	2	G	C	C	T	C	G	C	T	T
j	0.0009	2	2	G	C	C	T	C	C	C	T	T
e	0.0009	2	2	G	C	C	T	C	C	C	T	T
c	0.0009	2	2	G	C	G	T	C	G	C	T	T
i	0.0009	2	2	G	C	C	T	C	G	C	T	T
k	0.00045	1	1	C	T	C	T	C	G	G	T	T
g	0.00045	1	1	G	C	C	■	C	G	C	T	T
q	0.00045	1	1	G	T	C	T	C	G	C	T	T
r	0.00045	1	1	G	C	C	T	C	G	C	G	T
h	0.00045	1	1	G	C	C	■	C	G	C	T	T
n	0.00045	1	1	G	C	C	T	C	G	C	T	T
m	0.00045	1	1	G	C	C	T	C	G	C	T	T
l	0.00045	1	1	G	C	G	T	C	G	G	T	T
p	0.00045	1	1	G	C	C	T	T	C	C	T	T
o	0.00045	1	1	G	C	C	T	C	G	C	T	■

Table 13.93 - IGKV3-20

h.type	cum.freq	# indvs	dops	2:89442344	2:89442350	2:89442352	2:89442354	2:89442355	
				rs183373822	rs187696503	rs182188819	rs187989142	rs192831943	
				Glu->Gln	Thr->Ser	Thr->Ile	Asp->GluAsp->Glu	Asp->AlaAsp->Ala	ENST00000492167
ref	0.9847	-	-	C	T	G	A	T	
a	0.00225	5	3	C	T	G	A	T	1,9,14
b	0.00225	5	4	C	T	G	A	T	5,7,10,14
d	0.00135	3	3	C	T	**-**	A	T	2,6,14
f	0.00135	3	2	C	**-**	G	A	T	1,14
j	0.0009	2	2	C	T	**-**	A	T	1,10
e	0.0009	2	2	C	T	G	A	T	9,13
c	0.0009	2	2	C	T	G	A	T	10,11
i	0.0009	2	2	C	T	G	**T**	T	9,14
k	0.00045	1	1	C	T	G	A	T	11
g	0.00045	1	1	**G**	T	G	A	T	14
q	0.00045	1	1	C	T	G	A	T	13
r	0.00045	1	1	C	T	G	A	T	9
h	0.00045	1	1	C	T	G	A	T	14
n	0.00045	1	1	C	T	G	A	**G**	10
m	0.00045	1	1	**G**	T	G	A	T	13
l	0.00045	1	1	C	T	G	A	T	10
p	0.00045	1	1	C	T	G	A	T	2
o	0.00045	1	1	C	T	G	A	T	4

Table 13.94 - IGKV3-7

h.type	cum.fre	# indvs	pops	2:89278103	2:89278106	2:89278219	2:89278262	
				rs144594870	rs189211039	rs2458828	rs181802934	
				Ser->Gly	Thr->Ser	Val->Ala	Thr->Ala	ENST00000390247
ref	0.78325	-	-	T	T	A	T	
a	0.2018	349	14	C	T	A	T	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.01	19	3	T	T	G	T	1,10,14
c	0.00225	5	3	T	[redacted]	A	T	1,10,14
d	0.00225	5	3	T	T	G	C	1,10,14
e	0.00045	1	1	C	T	G	T	9

Table 13.95 - IGKV3D-11

h.type	cum.fr	# indvs	pops	2:90212050 rs58164053 Gly->Ser	2:90212051 rs188947989 Gly->Val	2:90212126 rs7578124 Asn->Thr	2:90212138 rs1916178148 Gly->Val	2:90212167 rs1854410352 Pro->Ser	2:90212198 rs189915304 Ser->Asn	2:90212240 rs117046277 Arg->His	2:90212243 rs181647122 Ser->Ile	ENST00000390277
ref	0.635	-	-	G	G	A	G	C	G	G	G	
a	0.3511	561	14	■	G	A	G	C	G	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0036	8	5	G	G	A	G	T	G	G	G	4,6,7,8,9
b	0.0023	5	3	G	G	A	G	C	G	■	G	3,4,9
g	0.0018	3	2	■	G	A	G	C	G	G	T	1,10
f	0.0018	4	3	■	G	A	G	C	G	■	G	1,3,4
d	0.0014	3	2	G	G	A	G	C	■	G	G	10,14
h	0.0009	2	1	G	G	C	G	C	G	G	G	4
i	0.0009	2	1	G	G	A	T	C	G	G	G	10
j	0.0005	1	1	■	G	A	G	T	G	■	G	14
k	0.0005	1	1	G	T	A	G	C	G	G	G	9
e	0.0005	1	1	■	G	A	G	T	G	G	G	1

Table 13.96 - IGKV3D-15

h.type	cum.freq	# indvs	pops	2:90154131	rs146414056	Thr->Ile	ENST00000417279
ref	0.7939	-	-			C	
a	0.2061	368	14			T	1,2,3,4,5,6,7,8,9,10,11,12,13,14

Table 13.97 - IGKV3D-20

h.type	cum.freq	# indvs	pops	2:90077781 rs183765055 Glu->Gly	2:90078037 rs189772344 Thr->Met	2:90078079 rs182769673 Ala->Gly	2:90078093 rs2555982 Gly->Arg	ENST00000390270
ref	0.9845	-	-	A	C	C	G	
a	0.0073	16	9	A	C	C	■	1,2,4,6,7,8,9,10,11
b	0.0064	12	4	**G**	C	C	G	1,10,11,14
d	0.0009	2	2	A	**T**	C	■	7,12
e	0.00045	1	1	A	**T**	C	G	11
c	0.00045	1	1	A	C	**G**	G	13

Table 13.98 - IGKV4-1

h.type	cum.freq	# indvs	pops	2:89185362 rs180851255 Tyr->Phe	2:89185392 rs116328981 Asp->Gly	2:89185400 rs148071875 Ala->Thr	2:89185431 rs141793384 Asn->Ser	2:89185450 rs189675840 Ser->Arg	2:89185451 rs138603647 Val->Ile	2:89185455 rs181816666 Leu->Ser	2:89185457 rs141709704 Tyr->His	2:89185461 rs78280375 Ser->Asn
ref	0.9516	-	-	A	A	G	A	T	G	T	T	G
a	0.01185	26	11	A	A	G	A	T	G	T	T	G
b	0.00545	11	9	A	A	G	A	T	[redacted]	T	T	G
c	0.00365	8	7	A	A	G	A	T	G	T	T	G
d	0.00315	7	6	A	A	[redacted]	A	T	G	T	T	G
k	0.0018	4	4	A	A	G	A	T	G	T	T	[redacted]
w	0.0018	3	3	A	A	G	A	T	G	T	T	G
h	0.0018	4	4	T	A	G	A	T	G	T	T	G
f	0.0018	4	4	A	A	G	A	T	G	T	C	G
e	0.00135	3	3	A	A	G	A	T	G	T	T	G
r	0.00135	3	3	A	A	G	A	T	G	T	T	G
x	0.00135	3	2	A	A	G	A	T	G	T	T	G
o	0.00135	3	2	A	A	G	A	T	G	T	T	G
y	0.0009	2	2	A	A	G	A	T	G	T	T	G
g	0.0009	2	2	A	A	G	A	T	G	T	T	G
t	0.0009	2	2	A	A	G	A	T	G	T	T	G
i	0.0009	2	2	A	A	G	A	T	G	T	T	G
n	0.0009	2	2	A	A	G	A	T	G	T	T	G
l	0.0009	2	2	A	A	G	G	T	G	T	T	G
j	0.00045	1	1	A	A	G	A	T	G	T	T	G
u	0.00045	1	1	A	A	G	A	T	G	T	T	G
aa	0.00045	1	1	A	A	G	A	T	G	C	T	G
v	0.00045	1	1	A	A	[redacted]	A	T	[redacted]	T	T	G
s	0.00045	1	1	A	A	G	A	[redacted]	G	T	T	G
ab	0.00045	1	1	A	A	G	A	T	G	T	T	G
q	0.00045	1	1	T	A	G	A	T	G	T	T	G
z	0.00045	1	1	A	A	G	A	T	G	T	T	G
af	0.00045	1	1	A	A	G	A	T	G	T	T	G
ad	0.00045	1	1	A	A	G	A	T	G	T	T	G

Table 13.98 - IGKV4-1

ac	0.00045	1	1	A	A	G	A	T	G	T	T	G
m	0.00045	1	1	A	A	G	A	T	■	T	T	■
p	0.00045	1	1	A	**G**	G	A	T	G	T	T	G
ae	0.00045	1	1	A	A	G	A	T	G	T	T	G

Table 13.98 - IGKV4-1

h.type	cum.freq	# indvs	dpds	2:89185464 rs139176691 Ser->Phe	2:89185466 rs186837913 Asn->His	2:89185474 rs142402063 Lys->Asn	2:89185476 rs191518466 Asn->Ser	2:89185482 rs143877691 Leu->Ser	2:89185484 rs148599313 Ala->Pro	2:89185485 rs142986022 Ala->Val	2:89185491 rs140954172 Tyr->Phe	2:89185511 rs186251122 Pro->Thr
ref	0.9516	-	-	C	A	G	A	T	G	C	A	C
a	0.01185	26	11	C	A	G	A	T	G	C	A	C
b	0.00545	11	9	C	A	G	A	T	G	C	A	C
c	0.00365	8	7	C	A	G	A	T	G	**T**	A	C
d	0.00315	7	6	C	A	G	A	T	G	C	A	C
k	0.0018	4	4	C	A	G	A	T	G	C	A	C
w	0.0018	3	3	C	A	G	A	T	**C**	C	A	C
h	0.0018	4	4	C	A	G	A	T	G	C	A	C
f	0.0018	4	4	C	A	G	A	T	G	C	A	C
e	0.00135	3	3	C	A	G	A	T	G	C	**T**	C
r	0.00135	3	3	C	A	G	A	T	G	C	A	C
x	0.00135	3	2	C	A	G	A	T	G	C	A	C
o	0.00135	3	2	C	A	G	**G**	T	G	C	A	C
y	0.0009	2	2	C	A	**C**	A	T	**C**	C	A	C
g	0.0009	2	2	C	A	G	A	T	G	C	A	C
t	0.0009	2	2	C	A	**C**	A	T	G	C	A	C
i	0.0009	2	2	C	A	G	A	T	G	C	A	C
n	0.0009	2	2	C	A	G	A	T	G	C	A	C
l	0.0009	2	2	C	A	G	A	T	G	C	A	C
j	0.00045	1	1	C	A	G	A	T	G	C	A	**N**
u	0.00045	1	1	C	A	**C**	A	T	G	C	A	C
aa	0.00045	1	1	C	A	G	A	T	G	C	A	C
v	0.00045	1	1	C	A	G	A	T	G	C	A	C
s	0.00045	1	1	C	A	G	A	T	G	C	A	C
ab	0.00045	1	1	C	A	G	A	T	G	**T**	A	C
q	0.00045	1	1	C	**C**	G	A	T	G	C	A	C
z	0.00045	1	1	C	A	G	A	T	G	C	A	C
af	0.00045	1	1	C	A	G	A	**C**	G	C	A	C
ad	0.00045	1	1	C	A	G	A	T	G	C	A	C

Table 13.98 - IGKV4-1

ac	0.00045	1	1	C	A	G	A	T	G	C	A	C
m	0.00045	1	1	C	A	G	A	T	G	C	**T**	C
p	0.00045	1	1	C	A	G	A	T	G	C	A	C
ae	0.00045	1	1	**T**	A	**C**	A	T	G	C	A	C

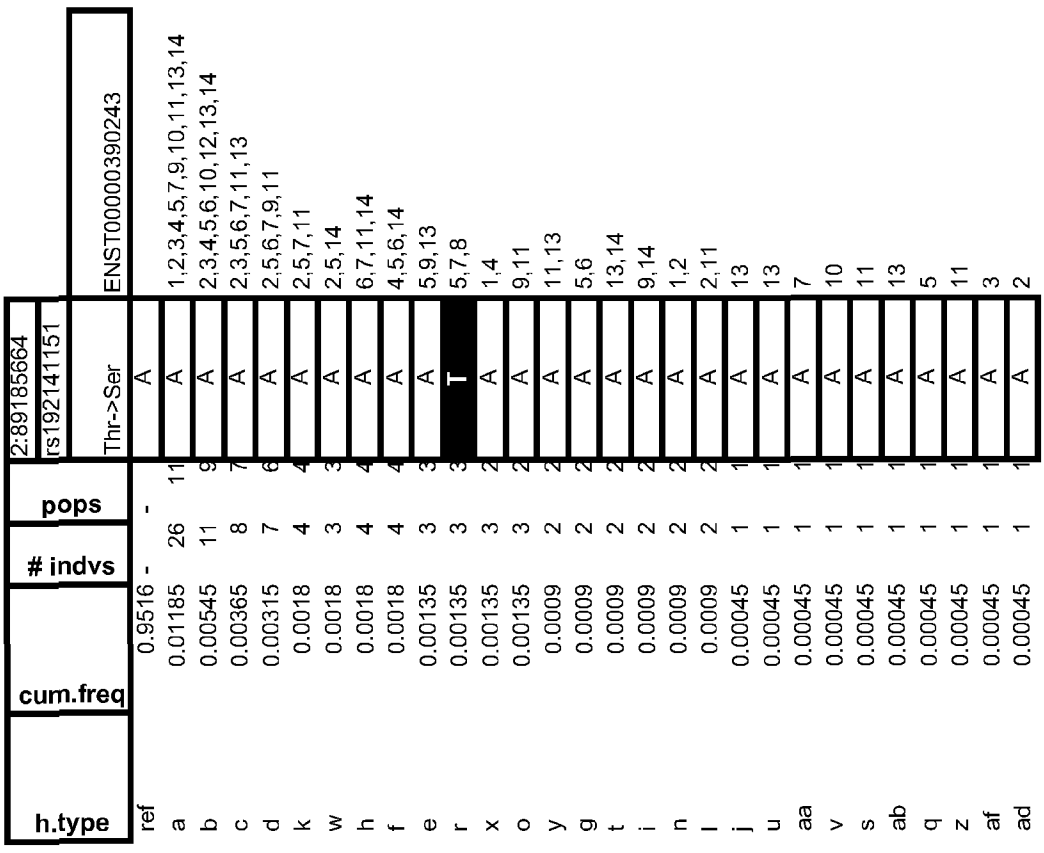
Table 13.98 - IGKV4-1

h.type	cum.freq	# indvs	sdos	2:89185518 rs141640067 Lys->Arg	2:89185529 rs145197543 Tyr->Asp	2:89185530 rs188057788 Tyr->Cys	2:89185563 rs192839829 Asp->Gly	2:89185577 rs138751137 Ser->Arg	2:89185586 rs141777505 Gly->Arg	2:89185607 rs182919936 Ile->Phe	2:89185622 rs188769218 Ala->Thr	2:89185637 rs145449257 Val->Leu
ref	0.9516	-	-	A	T	A	A	A	G	A	G	G
a	0.01185	26	11	A	T	A	A	A	G	A	■	G
b	0.00545	11	9	A	T	A	A	A	G	A	G	G
c	0.00365	8	7	A	T	A	A	A	G	A	G	G
d	0.00315	7	6	A	T	A	A	A	G	A	G	G
k	0.0018	4	4	A	T	A	A	A	G	A	G	G
w	0.0018	3	3	A	T	A	A	A	G	A	G	G
h	0.0018	4	4	A	T	A	A	A	G	A	G	G
f	0.0018	4	4	A	T	A	A	A	G	A	G	G
e	0.00135	3	3	A	T	A	A	A	G	A	G	G
r	0.00135	3	3	A	T	A	A	A	G	A	G	G
x	0.00135	3	2	A	T	A	A	A	G	A	G	**C**
o	0.00135	3	2	A	T	A	A	A	G	A	G	G
y	0.0009	2	2	A	T	A	A	A	G	A	G	G
g	0.0009	2	2	A	T	**G**	A	A	G	A	G	G
t	0.0009	2	2	A	T	A	A	A	G	A	G	G
i	0.0009	2	2	A	**G**	A	A	A	G	A	G	G
n	0.0009	2	2	**G**	T	A	A	A	G	A	G	G
l	0.0009	2	2	A	T	A	A	A	G	A	G	G
j	0.00045	1	1	A	T	A	A	A	G	A	G	G
u	0.00045	1	1	A	T	A	A	**C**	G	A	G	G
aa	0.00045	1	1	A	T	A	A	A	G	A	G	G
v	0.00045	1	1	A	T	A	A	A	G	A	G	G
s	0.00045	1	1	A	T	A	A	A	G	A	G	G
ab	0.00045	1	1	A	T	A	A	A	G	A	■	G
q	0.00045	1	1	A	T	A	A	A	G	A	G	G
z	0.00045	1	1	A	T	A	A	A	G	**T**	G	G
af	0.00045	1	1	A	T	A	A	A	G	A	G	G
ad	0.00045	1	1	A	T	A	**G**	A	G	A	G	G

Table 13.98 - IGKV4-1

ac	0.00045	1	1	A	T	A	A	A		A	G	G
m	0.00045	1	1	A	T	A	A	A	G	A	G	G
p	0.00045	1	1	A	T	A	A	A	G	A	G	G
ae	0.00045	1	1	A	T	A	A	A	G	A	G	G

Table 13.98 - IGKV4-1



h.type	cum.freq	# indvs	pops	2:89185664 rs1921141151 Thr->Ser	ENST00000390243
ref	0.9516	-	-	A	
a	0.01185	26	11	A	1,2,3,4,5,7,9,10,11,13,14
b	0.00545	11	9	A	2,3,4,5,6,10,12,13,14
c	0.00365	8	7	A	2,3,5,6,7,11,13
d	0.00315	7	6	A	2,5,6,7,9,11
k	0.0018	4	4	A	2,5,7,11
w	0.0018	3	3	A	2,5,14
h	0.0018	4	4	A	6,7,11,14
f	0.0018	4	4	A	4,5,6,14
e	0.00135	3	3	A	5,9,13
r	0.00135	3	3	T	5,7,8
x	0.00135	3	2	A	1,4
o	0.00135	3	2	A	9,11
y	0.0009	2	2	A	11,13
g	0.0009	2	2	A	5,6
t	0.0009	2	2	A	13,14
i	0.0009	2	2	A	9,14
n	0.0009	2	2	A	1,2
l	0.0009	2	2	A	2,11
j	0.00045	1	1	A	13
u	0.00045	1	1	A	13
aa	0.00045	1	1	A	7
v	0.00045	1	1	A	10
s	0.00045	1	1	A	11
ab	0.00045	1	1	A	13
q	0.00045	1	1	A	5
z	0.00045	1	1	A	11
af	0.00045	1	1	A	3
ad	0.00045	1	1	A	2

Table 13.98 - IGKV4-1

ac	0.00045	1	1	A	1
m	0.00045	1	1	A	12
p	0.00045	1	1	A	14
ae	0.00045	1	1	A	9

Table 13.99 - IGKV5-2

h.type	cum.freq	# indvs	pops	2:89197020 rs1838911175 Thr->Arg	2:89197105 rs189117859 Asp->Glu	2:89197190 rs55661410 Pro->Ser	ENST00000390244
ref	0.6714	-	-	C	T	C	
a	0.32315	501	14	C	T	**T**	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.005	11	6	C	**G**	**T**	2,5,6,7,11,13
c	0.00045	1	1	**G**	T	C	11

Table 13.100 - IGKV6-21

h.type	cum.freq	# indvs	pops	2:89459266 rs181879031 Thr->Met	2:89459267 rs395483 Thr->Ala	ENST00000390256
ref	0.9638	-	-	G	T	
a	0.03575	78	14	G	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.00045	1	1	■	T	14

Table 13.101 - IGKV6D-21

h.type	cum.freq	# indvs	pops	2:90060955 rs190780280 Asp->Gly	2:90060957 rs141599034 Ala->Thr	2:90060963 rs150509294 Ala->Thr	ENST00000436451 Ala->Thr	ENST00000557853 Ala->Thr	ENST00000558079 Ala->Thr
ref	0.9735	-	-	A	G	G			
a	0.0192	42	13	A	G				1,2,3,4,5,6,7,8,9,10,11,12,14
b	0.00685	14	4	A		G			1,5,10,14
c	0.00045	1	1	**G**	G	G			3

Table 13.102 - IGKV6D-41

h.type	cum.freq	# indvs	pops	2:90108534 rs191316044 Val->Ala	2:90108797 . ->Asp	2:90109001 rs192662836 Thr->Ile	ENST00000390271
ref	0.99865	-	-	T	ATGT	C	
a	0.00045	1	1	**C**	ATGT	C	3
c	0.00045	1	1	T	ATGT	**T**	6
b	0.00045	1	1	T	■	C	9

Table 13.103 - IGLJ1

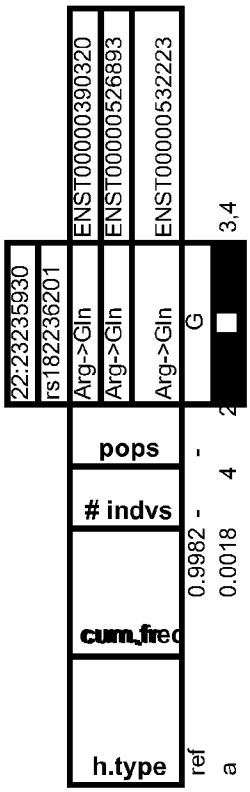


Table 13.104 - IGLJ2

h.type	cum.freq	# indvs	pops	22:23241713 rs182275122 Ala->Gly	22:23241772 rs148186519 Phe->Leu	22:23241773 rs186547678 Phe->Tyr	22:23241777 rs192641205 Leu->Phe	22:23241802 . ->Val	22:23241802 rs185471925 Val->Ile	ENST00000390322
ref	0.90495	-	-	C	T	T	G	GTAT	G	
a	0.0741	160	13	G	T	T	G	GTAT	G	1,2,3,4,5,6,7,9,10,11,12,13,14
b	0.0178	38	12	C	C	T	G	GTAT	G	1,2,3,4,5,7,9,10,11,12,13,14
e	0.0009	2	2	G	C	T	G	GTAT	G	4,9
c	0.0009	2	2	C	T	T	C	GTAT	G	1,10
d	0.00045	1	1	C	T	T	G	GTAT	■	2
g	0.00045	1	1	C	T	■	G	GTAT	G	13
f	0.00045	1	1	C	T	T	G	G	G	2

Table 13.105 - IGLJ3

h.type	cum.freq	# indvs	pops	22:23247139	22:23247164		ENST00000390324	ENST00000448061
				rs185223887	rs139014240			
				Gly->Arg	Gln->Arg			
					->			
ref	0.9973	-	-	G	A			
a	0.00225	4	2	G	G		2,13	
b	0.00045	1	1	■	A		2	

Table 13.106 - IGLJ5

h.type	cum.freq	# indvs	pops	22:23256422 rs1893443862 Leu->Phe	22:23256465 rs192396255 Glu->Lys	ENST00000390327
ref	0.99545	-	-	G	G	1,6,7,12,13
a	0.00365	8	5	T	G	14
b	0.0009	2	1	G	■	

Table 13.107 - IGLJ6

h.type	cum.freq	# indvs	pops	22:23260329 rs76121504 Ser->Leu	22:23260345 rs114500104 Phe->Leu	22:23260367 rs184683407 Val->Ile	ENST00000390328
ref	0.9895	-	-	C	C	G	
a	0.00595	13	3	**T**	C	G	1,12,14
b	0.0032	7	3	C	**G**	G	1,10,14
c	0.00135	3	2	C	C	■	1,10

Table 13.108 - IGLJ7

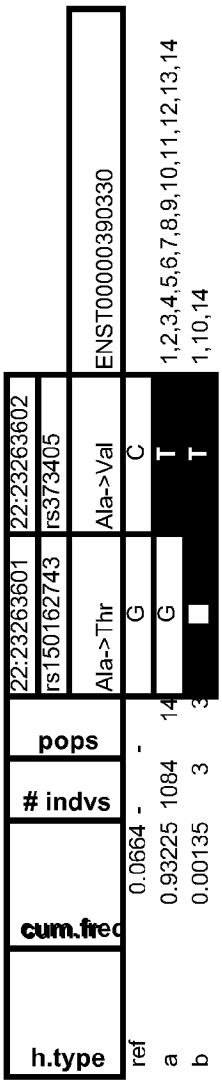


Table 13.109 - IGLV10-54

h.type	cum.fre	# indvs	pops	22:22569242 rs191472679 Val->Leu	22:22569357 rs1450111998 Ser->Thr	22:22569442 rs183335857 Ser->Thr	22:22569448 rs5750741 Ile->Asn	22:22569451 rs76512804 Val->Ala	22:22569468 rs143187632 Ala->Thr	22:22569554 rs6001413 Phe->Leu	22:22569559 rs7291275 Ala->Val
ref	0.3126	-	-	G	T	G	T	T	G	C	C
a	0.50725	812	14	G	T	G	■	T	G	■	C
b	0.09425	196	14	G	T	G	■	T	G	C	C
c	0.05525	109	9	G	T	G	T	T	G	C	C
d	0.01265	27	5	G	T	G	■	C	G	C	C
e	0.0036	8	3	G	T	G	■	T	G	■	C
f	0.00315	7	2	G	T	G	■	T	G	■	C
h	0.0027	6	2	G	T	G	■	T	■	C	C
i	0.0027	6	2	G	T	G	■	T	G	C	C
k	0.0018	4	2	G	T	G	T	T	G	C	T
g	0.00135	3	1	G	■	G	■	T	G	C	C
j	0.0009	2	2	G	T	G	■	T	G	C	C
n	0.0009	2	2	G	T	C	■	T	G	C	C
m	0.00045	1	1	C	T	G	■	T	G	C	C
l	0.00045	1	1	G	T	G	T	T	G	C	C

Table 13.109 - IGLV10-54

h.type	cum.freq	# indivs	pdos	22:22569562	22:22569609	22:22569618	22:22569640	22:22569649	
				rs186196532	rs140258196	rs191049703	rs6001415	rs116733113	
				Ser->Cys	Glu->Lys	Ala->Pro	Leu->Trp	Ser->Asn	ENST00000390287
ref	0.3126	-	-	C	G	G	T	G	
a	0.50725	812	14	C	G	G	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.09425	196	14	C	G	G	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.05525	109	9	C	G	G	G	G	1,3,5,6,9,10,11,12,14
d	0.01265	27	5	C	G	G	G	G	1,5,10,12,14
e	0.0036	8	3	C		G	G	G	10,11,14
f	0.00315	7	2	C	G	G	G		10,14
h	0.0027	6	2	C	G	G	G	G	10,14
i	0.0027	6	2	C	G	C	G	G	1,14
k	0.0018	4	2	C	G	G	G	G	10,14
g	0.00135	3	1	C	G	G	G	G	10
j	0.0009	2	2	C	G	G	T	G	11,13
n	0.0009	2	2	C	G	G	G	G	2,14
m	0.00045	1	1	C	G	G	G	G	13
l	0.00045	1	1	G	G	G	T	G	11

Table 13.110 - IGLV11-55

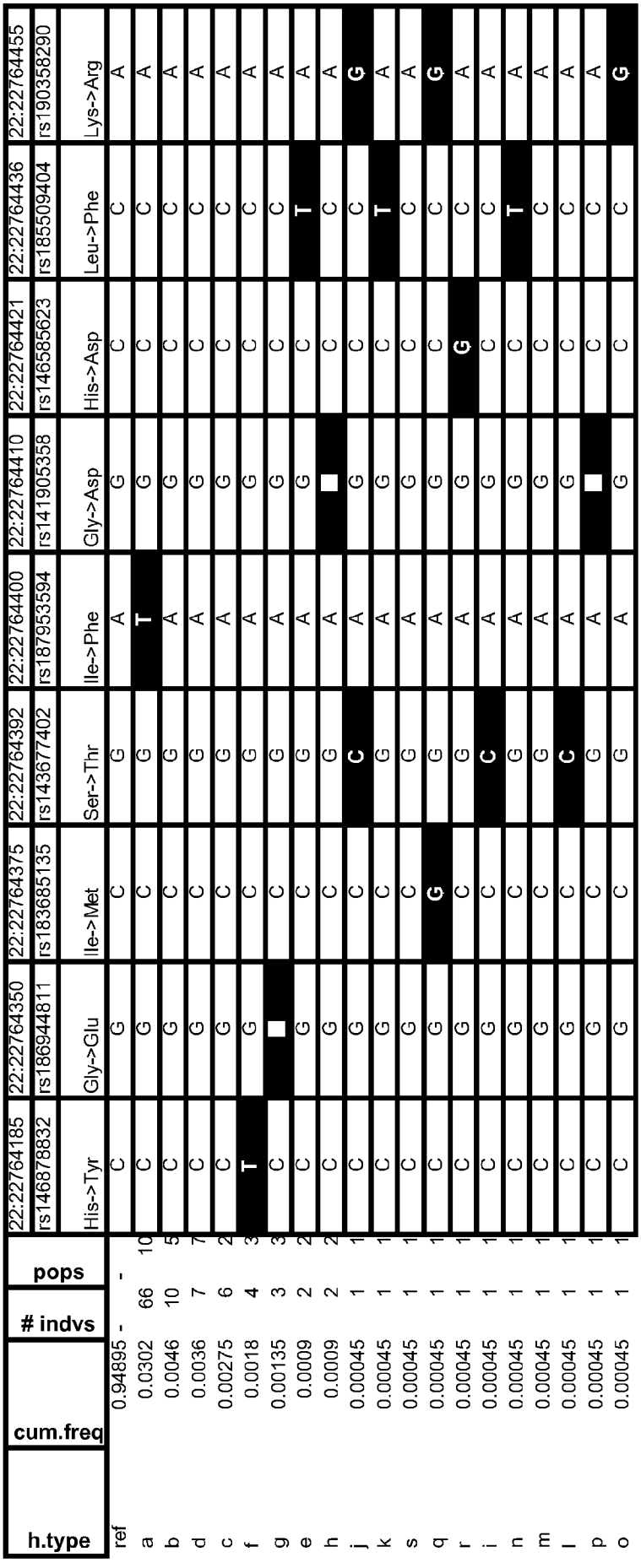
h.type	cum.freq	# indvs	pops	22:22556076 rs117609231 Leu->Gln	22:22556114 rs61749483 Gly->ArgGly->Arg	22:22556368 rs9306341 Pro->Leu	22:22556443 rs188903924 Arg->Gln	22:22556479 rs61744363 Ala->Val	22:22556517 rs151095711 Ala->Pro	22:22556577 rs113498536 Val->Phe	ENST00000390286
ref	0.4994	-	-	T	G	C	G	C	G	G	
a	0.4463	728	14	T	G	**T**	G	C	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0325	71	12	T	**C**	**T**	G	C	G	G	1,2,3,5,6,7,8,9,10,11,13,14
c	0.0096	21	6	T	G	**T**	G	**T**	G	G	3,4,6,9,12,13
d	0.00545	12	6	T	**C**	C	G	C	G	G	2,4,9,10,13,14
g	0.00225	5	3	T	G	**T**	G	**T**	G	**T**	1,10,14
e	0.0018	4	3	[redacted]	G	**T**	G	C	G	G	2,5,13
h	0.00135	3	2	T	G	C	G	C	**C**	G	1,10
f	0.00135	3	2	T	G	**T**	[redacted]	C	G	G	4,9

Table 13.111 - IGLV1-36

h.type	cum.freq	# indvs	pops	22:22786604 rs61735471 Ala->Thr	22:22786673 rs190488947 Pro->Thr	22:22786674 rs1180061119 Pro->Leu	22:22786692 rs182633696 Arg->Gln	22:22786709 rs1391555569 Ser->Pro	22:22786750 rs59686180	22:22786792 rs57337788 Ser->Arg	22:22786797 rs61735468 Asn->Ser
ref	0.8795	-	-	G	C	C	G	T	TGAG	C	A
a	0.0942	185	14	G	C	C	G	T	TGAG	**G**	A
b	0.00685	15	5	G	C	**T**	G	T	TGAG	C	A
d	0.0068	15	4	G	C	C	G	T	TGAG	C	**G**
c	0.00455	10	2	■	C	C	G	T	TGAG	C	A
f	0.0036	8	2	G	■	C	G	**C**	TGAG	C	A
g	0.0027	6	3	G	C	C	G	T	**T**	C	A
e	0.0018	4	3	G	C	C	■	T	TGAG	C	A

ENST00000390301	1,2,3,4,5,6,7,8,9,10,11,12,13,14
	2,6,7,12,13
	1,10,12,14
	10,14
	10,14
	1,4,10
	2,6,7

Table 13.112 - IGLV1-40



h.type	cum.freq	# indvs	pops	22:22764185 rs146878832 His->Tyr	22:22764350 rs186944811 Gly->Glu	22:22764375 rs1836885135 Ile->Met	22:22764392 rs143677402 Ser->Thr	22:22764400 rs187953594 Ile->Phe	22:22764410 rs141905358 Gly->Asp	22:22764421 rs146585623 His->Asp	22:22764436 rs185509404 Leu->Phe	22:22764455 rs190358290 Lys->Arg
ref	0.94895	-	-	C	G	C	G	A	G	C	C	A
a	0.0302	66	10	C	G	C	G	**T**	G	C	C	A
b	0.0046	10	5	C	G	C	G	A	G	C	C	A
d	0.0036	7	7	C	G	C	G	A	G	C	C	A
c	0.00275	6	2	C	G	C	G	A	G	C	C	A
f	0.0018	4	3	**T**	G	C	G	A	G	C	C	A
g	0.00135	3	3	C	■	C	G	A	G	C	C	A
e	0.0009	2	2	C	G	C	G	A	G	C	**T**	A
h	0.0009	2	2	C	G	C	G	A	■	C	C	A
j	0.00045	1	1	C	G	C	**C**	A	G	C	C	**G**
k	0.00045	1	1	C	G	C	G	A	G	C	**T**	A
s	0.00045	1	1	C	G	C	G	A	G	C	C	A
q	0.00045	1	1	C	G	**G**	G	A	G	C	C	**G**
r	0.00045	1	1	C	G	C	G	A	G	**G**	C	A
i	0.00045	1	1	C	G	C	**C**	A	G	C	C	A
n	0.00045	1	1	C	G	C	G	A	G	C	**T**	A
m	0.00045	1	1	C	G	C	G	A	G	C	C	A
l	0.00045	1	1	C	G	C	**C**	A	G	C	C	A
p	0.00045	1	1	C	G	C	G	A	■	C	C	A
o	0.00045	1	1	C	G	C	G	A	G	C	C	**G**

Table 13.112 - IGLV1-40

h.type	cum.freq	# indvs	pops	22:22764472 rs181685335 Asn->Tyr	22:22764473 rs185969886 Asn->Ser	22:22764476 rs111666735 Ser->Asn	22:22764487 rs138088831 Ser->Ala	22:22764512 rs181089093 Gly->Ala	22:22764541 rs190749695 Ala->Thr	22:22764548 rs143634461 Thr->Ser	ENST00000390299
ref	0.94895	-	-	A	A	G	T	G	G	C	
a	0.0302	66	10	A	A	G	T	G	G	C	1,2,3,4,5,6,7,10,13,14
b	0.0046	10	5	A	A	G	T	G	G	G	2,6,7,12,13
d	0.0036	7	7	A	A	[redacted]	T	G	G	C	1,2,5,10,11,13,14
c	0.00275	6	2	A	A	G	G	G	G	C	10,14
f	0.0018	4	3	A	A	G	T	G	G	C	1,10,14
g	0.00135	3	3	A	A	G	T	G	G	C	2,7,12
e	0.0009	2	2	A	A	G	T	G	G	C	10,13
h	0.0009	2	2	A	A	G	T	G	G	C	2,3
j	0.00045	1	1	A	A	G	T	G	G	C	1
k	0.00045	1	1	A	G	G	T	G	G	C	1
s	0.00045	1	1	T	A	G	T	G	G	C	13
q	0.00045	1	1	A	A	G	T	G	G	C	10
r	0.00045	1	1	A	A	G	T	G	G	C	10
i	0.00045	1	1	A	A	[redacted]	T	G	G	C	10
n	0.00045	1	1	A	A	[redacted]	T	G	G	C	2
m	0.00045	1	1	A	A	G	T	C	G	C	2
l	0.00045	1	1	A	A	G	T	G	G	C	10
p	0.00045	1	1	A	A	G	T	G	[redacted]	C	1
o	0.00045	1	1	A	A	G	T	G	G	C	5

Table 13.113 - IGLV1-44

h.type	cum.fre	# indvs	pops	22:22735255 rs183817639 Ser->Gly	22:22735448 rs180819320 Ala->Thr	22:22735517 rs61731361 Thr->Ala	22:22735551 rs190708617 Ala->Gly	22:22735573 rs111733579 Ser->Arg	22:22735661 rs6002086 Ser->Ala	22:22735678 rs181994204 Asp->Glu	22:22735710 rs149125249 Asn->Ser	ENST00000390297
ref	0.9531	-	-	A	G	A	C	T	T	T	A	
a	0.0292	60	12	A	G	A	C	**G**	T	T	A	1,2,3,4,5,6,7,9,10,11,13,14
b	0.00775	16	2	A	G	A	C	**G**	**G**	T	A	10,14
c	0.00365	8	4	A	G	A	**G**	T	T	T	A	6,7,11,13
d	0.0018	4	3	**G**	G	A	C	T	T	T	A	4,11,13
e	0.0018	4	3	A	G	A	C	T	T	■	A	7,13,14
f	0.00135	3	2	A	G	A	C	T	T	T	**G**	1,14
g	0.00045	1	1	A	G	**G**	C	T	T	T	A	9
h	0.00045	1	1	A	■	A	C	T	T	T	A	9
i	0.00045	1	1	A	G	A	C	**G**	T	T	**G**	3

Table 13.114 - IGLV1-47

h.type	cum.fre	# indvs	pops	22:22712158	22:22712333	22:22712403	22:22712410	22:22712444	22:22712467	22:22712552	22:22712562	
				rs12160269	rs185378771	rs181574862	rs185842267	rs8142074	rs5757973	rs61231611	rs184225865	
				Leu->Ile	Pro->Ala	Gly->Glu	Asn->Lys	Ala->Thr	Ser->Arg	Arg->Trp	Asp->Ala	ENST00000390294
ref	0.11455	-	-	C	C	G	T	G	T	C	A	
a	0.8413	1059	14	C	C	G	T	G	G	C	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0182	40	4		C	G	T	G	G	C	A	1,10,12,14
c	0.01695	33	4	C	C	G	T		G	C	A	1,5,10,14
d	0.00405	9	3	C	C	G	T	G	G	T	A	1,10,14
g	0.0018	4	2	C	C	G	G	G	G	C	A	10,14
f	0.00135	3	2	C	C		T	G	G	C	A	8,14
e	0.0009	2	2	C	C	G	T		T	C	A	1,14
h	0.00045	1	1	C	G	G	T	G	G	C	A	13
i	0.00045	1	1	C	C	G	T	G	G	C	C	10

Table 13.115 - IGLV1-50

h.type	cum.freq	# indvs	pops	22:22681721 rs12484322 Ser->Pro	22:22681967 rs140594791 Ala->Val	22:22682039 rs150490781 Asn->Ser	22:22682063 rs73880621 Gln->Arg	ENST00000390291
ref	0.7922	-	-	T	C	A	A	
a	0.19325	373	14	**C**	C	A	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0105	23	5	T	C	A	**G**	1,10,11,12,14
c	0.00225	5	3	T	C	**G**	A	10,12,13
d	0.0018	4	2	T	**T**	A	A	3,9

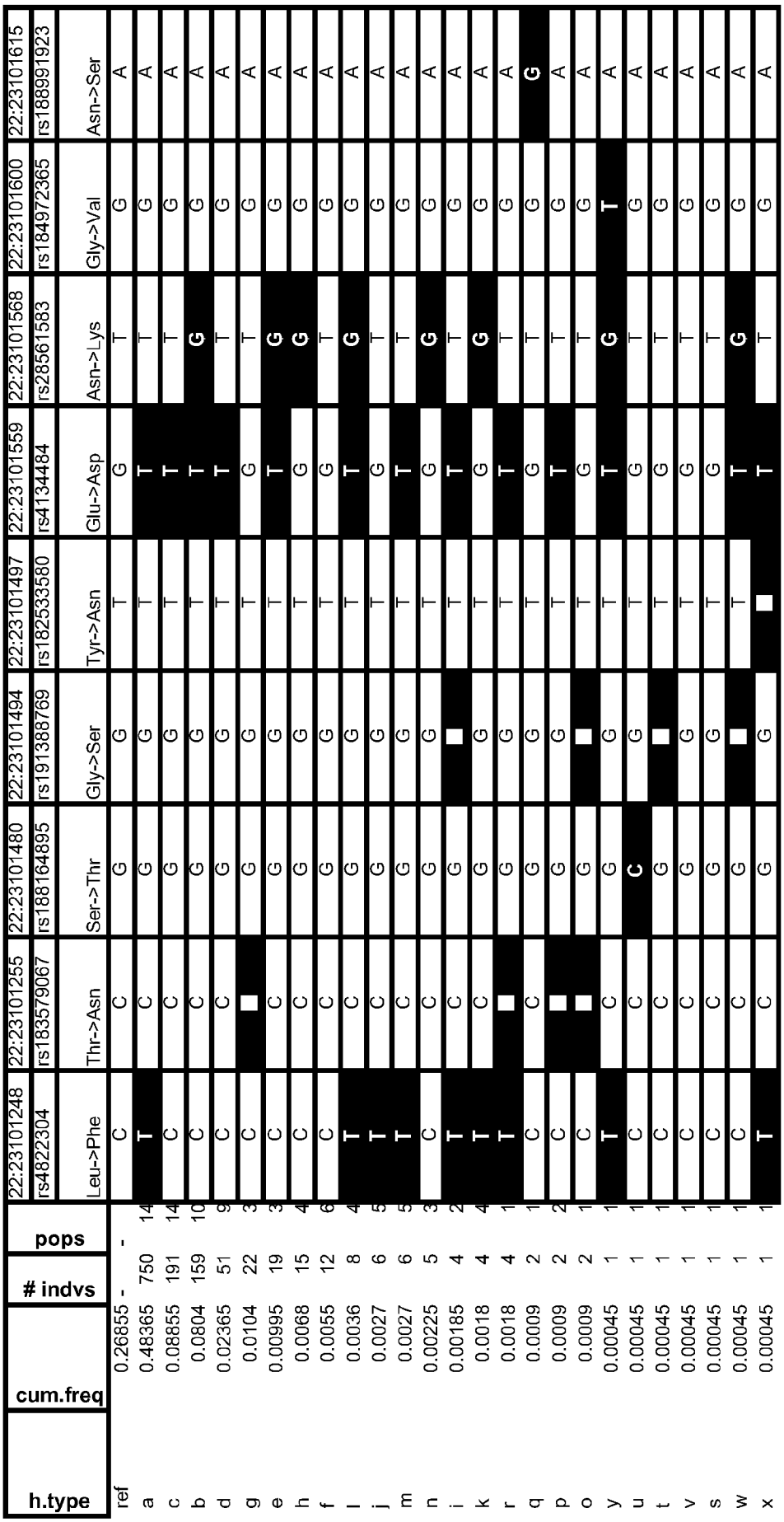
Table 13.116 - IGLV1-51

h.type	cum.freq	# indvs	pops	22:22677182 rs11703618 Asp->Glu	22:22677235 rs186189114 Gly->Asp	22:22677285 rs139024634 Asp->Asn	22:22677298 rs190511809 Gly->Val	22:22677307 rs181313572 Asp->Gly	22:22677309 rs73401707 Ser->Arg	22:22677318 rs190934387 Ser->Arg	ENST00000390290
ref	0.8737	-	-	C	G	G	G	A	A	A	
a	0.1177	221	11	■	G	G	G	A	A	A	1,2,5,6,7,8,10,11,12,13,14
b	0.00275	6	3	C	G	■	G	A	A	A	2,5,12
d	0.0018	4	2	■	G	G	G	A	A	**C**	3,4
c	0.0018	4	3	C	G	G	G	A	A	**C**	3,4,9
g	0.0009	2	2	C	■	G	G	A	A	A	6,10
e	0.00045	1	1	C	G	G	G	A	**C**	A	6
h	0.00045	1	1	C	G	G	**T**	A	A	A	13
f	0.00045	1	1	C	G	G	G	**G**	A	A	6

Table 13.117 - IGLV2-11

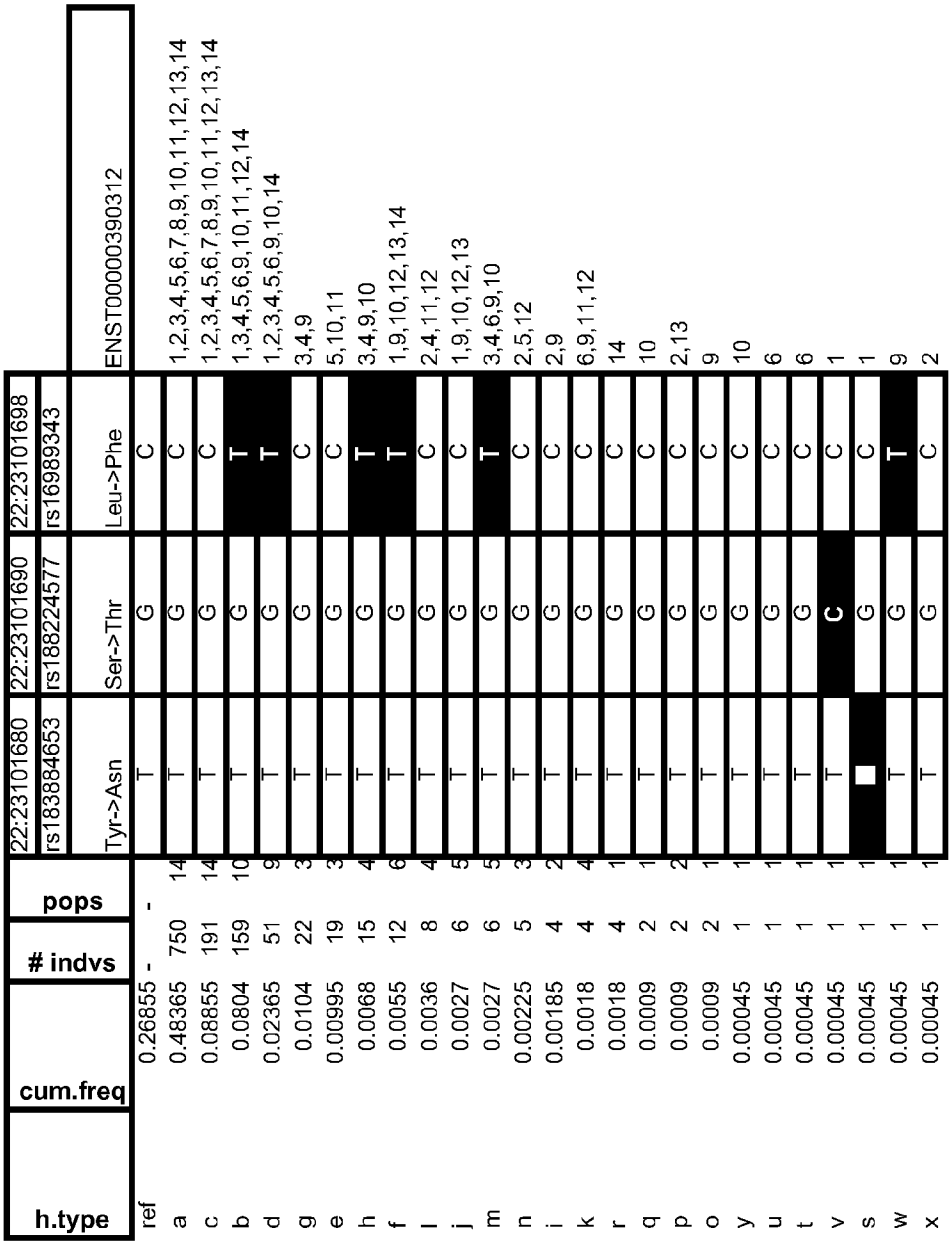
h.type	cum.freq	# indvs	pops	22:23135047 rs185062160 Ser->Asn	22:23135342 rs187815581 Met->Ile	22:23135358 rs193000225 Lys->Gln	22:23135360 rs185237664 Lys->Asn	22:23135376 rs189822640 Pro->Ser	22:23135379 rs181371196 Asp->Asn	22:23135382 rs184627661 Arg->Cys	22:23135466 rs142176771 Cys->Ser	ENST00000390314
ref	0.9745	-	-	G	G	A	G	C	G	C	T	
a	0.00775	17	3	G	G	A	G	C	G	C	■	3,4,9
b	0.00455	10	7	G	G	A	T	C	G	C	T	2,6,7,10,11,12,13
d	0.0037	8	5	G	G	A	G	T	■	C	T	1,5,10,11,14
c	0.0032	7	5	G	G	A	T	T	■	C	T	2,3,9,12,14
g	0.00225	5	3	G	G	A	G	C	■	C	T	2,10,14
j	0.0009	2	2	G	G	A	T	C	G	C	■	3,9
e	0.0009	2	1	■	G	A	G	C	G	C	T	6
f	0.0009	2	2	G	G	C	G	C	G	C	T	6,13
i	0.0009	2	2	G	G	A	G	C	G	T	T	2,11
h	0.00045	1	1	G	C	A	G	C	G	C	T	14

Table 13.118 - IGLV2-14



h.type	cum.freq	# indvs	pops	22:23101248 rs4822304 Leu->Phe	22:23101255 rs183579067 Thr->Asn	22:23101480 rs188164895 Ser->Thr	22:23101494 rs191388769 Gly->Ser	22:23101497 rs182533580 Tyr->Asn	22:23101559 rs4134484 Glu->Asp	22:23101568 rs28561583 Asn->Lys	22:23101600 rs184972365 Gly->Val	22:23101615 rs188991923 Asn->Ser
ref	0.26855	-	-	C	C	G	G	T	G	T	G	A
a	0.48365	750	14	T	C	G	G	T	T	T	G	A
c	0.08855	191	14	C	C	G	G	T	T	T	G	A
b	0.0804	159	10	C	C	G	G	T	T	G	G	A
d	0.02365	51	9	C	C	G	G	T	T	T	G	A
g	0.0104	22	3	C	■	G	G	T	G	T	G	A
e	0.00995	19	3	C	C	G	G	T	T	G	G	A
h	0.0068	15	4	C	C	G	G	T	G	G	G	A
f	0.0055	12	6	C	C	G	G	T	G	T	G	A
l	0.0036	8	4	T	C	G	G	T	T	G	G	A
j	0.0027	6	5	T	C	G	G	T	G	T	G	A
m	0.0027	6	5	T	C	G	G	T	T	T	G	A
n	0.00225	5	3	C	C	G	G	T	G	G	G	A
i	0.00185	4	2	T	C	G	■	T	T	T	G	A
k	0.0018	4	4	T	C	G	G	T	G	G	G	A
r	0.0018	4	1	T	■	G	G	T	T	T	G	A
q	0.0009	2	1	C	C	G	G	T	G	T	G	G
p	0.0009	2	2	C	■	G	G	T	T	T	G	A
o	0.0009	2	1	C	■	G	■	T	G	T	G	A
y	0.00045	1	1	T	C	G	G	T	T	G	T	A
u	0.00045	1	1	C	C	C	G	T	G	T	G	A
t	0.00045	1	1	C	C	G	■	T	G	T	G	A
v	0.00045	1	1	C	C	G	G	T	G	T	G	A
s	0.00045	1	1	C	C	G	G	T	G	T	G	A
w	0.00045	1	1	C	C	G	■	T	T	G	G	A
x	0.00045	1	1	T	C	G	G	■	T	T	G	A

Table 13.118 - IGLV2-14



h.type	cum.freq	# indvs	pops	22:23101680 rs183884653 Tyr->Asn	22:23101690 rs188224577 Ser->Thr	22:23101698 rs169893343 Leu->Phe	ENST00000390312
ref	0.26855	-	-	T	G	C	
a	0.48365	750	14	T	G	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.08855	191	14	T	G	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0804	159	10	T	G	T	1,3,4,5,6,9,10,11,12,14
d	0.02365	51	9	T	G	T	1,2,3,4,5,6,9,10,14
g	0.0104	22	3	T	G	C	3,4,9
e	0.00995	19	3	T	G	C	5,10,11
h	0.0068	15	4	T	G	T	3,4,9,10
f	0.0055	12	6	T	G	T	1,9,10,12,13,14
l	0.0036	8	4	T	G	C	2,4,11,12
j	0.0027	6	5	T	G	C	1,9,10,12,13
m	0.0027	6	5	T	G	T	3,4,6,9,10
n	0.00225	5	3	T	G	C	2,5,12
i	0.00185	4	2	T	G	C	2,9
k	0.0018	4	4	T	G	C	6,9,11,12
r	0.0018	4	1	T	G	C	14
q	0.0009	2	1	T	G	C	10
p	0.0009	2	2	T	G	C	2,13
o	0.0009	2	1	T	G	C	9
y	0.00045	1	1	T	G	C	10
u	0.00045	1	1	T	G	C	6
t	0.00045	1	1	T	G	C	6
v	0.00045	1	1	T	C	C	1
s	0.00045	1	1	[redacted]	G	C	1
w	0.00045	1	1	T	G	T	9
x	0.00045	1	1	T	G	C	2

Table 13.119 - IGLV2-18

h.type	cum.freq	# indvs	pops	22:23077436 rs149656926 Glu->Asp	22:23077555 rs4822296 Leu->Ser	ENST00000390310
ref	0.2279	-	-	G	T	
a	0.763	998	14	G	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0091	19	6	T	C	1,3,4,7,10,14

Table 13.120 - IGLV2-23

h.type	cum.freq	# indvs	pops	22:23040461 rs12166262 Thr->Asn	22:23040637 rs192197296 Val->Met	22:23040638 rs185785000 Val->Glu	22:23040685 rs190209511 Ser->Cys	22:23040697 rs193192358 Gly->Arg	22:23040700 rs61749282 Ser->Arg	22:23040724 rs188648111 Gln->Glu	22:23040765 rs61732325 Glu->Asp	22:23040767 rs5751516 Gly->Val
ref	0.58265	-	-	C	G	T	A	G	A	C	G	G
a	0.34655	593	14	C	G	T	A	G	A	C	G	T
b	0.0405	82	12	■	G	T	A	G	A	C	T	T
c	0.01545	31	5	C	G	T	A	G	A	C	G	T
d	0.00315	7	5	C	G	■	A	G	A	C	G	G
g	0.00315	7	5	C	G	T	A	G	A	C	T	T
f	0.00225	5	2	■	G	T	A	G	A	C	T	G
e	0.00135	3	3	C	G	T	A	G	A	C	G	G
k	0.0009	2	2	■	G	T	A	G	A	C	G	G
o	0.0009	2	2	C	G	T	A	G	A	C	T	G
j	0.00045	1	1	C	G	T	T	G	A	C	G	G
h	0.00045	1	1	C	■	T	A	G	A	C	G	G
i	0.00045	1	1	C	G	T	A	■	A	C	G	G
n	0.00045	1	1	C	G	T	A	G	A	G	G	G
m	0.00045	1	1	C	■	T	A	G	A	C	G	T
l	0.00045	1	1	C	G	T	A	G	C	C	G	G
p	0.00045	1	1	C	G	T	A	G	A	C	T	G

Table 13.120 - IGLV2-23

h.type	cum.freq	# indvs	pops	22:23040887 rs5759376 Tyr->Cys	ENST00000390306
ref	0.58265	-	-	A	
a	0.34655	593	14	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.0405	82	12	A	1,2,3,4,5,7,9,10,11,12,13,14
c	0.01545	31	5	**G**	3,4,5,9,11
d	0.00315	7	5	A	2,5,6,7,12
g	0.00315	7	5	A	1,2,9,10,14
f	0.00225	5	2	A	10,13
e	0.00135	3	3	**G**	3,9,11
k	0.0009	2	2	A	8,10
o	0.0009	2	2	A	3,10
j	0.00045	1	1	A	9
h	0.00045	1	1	A	3
i	0.00045	1	1	A	9
n	0.00045	1	1	A	14
m	0.00045	1	1	A	13
l	0.00045	1	1	A	5
p	0.00045	1	1	**G**	11

Table 13.121 - IGLV2-33

h.type	cum.freq	# indvs	pops	22:22930684 rs11914178 Met->Val	22:22930896 rs143638085 Gly->Arg	22:22931059 rs185353299 Gly->Val	ENST00000390302
ref	0.98955	-	-	A	G	G	
a	0.00775	16	4	**G**	G	G	10,11,12,14
b	0.0018	4	3	A	G	**T**	2,5,13
c	0.0009	2	1	A	■	G	14

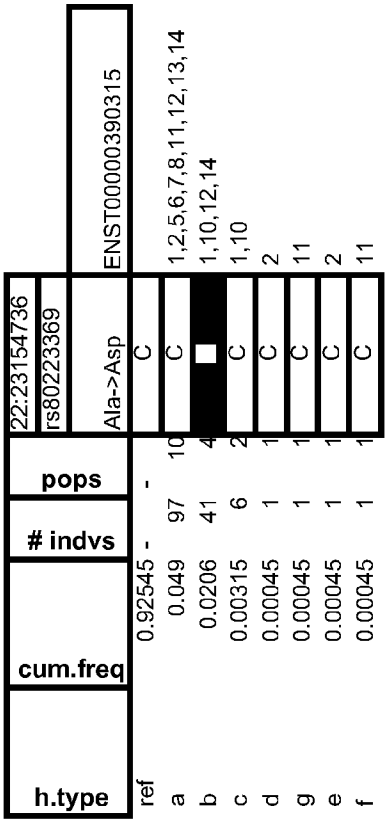
Table 13.122 - IGLV2-8

h.type	cum.freq	# indvs	pops	22:23165340 rs5996397 Thr->Ser	22:23165526 rs191286893 Pro->Ser	22:23165610 rs1850026152 Pro->Ser	22:23165680 rs149347098 Ser->Phe	22:23165701 rs144694239 Thr->Met	22:23165757 rs11558656 Ser->Cys	ENST00000390317
ref	0.8422	-	-	C	C	C	C	C	A	
a	0.1451	267	11	**G**	C	C	C	C	A	1,2,5,6,7,8,10,11,12,13,14
b	0.0046	10	7	C	C	C	**T**	C	A	2,5,6,7,11,12,13
c	0.00315	7	3	C	**T**	C	C	C	A	3,4,9
d	0.0027	6	3	**G**	C	C	C	C	**T**	1,10,14
e	0.0018	4	2	C	C	C	C	**T**	A	4,7
f	0.00045	1	1	C	C	**T**	C	C	A	2

Table 13.123 - IGLV3-10

h.type	cum.freq	# indvs	pops	22:23154549 rs182282578 Thr->Ala	22:23154570 rs192547759 Pro->Thr	22:23154580 rs183865794 Tyr->Cys	22:23154582 rs1065464 Ala->Pro	22:23154583 rs188439978 Ala->Val	22:23154592 rs184893196 Tyr->Cys	22:23154646 rs141938356 Arg->Gln	22:23154659 rs188537134 Ile->Met	22:23154667 rs191450268 Arg->Lys
ref	0.92545	-	-	A	C	A	G	C	A	G	C	G
a	0.049	97	10	A	C	A	**C**	C	A	G	C	G
b	0.0206	41	4	A	C	A	G	C	A	G	C	G
c	0.00315	6	2	A	C	A	G	C	A	■	C	G
d	0.00045	1	1	A	C	**G**	G	C	A	G	C	G
g	0.00045	1	1	A	■	A	G	C	A	G	**G**	G
e	0.00045	1	1	**G**	C	A	G	C	**G**	G	C	■
f	0.00045	1	1	A	C	A	G	**T**	A	G	C	G

Table 13.123 - IGLV3-10



h.type	cum.freq	# indvs	pops	22:23154736 rs80223369 Ala->Asp	ENST00000390315
ref	0.92545	-	-	C	
a	0.049	97	10	C	1,2,5,6,7,8,11,12,13,14
b	0.0206	41	4	[redacted]	1,10,12,14
c	0.00315	6	2	C	1,10
d	0.00045	1	1	C	2
g	0.00045	1	1	C	11
e	0.00045	1	1	C	2
f	0.00045	1	1	C	11

Table 13.124 - IGLV3-12

h.type	cum.freq	# indvs	pops	22:23114327 rs6003299 Met->Thr	22:23114350 rs5996382 Ser->Gly	22:23114938 rs2073450 Ser->Arg	22:23114993 rs2073451 Ala->Thr	22:23115059 rs190626294 Ser->Gly	ENST00000390313
ref	0.42115	-	-	T	A	C	G	A	
a	0.38715	653	14	C	G	C	■	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.1116	220	14	C	G	C	G	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0774	158	14	C	G	■	G	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.00135	3	3	T	A	C	■	A	2,4,14
g	0.00045	1	1	C	G	C	■	G	1
e	0.00045	1	1	C	A	C	G	A	4
f	0.00045	1	1	C	A	C	■	A	10

Table 13.125 - IGLV3-16

h.type	cum.fre	# indvs	pops	22:23090155 rs191152677 Ser->Thr	22:23090222 rs186816034 Tyr->Phe	22:23090242 rs11090180 Phe->Val	22:23090243 rs11090181 Phe->Ser	22:23090248 rs191325224 Val->Met	22:23090260 rs149984128 Tyr->Asn	22:23090327 rs145232556 Val->Asp	22:23090381 rs149168687 Leu->Gln	ENST00000390311
ref	0.9395	-	-	T	A	T	T	G	T	T	T	
a	0.04295	82	6	T	A	**G**	**C**	G	T	T	T	1,5,10,11,12,14
b	0.0081	18	9	T	A	T	T	G	T	T	**N**	1,2,5,7,9,10,12,13,14
c	0.0045	10	5	T	A	T	T	G	**N**	T	T	1,9,10,11,14
h	0.0018	4	3	T	A	**G**	**C**	G	T	T	**N**	1,10,14
d	0.00135	3	3	T	A	T	T	G	T	**N**	T	2,5,9
g	0.00045	1	1	T	A	T	T	**N**	T	T	T	5
e	0.00045	1	1	T	**T**	T	T	G	**N**	T	T	10
f	0.00045	1	1	**N**	A	T	T	G	T	T	T	13
i	0.00045	1	1	T	**T**	**G**	**C**	G	T	T	T	10

Table 13.126 - IGLV3-19

h.type	cum.freq	# indvs	pops	22:23063178 rs12158984 Leu->Phe	22:23063348 rs117660534 Val->Ile	22:23063367 rs144678284 Gln->Pro	22:23063406 rs193091235 Arg->Thr	22:23063411 rs144208459 Thr->Ala	22:23063421 rs191985323 Gly->Glu	22:23063436 rs117237784 Ser->Asn	22:23063448 rs76281239 Ser->Asn	22:23063480 rs188492127 Val->Ile
ref	0.86895	-	-	C	G	A	G	A	G	G	G	G
a	0.10015	190	12	**T**	G	A	G	A	G	G	G	G
b	0.01785	38	7	C	G	A	G	A	G	G	G	G
c	0.00225	5	2	C	G	A	G	A	**-**	G	G	G
d	0.00135	3	1	C	G	A	G	A	G	G	G	G
l	0.00135	3	2	C	G	A	G	A	G	G	**-**	G
j	0.0009	2	2	C	**-**	A	G	A	G	G	G	G
k	0.0009	2	2	C	G	A	G	A	G	**-**	G	G
h	0.0009	2	1	C	G	A	G	A	G	G	G	**-**
i	0.0009	2	1	C	G	A	G	A	G	G	**-**	**-**
p	0.0009	2	2	C	G	A	G	A	G	**-**	G	**-**
g	0.00045	1	1	C	G	A	G	A	G	**-**	**-**	G
e	0.00045	1	1	C	G	A	G	A	G	G	G	G
q	0.00045	1	1	C	G	**C**	G	A	G	G	G	G
r	0.00045	1	1	C	G	A	G	A	G	G	G	G
f	0.00045	1	1	C	G	A	G	A	G	G	G	G
n	0.00045	1	1	C	G	A	G	A	G	G	G	G
m	0.00045	1	1	C	G	A	G	**G**	G	**-**	**-**	G
o	0.00045	1	1	C	G	A	**C**	A	G	G	G	G

Table 13.126 - IGLV3-19

h.type	cum.freq	# indvs	pops	22:23063484 rs79338593 Leu->Arg	22:23063531 rs181016062 Phe->Ile	22:23063577 rs184342792 Gly->Glu	22:23063612 rs182162941 Asn->Asp	22:23063618 rs77723964 Arg->Trp	ENST00000390309
ref	0.86895	-	-	T	T	G	A	C	
a	0.10015	190	12	T	T	G	A	C	1,2,3,4,5,6,9,10,11,12,13,14
b	0.01785	38	7	G	T	G	A	T	1,3,4,9,10,12,14
c	0.00225	5	2	T	T	G	A	C	1,14
d	0.00135	3	1	T	T	G	A	T	10
l	0.00135	3	2	T	T	G	A	C	10,14
j	0.0009	2	2	T	T	G	A	C	2,14
k	0.0009	2	2	T	T	G	A	C	4,13
h	0.0009	2	1	T	T	G	A	C	9
i	0.0009	2	1	T	T	G	A	C	10
p	0.0009	2	2	T	T	G	A	C	2,6
g	0.00045	1	1	G	T	G	A	T	9
e	0.00045	1	1	T	T	-	A	C	4
q	0.00045	1	1	T	T	G	A	C	14
r	0.00045	1	1	G	T	G	A	C	4
f	0.00045	1	1	T	T	G	G	C	4
n	0.00045	1	1	T	-	G	A	C	13
m	0.00045	1	1	T	T	G	A	C	13
o	0.00045	1	1	T	T	G	A	C	10

Table 13.127 - IGLV3-1

h.type	cum.fre	# indvs	pops	22:23223277 rs144850905 Gly->AlaGly->Ala	22:23223280 rs184648956 Ser->Cys	22:23223283 rs61736464 Val->Glu	22:23223292 rs145382269 Tyr->Ser	22:23223297 rs181792449 Leu->Val	22:23223301 rs186574328 Thr->Ile	22:23223306 rs61736463 Pro->Ala	22:23223330 rs145228815 Gly->Arg
ref	0.9255	-	-	G	C	T	A	C	C	C	G
a	0.01865	40	13	G	C	T	A	C	C	C	G
c	0.00585	13	8	G	C	T	A	C	C	C	G
b	0.005	11	8	G	C	T	A	C	C	C	G
h	0.00315	7	7	G	C	T	A	C	C	C	G
o	0.00315	7	4	G	C	T	A	C	C	C	G
e	0.0027	6	5	G	C	T	**C**	C	C	C	G
t	0.00225	5	5	G	C	T	A	C	C	C	G
v	0.00225	5	5	G	C	T	A	C	C	C	G
f	0.00225	5	4	G	C	T	A	**G**	C	C	G
i	0.0018	4	4	G	C	T	A	C	C	C	G
n	0.0018	4	4	G	C	T	A	C	C	C	G
ag	0.00135	3	3	G	C	T	A	C	C	C	G
y	0.00135	3	3	G	C	T	A	C	C	C	G
s	0.00135	3	3	G	C	T	A	C	C	**G**	G
x	0.00135	3	3	G	C	T	A	C	**T**	C	G
ad	0.00135	3	2	G	C	T	A	C	C	C	G
l	0.00135	3	3	G	C	T	A	C	C	C	G
p	0.00135	3	2	G	C	T	A	C	C	C	G
ah	0.0009	2	2	G	C	T	A	C	C	C	G
d	0.0009	2	2	G	C	T	A	C	C	C	G
j	0.0009	2	1	G	C	T	A	C	C	C	G
g	0.0009	2	2	G	C	T	A	C	C	C	G
z	0.0009	2	2	G	C	T	A	C	C	C	G
ac	0.0009	2	2	G	C	T	A	C	C	C	G
ae	0.0009	2	2	G	C	T	A	C	C	C	G
u	0.00045	1	1	G	C	T	A	**G**	C	C	G
k	0.00045	1	1	G	C	T	A	C	C	C	G
aj	0.00045	1	1	G	C	T	A	C	C	C	**C**

Table 13.127 - IGLV3-1

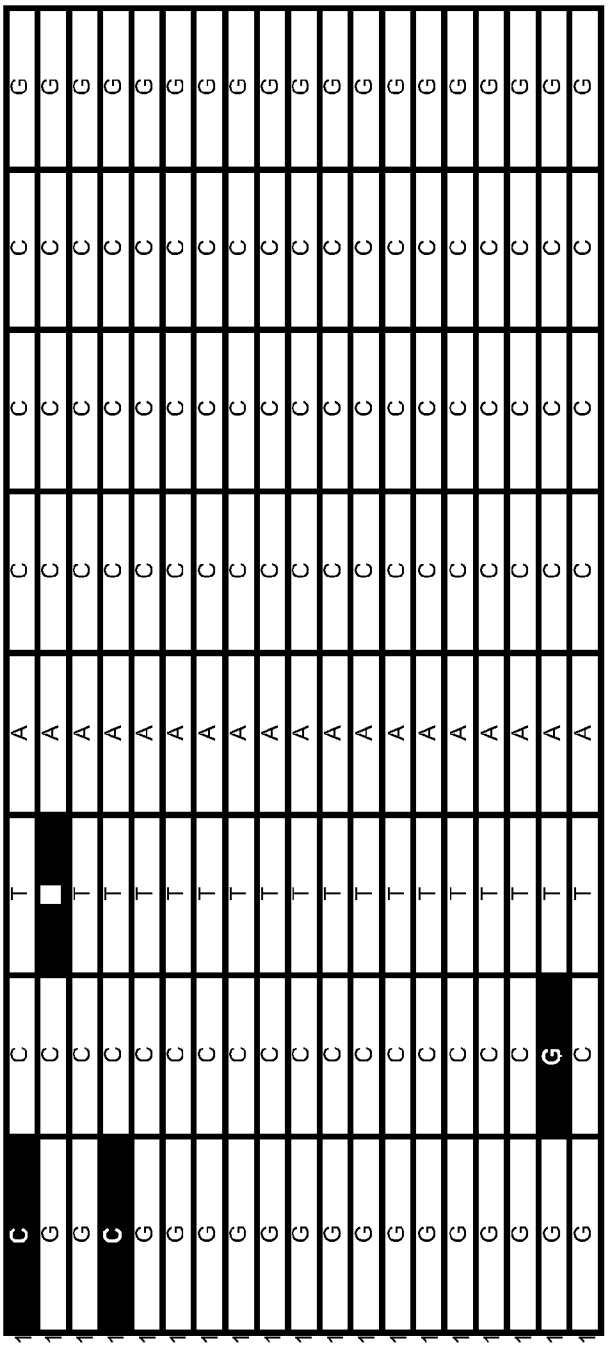


Table 13.127 - IGLV3-1

h.type	cum.fre	# indvs	pops	22:23223348 rs189772470 Thr->Pro	22:23223355 rs183495921 Ser->Cys	22:23223363 rs187892119 Lys->Glu	22:23223364 rs191755097 Lys->Ile	22:23223369 rs188197731 Gly->Arg	22:23223375 rs192957059 Lys->Glu	22:23223376 rs61731694 Lys->Thr	22:23223382 rs138666033 Ala->Val	22:23223385 rs61731693 Cys->Tyr
ref	0.9255	-	-	A	C	A	A	G	A	A	C	G
a	0.01865	40	13	A	C	A	A	G	A	A	C	G
c	0.00585	13	8	A	C	A	A	G	A	A	**T**	G
b	0.005	11	8	A	C	A	A	G	A	A	C	**N**
h	0.00315	7	7	A	C	A	A	G	A	A	C	G
o	0.00315	7	4	A	C	A	A	G	A	A	C	G
e	0.0027	6	5	A	C	A	A	G	A	A	C	G
t	0.00225	5	5	A	C	A	A	G	A	A	C	G
v	0.00225	5	5	A	C	A	A	G	A	A	C	G
f	0.00225	5	4	A	C	A	A	G	A	A	C	G
i	0.0018	4	4	**C**	C	A	A	G	A	A	C	G
n	0.0018	4	4	A	C	A	A	G	A	A	C	G
ag	0.00135	3	3	A	C	A	A	G	A	A	C	G
y	0.00135	3	3	A	C	A	A	G	A	**C**	C	G
s	0.00135	3	3	A	C	A	A	G	A	A	C	G
x	0.00135	3	3	A	C	A	A	G	A	A	C	G
ad	0.00135	3	2	A	C	A	A	G	A	A	C	G
l	0.00135	3	3	A	C	A	A	G	A	A	C	**N**
p	0.00135	3	2	A	C	A	A	G	A	A	C	G
ah	0.0009	2	2	A	C	**G**	A	G	A	A	C	G
d	0.0009	2	2	A	C	A	A	G	A	A	C	G
j	0.0009	2	1	A	C	A	A	G	A	A	C	G
g	0.0009	2	2	A	C	A	A	G	A	A	C	G
z	0.0009	2	2	A	C	A	A	G	A	A	C	G
ac	0.0009	2	2	A	C	A	A	G	A	A	C	G
ae	0.0009	2	2	A	C	A	A	G	A	A	C	G
u	0.00045	1	1	A	C	A	A	G	A	A	C	G
k	0.00045	1	1	A	C	A	A	G	**G**	A	C	G
aj	0.00045	1	1	A	C	A	A	G	A	A	C	G

Table 13.127 - IGLV3-1

	1	A	C	A	A	G	A	A	C	G		
aa	0.00045	1	A	C	A	A	G	A	A	C	G	
aq	0.00045	1	A	C	A	A	G	A	A	C	G	
al	0.00045	1	A	C	A	A	G	A	A	C	G	
ab	0.00045	1	A	C	A	A	G	**G**	A	C	G	
ak	0.00045	1	A	C	A	A	G	A	A	C	G	
q	0.00045	1	**C**	C	A	A	G	A	A	**T**	G	
ap	0.00045	1	A	C	A	A	G	**G**	A	C	G	
w	0.00045	1	A	**G**	A	A	G	A	A	C	G	
ao	0.00045	1	A	C	A	A	G	A	A	C	G	
r	0.00045	1	A	C	A	**T**	G	A	A	C	G	
af	0.00045	1	A	C	A	A	G	A	A	C	G	
au	0.00045	1	A	C	A	A	G	A	A	**T**	G	
am	0.00045	1	A	C	A	A	G	A	A	C	G	
at	0.00045	1	A	C	A	**T**	G	A	A	C	G	
m	0.00045	1	A	C	A	A	G	**G**	A	**T**	G	
ar	0.00045	1	A	C	A	A	G	A	A	C	G	
as	0.00045	1	A	C	A	A	■	A	A	C	G	
ai	0.00045	1	A	C	A	A	G	A	A	C	G	
an	0.00045	1	A	C	A	A	G	A	A	C	G	

Table 13.127 - IGLV3-1

h.type	cum.fre	# indvs	pops	22:23223386 rs61735526 Cys->Trp	22:23223387 rs186777494 Trp->Arg	22:23223400 rs61731688 Lys->Arg	22:23223402 rs61731687 Pro->Ala	22:23223417 rs190667156 Val->Met	22:23223432 rs73164939 Gln->Glu	22:23223438 rs143830449 Ser->Gly	22:23223439 rs61731685 Ser->Asn	22:23223525 rs61731681 Met->Leu
ref	0.9255	-	-	C	T	A	C	G	C	A	G	A
a	0.01865	40	13	C	T	A	C	G	C	A	■	A
c	0.00585	13	8	C	T	A	C	G	C	A	G	A
b	0.005	11	8	C	T	A	C	G	C	A	G	A
h	0.00315	7	7	C	T	A	C	■	C	A	G	A
o	0.00315	7	4	C	T	G	C	G	C	A	G	A
e	0.0027	6	5	C	T	A	C	G	C	A	G	A
t	0.00225	5	5	C	T	A	G	G	C	A	G	A
v	0.00225	5	5	C	T	A	C	G	G	A	G	A
f	0.00225	5	4	C	T	A	C	G	C	A	G	A
i	0.0018	4	4	C	T	A	C	G	C	A	G	A
n	0.0018	4	4	C	T	A	C	G	C	G	G	A
ag	0.00135	3	3	C	T	A	C	G	C	G	■	A
y	0.00135	3	3	C	T	A	C	G	C	A	G	A
s	0.00135	3	3	C	T	A	C	G	C	A	G	A
x	0.00135	3	3	C	T	A	C	G	C	A	G	A
ad	0.00135	3	2	C	T	A	C	G	C	A	G	A
l	0.00135	3	3	C	T	A	C	G	C	A	■	A
p	0.00135	3	2	C	T	A	C	G	C	A	G	A
ah	0.0009	2	2	C	T	A	C	G	C	A	G	A
d	0.0009	2	2	C	T	A	C	G	C	A	G	C
j	0.0009	2	1	C	T	A	C	G	C	A	G	A
g	0.0009	2	2	G	T	A	C	G	C	A	■	A
z	0.0009	2	2	C	T	A	C	G	C	A	G	A
ac	0.0009	2	2	G	T	A	C	G	C	A	G	A
ae	0.0009	2	2	C	T	A	C	G	C	A	G	A
u	0.00045	1	1	C	T	A	C	G	C	G	G	A
k	0.00045	1	1	C	T	G	C	G	C	A	G	A
aj	0.00045	1	1	C	T	A	C	G	C	A	G	A

Table 13.127 - IGLV3-1

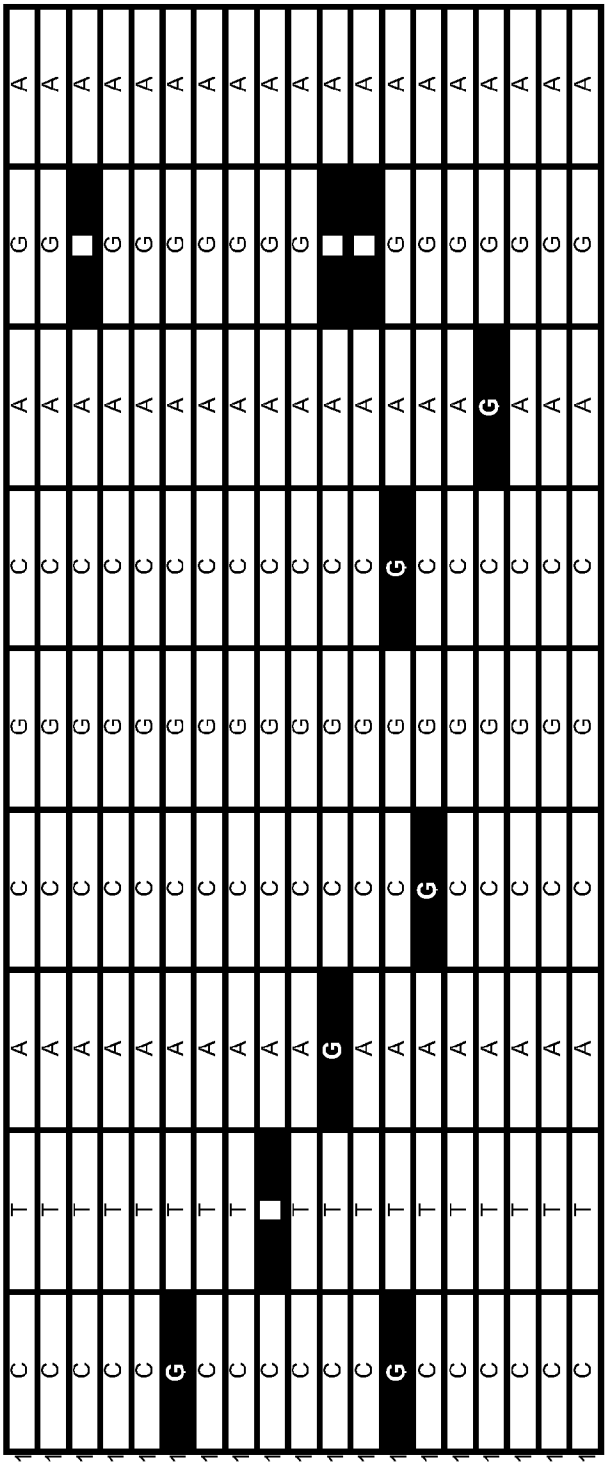


Table 13.127 - IGLV3-1

h.type	cum.fre	# indvs	pops	22:23223531 rs190043877 Glu->Lys	22:23223535 rs145774047 Ala->Gly	22:23223540 rs180853064 Tyr->Asp	22:23223541 rs142180628 Tyr->Ser	22:23223544 rs185803370 Tyr->Ser	22:23223550 rs192317432 Gln->Pro	22:23223573 rs145570740 His->Tyr	ENST00000390319
ref	0.9255	-	-	G	C	T	A	A	A	C	
a	0.01865	40	13	G	C	T	A	A	A	C	2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.00585	13	8	G	C	T	A	A	A	C	2,5,6,7,9,10,13,14
b	0.005	11	8	G	C	T	A	A	A	C	1,4,5,6,10,11,13,14
h	0.00315	7	7	G	C	T	A	A	A	C	1,3,4,5,9,10,14
o	0.00315	7	4	G	C	T	A	A	A	C	3,7,10,14
e	0.0027	6	5	G	C	T	A	A	A	C	1,2,3,10,11
t	0.00225	5	5	G	C	T	A	A	A	C	2,3,5,10,14
v	0.00225	5	5	G	C	T	A	A	A	C	1,5,6,7,9
f	0.00225	5	4	G	C	T	A	A	A	C	2,4,10,13
i	0.0018	4	4	G	C	T	A	A	A	C	1,7,10,11
n	0.0018	4	4	G	C	T	A	A	A	C	6,12,13,14
ag	0.00135	3	3	G	C	T	A	A	A	C	2,5,13
y	0.00135	3	3	G	C	T	A	A	A	C	5,10,12
s	0.00135	3	3	G	C	T	A	A	A	C	2,9,10
x	0.00135	3	3	G	C	T	A	A	A	C	3,10,14
ad	0.00135	3	2	G	C	T	A	A	A	**T**	2,13
l	0.00135	3	3	G	C	T	A	A	A	C	2,6,13
p	0.00135	3	2	G	C	T	A	**C**	A	C	5,9
ah	0.0009	2	2	G	C	T	A	A	A	C	9,14
d	0.0009	2	2	G	C	T	A	A	A	C	2,13
j	0.0009	2	1	G	C	T	**C**	A	A	C	2
g	0.0009	2	2	G	C	T	A	A	A	C	6,12
z	0.0009	2	2	G	C	T	A	A	**C**	C	1,4
ac	0.0009	2	2	G	C	T	A	A	A	C	6,11
ae	0.0009	2	2	G	**G**	T	A	A	A	C	13,14
u	0.00045	1	1	G	C	T	A	A	A	C	14
k	0.00045	1	1	G	C	T	A	A	A	C	11
aj	0.00045	1	1	G	C	T	A	A	A	C	2

Table 13.127 - IGLV3-1

aa	0.00045	1	1	G	C	T	A	A	A	C	5
aq	0.00045	1	1	G	C	T	A	A	A	C	13
al	0.00045	1	1	G	C	T	**C**	A	A	C	10
ab	0.00045	1	1	G	C	T	A	A	A	C	5
ak	0.00045	1	1	G	C	**G**	A	A	A	C	13
q	0.00045	1	1	G	C	T	A	A	A	C	2
ap	0.00045	1	1	G	C	T	A	A	A	C	5
w	0.00045	1	1	G	C	T	A	A	A	C	14
ao	0.00045	1	1	G	C	T	A	A	A	C	3
r	0.00045	1	1	G	C	T	A	A	A	C	2
af	0.00045	1	1	G	C	T	A	A	A	C	3
au	0.00045	1	1	G	C	T	A	A	A	C	1
am	0.00045	1	1	G	C	T	A	A	A	C	10
at	0.00045	1	1	G	C	T	A	A	A	C	10
m	0.00045	1	1	G	C	T	A	A	A	C	4
ar	0.00045	1	1	G	C	T	**C**	A	A	C	13
as	0.00045	1	1	G	C	T	A	A	A	C	1
ai	0.00045	1	1	G	C	T	A	A	A	C	2
an	0.00045	1	1	■	C	T	A	A	A	C	1

Table 13.128 - IGLV3-21

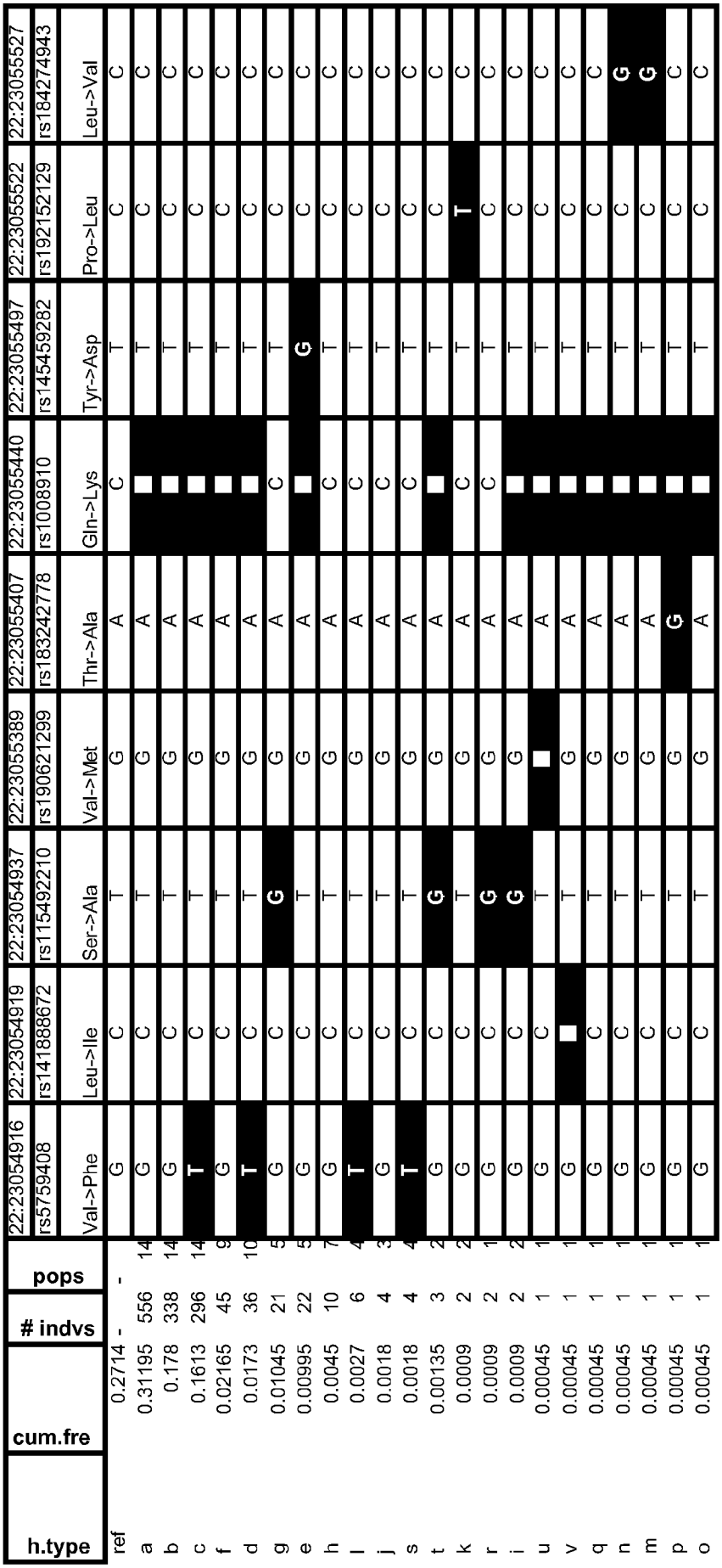
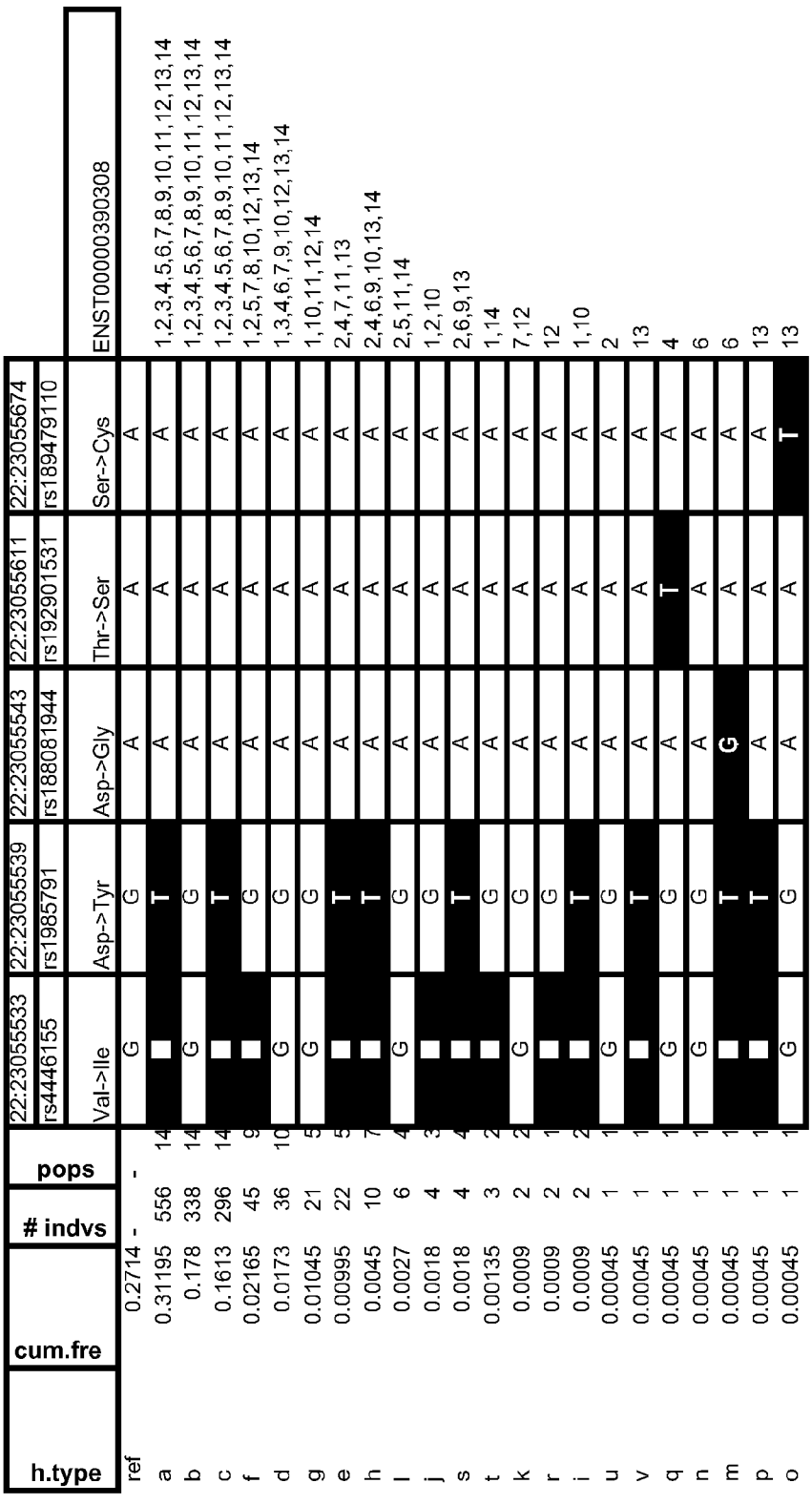


Table 13.128 - IGLV3-21



h.type	cum.fre	# indvs	pops	22:23055533 rs4446155 Val->Ile	22:23055539 rs1985791 Asp->Tyr	22:23055543 rs188081944 Asp->Gly	22:23055611 rs192901531 Thr->Ser	22:23055674 rs189479110 Ser->Cys	ENST00000390308
ref	0.2714	-	-	G	G	A	A	A	
a	0.31195	556	14		T	A	A	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.178	338	14	G	G	A	A	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.1613	296	14		T	A	A	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
f	0.02165	45	9		G	A	A	A	1,2,5,7,8,10,12,13,14
d	0.0173	36	10	G	G	A	A	A	1,3,4,6,7,9,10,12,13,14
g	0.01045	21	5	G	G	A	A	A	1,10,11,12,14
e	0.00995	22	5		T	A	A	A	2,4,7,11,13
h	0.0045	10	7		T	A	A	A	2,4,6,9,10,13,14
l	0.0027	6	4	G	G	A	A	A	2,5,11,14
j	0.0018	4	3		G	A	A	A	1,2,10
s	0.0018	4	4		T	A	A	A	2,6,9,13
t	0.00135	3	2		G	A	A	A	1,14
k	0.0009	2	2	G	G	A	A	A	7,12
r	0.0009	2	1		G	A	A	A	12
i	0.0009	2	2		T	A	A	A	1,10
u	0.00045	1	1	G	G	A	A	A	2
v	0.00045	1	1		T	A	A	A	13
q	0.00045	1	1	G	G	A	T	A	4
n	0.00045	1	1	G	G	A	A	A	6
m	0.00045	1	1		T	G	A	A	6
p	0.00045	1	1		T	A	A	A	13
o	0.00045	1	1	G	G	A	A	T	13

Table 13.129 - IGLV3-22

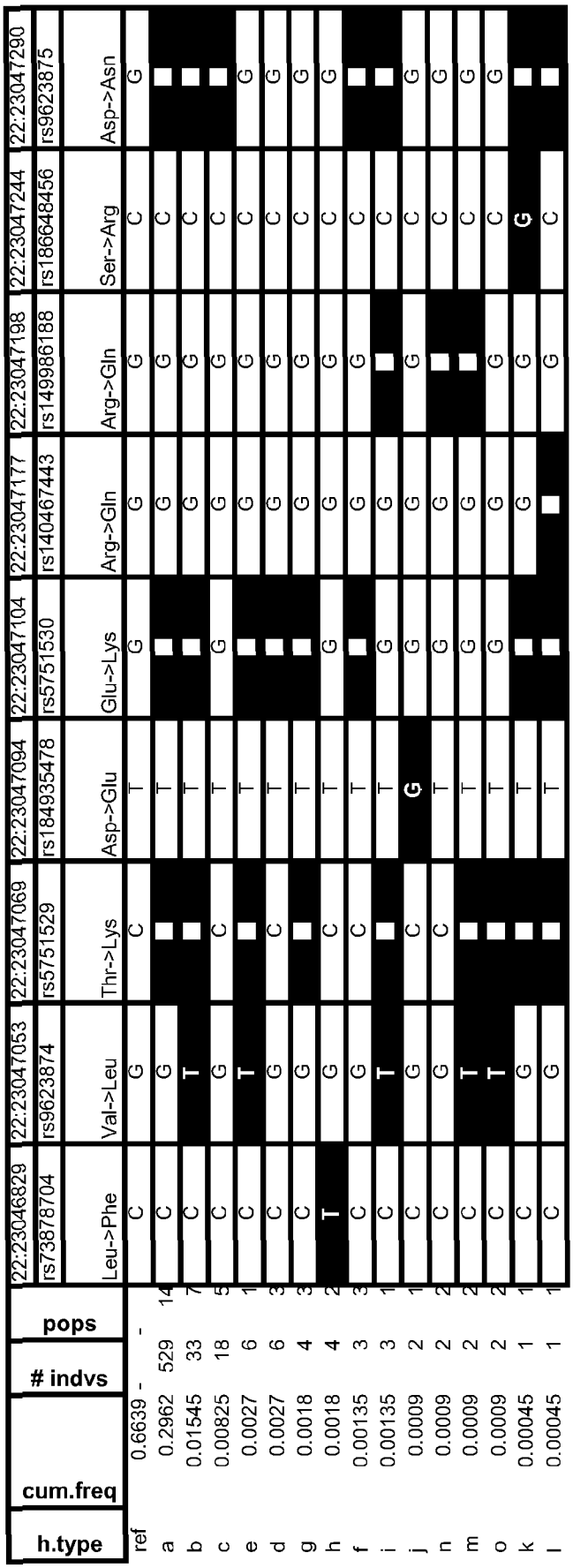


Table 13.129 - IGLV3-22

h.type	cum.freq	# indvs	pops	ENST00000390307
ref	0.6639	-	-	
a	0.2962	529	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.01545	33	7	1,3,4,9,10,12,14
c	0.00825	18	5	3,4,6,9,14
e	0.0027	6	1	10
d	0.0027	6	3	1,10,14
g	0.0018	4	3	3,4,10
h	0.0018	4	2	10,14
f	0.00135	3	3	2,10,12
i	0.00135	3	1	10
j	0.0009	2	1	4
n	0.0009	2	2	10,12
m	0.0009	2	2	1,14
o	0.0009	2	2	10,14
k	0.00045	1	1	11
l	0.00045	1	1	14

Table 13.130 - IGLV3-25

h.type	cum.freq	# indvs	pops	22:23029258 rs12628782 Pro->His	22:23029263 rs443102 Phe->Leu	22:23029478 rs190375718 Pro->His	22:23029526 rs3205084 Gly->Ala	22:23029537 rs182727015 Pro->Ser	22:23029586 rs188154997 Val->Glu	22:23029588 rs5751507 Leu->Val	22:23029595 rs182651541 Ile->Lys	22:23029597 rs188941516 Tyr->His
ref	0.1086	-	-	C	T	C	G	C	T	C	T	T
a	0.6581	943	14	C	C	C	G	C	T	C	T	T
b	0.19965	367	14	C	C	C	**C**	C	T	C	T	T
c	0.02105	41	6	■	C	C	G	C	T	C	T	T
d	0.00405	9	5	C	C	C	G	C	T	**G**	T	T
e	0.00135	3	3	C	C	C	G	C	T	C	T	T
g	0.0009	2	1	C	C	■	G	C	T	C	T	T
l	0.0009	2	2	C	C	C	G	C	T	C	T	**C**
j	0.00045	1	1	C	C	C	G	C	T	C	T	T
k	0.00045	1	1	C	C	C	**C**	C	T	C	T	**C**
s	0.00045	1	1	C	T	C	G	**T**	T	**G**	T	T
q	0.00045	1	1	C	C	C	G	C	T	**G**	T	**C**
r	0.00045	1	1	C	C	C	**C**	C	T	C	T	T
h	0.00045	1	1	C	C	C	G	C	T	C	T	T
f	0.00045	1	1	■	C	C	G	C	T	C	T	T
i	0.00045	1	1	C	C	C	G	C	T	C	T	T
n	0.00045	1	1	C	C	C	**C**	C	T	C	T	T
m	0.00045	1	1	C	C	C	**C**	**T**	T	C	T	T
p	0.00045	1	1	C	T	C	G	C	■	C	T	T
o	0.00045	1	1	C	C	C	G	C	T	C	■	T

Table 13.130 - IGLV3-25

h.type	cum.freq	# indvs	pops	22:23029598 rs192490624 Tyr->Ser	22:23029607 rs145801212 Ser->Ile	22:23029658 rs184773494 Thr->Arg	22:23029682 rs189679814 Gly->Ala	22:23029690 rs181030260 Ala->Ser	22:23029719 rs189572281 Gln->His	22:23029723 rs181874336 Ala->Thr
ref	0.1086	-	-	A	G	C	G	G	A	G
a	0.6581	943	14	A	G	C	G	G	A	G
b	0.19965	367	14	A	G	C	G	G	A	G
c	0.02105	41	6	A	G	C	G	G	A	G
d	0.00405	9	5	A	G	C	G	G	A	G
e	0.00135	3	3	A	**T**	C	G	G	A	G
g	0.0009	2	1	A	G	C	G	G	A	G
l	0.0009	2	2	A	G	C	G	G	A	G
j	0.00045	1	1	A	G	**G**	G	G	A	G
k	0.00045	1	1	A	G	C	G	G	A	G
s	0.00045	1	1	A	G	C	G	G	A	G
q	0.00045	1	1	A	G	C	G	G	**C**	G
r	0.00045	1	1	A	G	C	G	G	A	■
h	0.00045	1	1	A	G	C	G	**T**	A	G
f	0.00045	1	1	A	**T**	C	G	G	A	G
i	0.00045	1	1	A	G	C	**C**	G	A	G
n	0.00045	1	1	A	G	C	**C**	G	A	G
m	0.00045	1	1	A	G	C	G	G	A	G
p	0.00045	1	1	**C**	G	C	G	G	A	G
o	0.00045	1	1	A	G	C	G	G	A	G

Table 13.130 - IGLV3-25

h.type	cum.freq	# indvs	pops	ENST00000390305
ref	0.1086	-	-	
a	0.6581	943	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.19965	367	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.02105	41	6	2,3,4,5,9,11
d	0.00405	9	5	1,2,5,6,10
e	0.00135	3	3	2,5,14
g	0.0009	2	1	2
l	0.0009	2	2	3,5
j	0.00045	1	1	6
k	0.00045	1	1	5
s	0.00045	1	1	10
q	0.00045	1	1	2
r	0.00045	1	1	7
h	0.00045	1	1	10
f	0.00045	1	1	4
i	0.00045	1	1	6
n	0.00045	1	1	14
m	0.00045	1	1	8
p	0.00045	1	1	10
o	0.00045	1	1	10

Table 13.131 - IGLV3-27

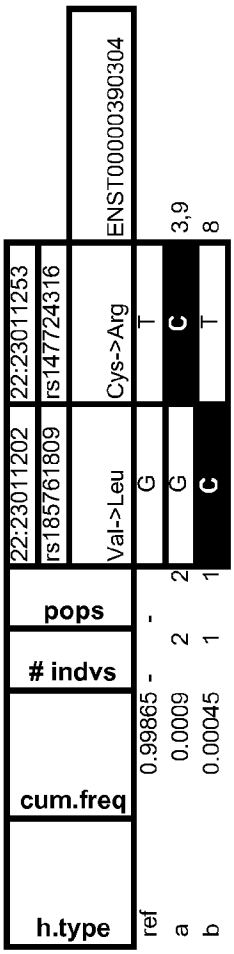


Table 13.132 - IGLV3-32

h.type	cum.freq	# indvs	pops	22:22937050 rs181933932 Pro->Leu	22:22937234 rs190288116 Ile->Val	22:22937306 rs144330202 Gly->Arg	22:22937351 rs182453787 Pro->Ala	22:22937394 rs147918941 Arg->Gln	ENST00000390303
ref	0.99455	-	-	C	A	G	C	G	
a	0.0032	7	3	C	A	G	C		3,4,9
c	0.0009	2	2	**T**	A	G	C	G	10,14
d	0.00045	1	1	C	A	G	**G**	G	3
e	0.00045	1	1	C	A		C	G	10
b	0.00045	1	1	C	**G**	G	C	G	11

Table 13.133 - IGLV3-9

h.type	cum.fre	# indvs	pops	22:23162047 rs187387271 Gly->Glu	22:23162116 rs111894401 Ser->Asn	ENST00000390316
ref	0.9982	-	-	G	G	1,13
a	0.0009	2	2	G	■	
b	0.0009	2	2	■	G	3,13

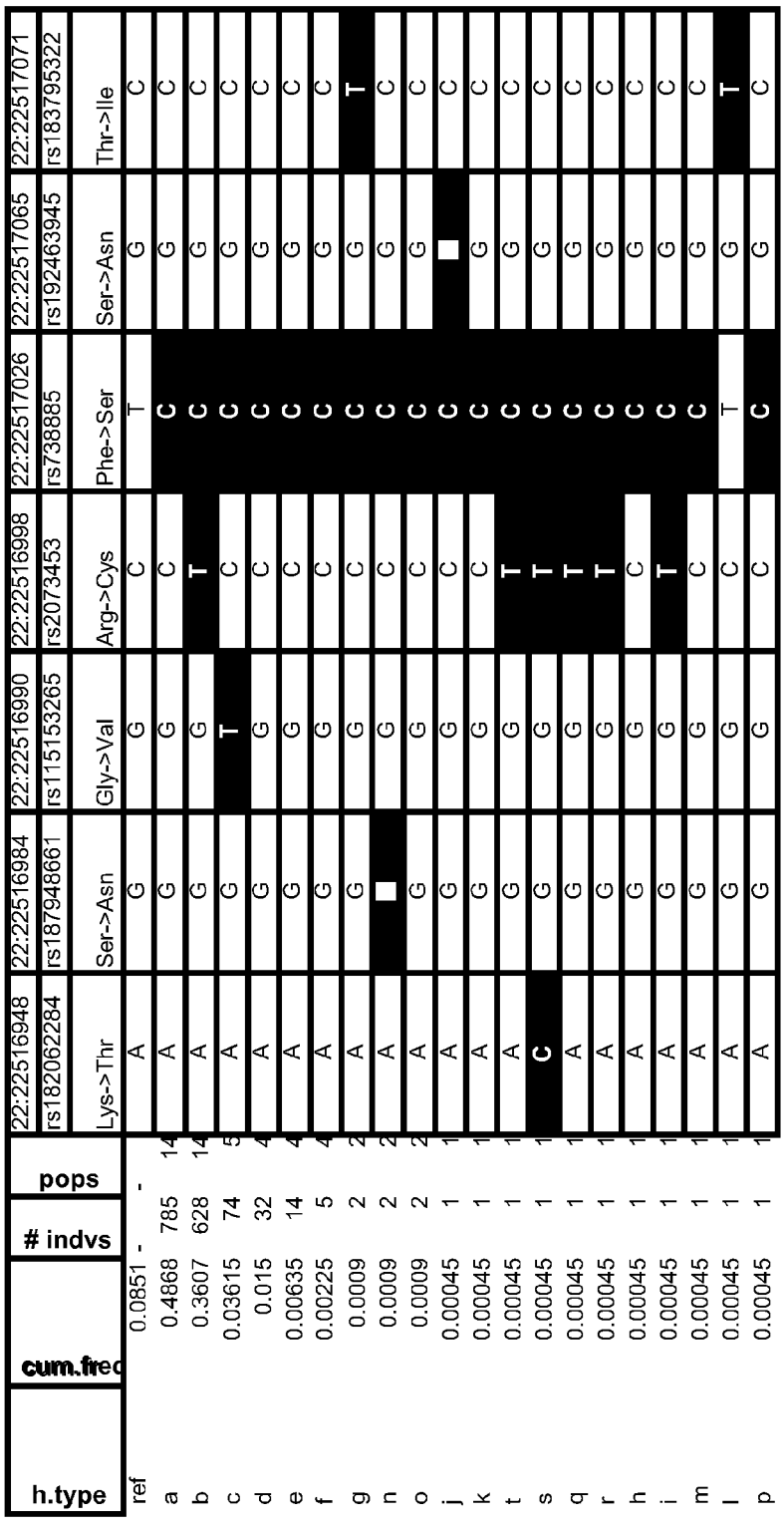
Table 13.134 - IGLV4-3

h.type	cum.freq	# indvs	pops	22:23213939 rs191969136 Leu->Phe	22:23213987 rs183930410 Ser->Arg	22:23214006 rs188425943 Tyr->His	22:23214144 rs1410259913 Leu->Phe	22:23214190 rs149834282 Thr->Lys	ENST00000390318
ref	0.9872	-	-	G	C	T	C	C	
a	0.01055	22	4	G	C	T	C	■	1,10,12,14
d	0.0009	2	2	G	C	**C**	C	C	7,14
e	0.00045	1	1	G	C	T	**T**	C	2
c	0.00045	1	1	**C**	C	T	C	C	14
b	0.00045	1	1	G	■	T	C	C	3

Table 13.135 - IGLV4-60

h.type	cum.freq	# indvs	pops	22:22516624	22:22516785	22:22516825	22:22516832	22:22516864	22:22516869	22:22516922	22:22516929	22:22516933
				.	rs190843139	rs114796103	rs186692431	rs78099525	rs183095995	rs189760597	rs138271962	rs74528122
					Leu->Val	Ser->Leu	Lys->Asn	Ser->Asn	Ile->Phe	Lys->Asn	Gly->Arg	Ser->Asn
ref	0.0851	-	-	TCTC	C	C	G	G	A	G	G	G
a	0.4868	785	14	TCTC	C	C	G	G	A	G	G	G
b	0.3607	628	14	TCTC	C	C	G	G	A	G	G	G
c	0.03615	74	5	TCTC	C	C	G	G	A	G	G	G
d	0.015	32	4	TCTC	C	T	G	G	A	G	G	G
e	0.00635	14	4	TCTC	C	C	G	G	A	G	G	■
f	0.00225	5	4	T	C	C	G	G	A	G	G	G
g	0.0009	2	2	TCTC	C	C	G	G	A	G	G	G
n	0.0009	2	2	TCTC	C	C	G	G	A	G	G	G
o	0.0009	2	2	TCTC	C	C	G	G	A	G	C	G
j	0.00045	1	1	TCTC	C	C	G	G	A	G	G	G
k	0.00045	1	1	TCTC	G	C	G	G	A	G	G	G
t	0.00045	1	1	TCTC	C	C	G	G	A	C	G	G
s	0.00045	1	1	TCTC	C	C	G	G	A	G	G	G
q	0.00045	1	1	TCTC	C	C	G	G	T	G	G	G
r	0.00045	1	1	TCTC	C	C	G	G	A	G	C	G
h	0.00045	1	1	TCTC	C	C	G	■	A	G	G	G
i	0.00045	1	1	TCTC	C	C	G	■	A	G	G	G
m	0.00045	1	1	TCTC	C	T	G	■	A	G	G	G
l	0.00045	1	1	TCTC	C	C	G	G	A	G	G	G
p	0.00045	1	1	TCTC	C	C	C	G	A	G	G	G

Table 13.135 - IGLV4-60



h.type	cum.freq	# indvs	pops	22:22516948 rs182062284 Lys->Thr	22:22516984 rs187948661 Ser->Asn	22:22516990 rs115153265 Gly->Val	22:22516998 rs2073453 Arg->Cys	22:22517026 rs738885 Phe->Ser	22:22517065 rs192463945 Ser->Asn	22:22517071 rs183795322 Thr->Ile
ref	0.0851	-	-	A	G	G	C	T	G	C
a	0.4668	785	14	A	G	G	C	C	G	C
b	0.3607	628	14	A	G	G	T	C	G	C
c	0.03615	74	5	A	G	T	C	C	G	C
d	0.015	32	4	A	G	G	C	C	G	C
e	0.00635	14	4	A	G	G	C	C	G	C
f	0.00225	5	4	A	G	G	C	C	G	C
g	0.0009	2	2	A	G	G	C	C	G	T
n	0.0009	2	2	A	■	G	C	C	G	C
o	0.0009	2	2	A	G	G	C	C	G	C
j	0.00045	1	1	A	G	G	C	C	■	C
k	0.00045	1	1	A	G	G	C	C	G	C
t	0.00045	1	1	A	G	G	T	C	G	C
s	0.00045	1	1	C	G	G	T	C	G	C
q	0.00045	1	1	A	G	G	T	C	G	C
r	0.00045	1	1	A	G	G	T	C	G	C
h	0.00045	1	1	A	G	G	C	C	G	C
i	0.00045	1	1	A	G	G	T	C	G	C
m	0.00045	1	1	A	G	G	C	C	G	C
l	0.00045	1	1	A	G	G	C	T	G	T
p	0.00045	1	1	A	G	G	C	C	G	C

Table 13.135 - IGLV4-60

h.type	cum.freq	# indvs	pops	ENST00000390284
ref	0.0851	-	-	
a	0.4668	785	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.3607	628	14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.03615	74	5	1,5,10,12,14
d	0.015	32	4	1,5,10,14
e	0.00635	14	4	1,10,11,14
f	0.00225	5	4	1,3,10,14
g	0.0009	2	2	5,11
n	0.0009	2	2	2,9
o	0.0009	2	2	5,10
j	0.00045	1	1	10
k	0.00045	1	1	5
t	0.00045	1	1	6
s	0.00045	1	1	13
q	0.00045	1	1	13
r	0.00045	1	1	13
h	0.00045	1	1	14
i	0.00045	1	1	10
m	0.00045	1	1	10
l	0.00045	1	1	5
p	0.00045	1	1	10

Table 13.136 - IGLV4-69

h.type	cum.freq	# indvs	pops	22:22385618 rs150132122 Ala->Asp	22:22385671 rs185429538 Ala->Pro	22:22385747 rs140354453 Gly->Glu	22:22385851 rs188851121 Thr->Pro	ENST00000390282
ref	0.98945	-	-	C	G	G	A	
a	0.0092	20	4	C	G	■	A	1,8,10,14
d	0.00045	1	1	C	**C**	G	A	5
c	0.00045	1	1	C	G	G	**C**	10
b	0.00045	1	1	■	G	G	A	2

Table 13.137 - IGLV5-37

h.type	cum.freq	# indvs	pops	22:22782150 rs56060925 Gly->Ser	22:22782240 rs181085584 Gly->Cys	22:22782258 rs111480336 Arg->Cys	ENST00000390300
ref	0.86435	-	-	G	G	C	
a	0.1266	251	13		G	C	1,2,3,4,5,6,7,8,9,10,11,12,13
b	0.0068	13	7	G	G	T	2,6,10,11,12,13,14
c	0.00225	4	2	G	T	C	10,14

Table 13.138 - IGLV5-45

h.type	cum.freq	# indvs	pops	22:22730582 rs12157664 Pro->Leu	22:22730584 rs1985918 Ser->Ala	22:22730635 rs61731372 Arg->Cys	22:22730636 rs149965450 Arg->His	22:22730663 rs191752433 Arg->Lys	22:22730725 rs192713801 Asp->His	22:22730732 rs117029072 Asp->Gly	22:22730786 rs147644645 Ser->Leu	ENST00000390296
ref	0.8017	-	-	C	T	C	G	G	G	A	C	
a	0.1706	312	12	C	**G**	C	G	G	G	A	C	1,2,3,5,6,7,8,10,11,12,13,14
b	0.0164	36	14	C	T	**T**	G	G	G	A	C	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.00545	12	2	**T**	**G**	C	G	G	G	A	C	10,14
d	0.0036	8	3	C	T	C	G	G	G	A	**T**	3,4,9
g	0.0009	2	2	C	T	C	■	G	G	A	C	2,13
e	0.00045	1	1	C	T	C	G	■	G	**G**	C	2
h	0.00045	1	1	C	**G**	C	G	G	**C**	A	C	10
f	0.00045	1	1	C	T	**T**	G	■	G	A	C	9

Table 13.139 - IGLV5-48

h.type	cum.freq	# indvs	pops	22:22707473	22:22707555	22:22707576	22:22707621	22:22707698	22:22707713	22:22707773	
				rs74549497	rs79612806	rs117042445	rs118041409	rs191708293	rs73880637	rs138488444	
				Pro->Ala	Asn->Ser	Phe->Tyr	Ser->Asn	Asn->Tyr	Val->Phe	Ser->Gly	ENST00000390293
ref	0.79525	-	-	C	A	T	G	A	G	A	
a	0.0873	174	14	G	G			A	T	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0301	66	13	G	G	T	G	A	G	A	1,2,3,4,5,6,7,8,9,10,11,13,14
b	0.02875	63	14	G	G	T	G	A	T	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
d	0.0123	27	10	G	G		G	A	T	A	1,2,4,5,6,9,10,12,13,14
e	0.0082	18	9	C	G		G	A	T	A	2,4,5,6,7,10,11,13,14
i	0.00595	13	7	C	G			A	T	A	2,3,4,7,9,11,13
f	0.0055	12	9	G	G	T		A	T	A	2,3,4,6,7,9,10,11,14
h	0.00545	12	8	C	A	T	G	A	T	A	1,2,3,6,10,12,13,14
g	0.0041	9	5	G	G			A	G	A	3,6,10,12,13
o	0.00315	7	6	C	A		G	A	T	A	1,2,7,8,11,13
k	0.00225	5	3	C	G	T	G	A	G	A	2,6,13
j	0.0018	4	3	C	A		G	A	G	A	4,10,12
t	0.00135	3	2	G	A	T	G	A	T	A	3,4
s	0.00135	3	3	G	A	T	G	A	G	A	2,5,13
u	0.0009	2	2	C	A			A	T	A	6,9
q	0.0009	2	2	C	A	T		A	G	A	5,12
r	0.0009	2	2	G	G		G	T	T	A	5,11
n	0.0009	2	2	C	G	T	G	A	T	A	2,10
m	0.0009	2	1	G	A	T		A	T	A	6
l	0.0009	2	2	C	G	T		A	T	A	12,13
v	0.00045	1	1	C	G		G	A	G	A	10
w	0.00045	1	1	G	G	T		A	G	A	2
x	0.00045	1	1	G	G		G	A	G	A	10
p	0.00045	1	1	C	A	T	G	A	G	G	3

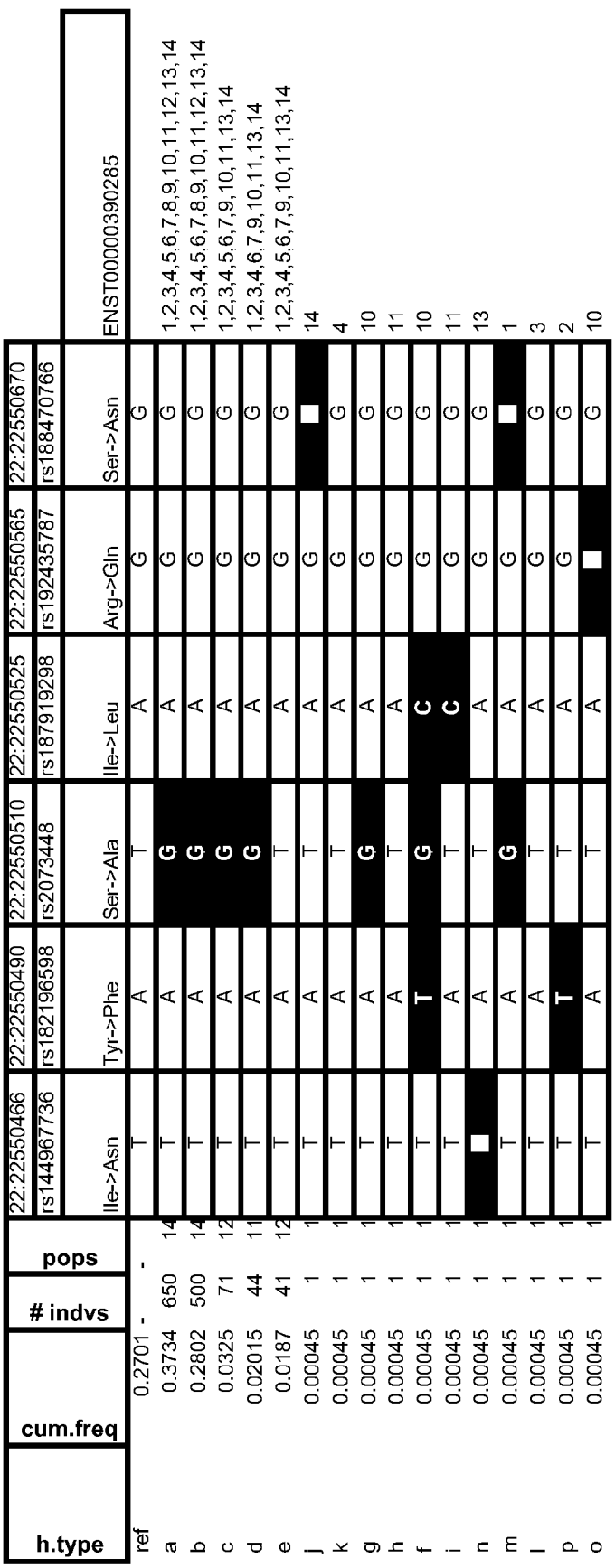
Table 13.140 - IGLV5-52

h.type	cum.freq	# indvs	pops	22:22673101	22:22673305	22:22673463	22:22673551	
				.	rs142368630	rs150977964	rs1872055544	ENST00000390289
					Ser->Phe	Arg->Cys	Cys->Tyr	
ref	0.9914	-	-	TCTC	C	C	G	
a	0.005	11	5	T	C	C	G	5,7,8,10,13
d	0.00135	3	2	TCTC	T	C	G	10,13
c	0.00135	3	1	TCTC	C	T	G	9
b	0.0009	2	1	TCTC	C	C	■	6

Table 13.141 - IGLV6-57

h.type	cum.freq	# indvs	pops	22:22550370 rs181723895 Gly->AlaGly->Ala	22:22550381 rs186024609 Asn->Asp	22:22550387 rs191529840 Met->Leu	22:22550399 rs183090103 Pro->Ser	22:22550405 rs186363981 Ser->Ala	22:22550415 rs190816426 Glu->Ala	22:22550436 rs149903410 Thr->Ser	22:22550450 rs2073447 Arg->Gly
ref	0.2701	-	-	G	A	A	C	T	A	C	C
a	0.3734	650	14	G	A	A	C	T	A	C	C
b	0.2602	500	14	G	A	A	C	T	A	C	**G**
c	0.0325	71	12	G	A	A	**T**	T	A	C	C
d	0.02015	44	11	G	A	A	**T**	T	A	C	**G**
e	0.0187	41	12	G	A	A	**T**	T	A	C	C
j	0.00045	1	1	G	A	A	**T**	T	A	C	C
k	0.00045	1	1	G	**G**	A	**T**	T	A	C	C
g	0.00045	1	1	G	A	**T**	C	T	A	C	C
h	0.00045	1	1	**C**	A	A	C	T	A	C	C
f	0.00045	1	1	G	A	A	C	T	A	C	C
i	0.00045	1	1	G	A	A	C	T	A	C	C
n	0.00045	1	1	G	A	A	C	T	A	C	C
m	0.00045	1	1	G	A	A	C	T	A	C	C
l	0.00045	1	1	G	A	A	C	T	A	**G**	C
p	0.00045	1	1	G	A	A	C	**G**	**C**	C	C
o	0.00045	1	1	G	A	A	C	T	A	C	C

Table 13.141 - IGLV6-57



h.type	cum.freq	# indvs	pops	22:22550466 rs144967736 Ile->Asn	22:22550490 rs182196598 Tyr->Phe	22:22550510 rs2073448 Ser->Ala	22:22550525 rs187919298 Ile->Leu	22:22550565 rs192435787 Arg->Gln	22:22550670 rs188470766 Ser->Asn	ENST00000390285
ref	0.2701	-	-	T	A	T	A	G	G	
a	0.3734	650	14	T	A	G	A	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
b	0.2602	500	14	T	A	G	A	G	G	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0325	71	12	T	A	G	A	G	G	1,2,3,4,5,6,7,9,10,11,13,14
d	0.02015	44	11	T	A	G	A	G	G	1,2,3,4,6,7,9,10,11,13,14
e	0.0187	41	12	T	A	T	A	G	G	1,2,3,4,5,6,7,9,10,11,13,14
j	0.00045	1	1	T	A	T	A	G	■	14
k	0.00045	1	1	T	A	T	A	G	G	4
g	0.00045	1	1	T	A	G	A	G	G	10
h	0.00045	1	1	T	A	T	A	G	G	11
f	0.00045	1	1	T	T	G	C	G	G	10
i	0.00045	1	1	T	A	T	C	G	G	11
n	0.00045	1	1	■	A	T	A	G	G	13
m	0.00045	1	1	T	A	G	A	G	■	1
l	0.00045	1	1	T	A	T	A	G	G	3
p	0.00045	1	1	T	T	T	A	G	G	2
o	0.00045	1	1	T	A	T	A	■	G	10

Table 13.142 - IGLV7-43

h.type	cum.freq	# indvs	pops	22:22749426 rs140166061 Cys->Arg	22:22749555 rs188344205 Pro->Thr	22:22749576 rs184825159 Gly->Arg	22:22749580 rs189388363 Gly->Glu	22:22749591 rs182275478 Leu->Val	22:22749676 rs73878432 Ala->Val	22:22749693 rs181430863 Ser->Gly	22:22749704 rs189954563 His->Gln	ENST00000390298
ref	0.97765	-	-	T	C	G	G	C	C	A	C	
a	0.01005	22	9	T	C	■	G	C	C	A	C	2,3,4,6,7,9,10,11,14
b	0.0087	19	4	T	C	G	G	C	**T**	A	C	1,5,10,14
c	0.00135	3	2	T	C	G	G	C	C	A	**G**	4,13
d	0.00045	1	1	T	■	G	G	C	C	A	C	1
g	0.00045	1	1	T	C	G	G	**G**	C	A	C	6
e	0.00045	1	1	T	C	G	G	C	C	**G**	C	1
h	0.00045	1	1	**C**	C	G	G	C	C	A	C	1
f	0.00045	1	1	T	C	G	■	C	C	A	C	1

Table 13.143 - IGLV7-46

h.type	cum.freq	# indvs	pops	22:22724244 rs188705617 Ala->Gly	22:22724306 rs141014218 Leu->Met	22:22724313 rs193180910 Tyr->Cys	22:22724332 rs185250547 His->Gln	22:22724394 rs7292000 Leu->Ser	22:22724424 rs191717127 Tyr->Phe	ENST00000390295
ref	0.0807	-	-	C	C	A	C	T	A	
a	0.9166	1080	14	C	C	A	C	C	A	1,2,3,4,5,6,7,8,9,10,11,12,13,14
c	0.0009	2	2	G	C	A	C	C	A	5,10
d	0.00045	1	1	C	C	G	C	C	A	10
e	0.00045	1	1	C	C	A	C	C	T	9
b	0.00045	1	1	C	■	A	C	C	A	13
f	0.00045	1	1	C	C	A	G	T	A	2

Table 13.144 - IGLV8-61

h.type	cum.freq	# indvs	pops	22:22453195 rs184643084 Gly->Arg	22:22453420 rs185633024 Ser->Gly	22:22453469 rs118030360 Thr->Met	22:22453495 rs181569364 Ser->Ala	22:22453513 rs9611014 Arg->Cys	22:22453546 rs188717570 Ala->Thr	ENST00000390283
ref	0.7574	-	-	G	A	C	T	C	G	
a	0.2193	407	12	G	A	C	T	**T**	G	1,2,3,4,5,6,7,8,9,11,12,13
b	0.02195	48	8	G	A	**T**	T	C	G	1,3,4,5,6,7,9,11
d	0.00045	1	1	G	**G**	C	T	**T**	G	2
e	0.00045	1	1	G	A	C	**G**	C	■	2
c	0.00045	1	1	■	A	C	T	**T**	G	5

Table 13.145 - IGLV9-49

h.type	cum.fre	# indvs	pops	22:22697585 rs146805069 Ala->Pro	22:22697782 rs1406612446 Thr->Ala	22:22697788 rs186017534 Pro->Ser	22:22697825 rs190876413 Thr->Lys	22:22697923 rs187756755 Thr->Ala	22:22697939 rs190665606 Gly->Glu	22:22697983 rs192733353 Ser->Pro	22:22697998 rs184707881 Tyr->His	22:22698031 rs148899698 Glu->Lys
ref	0.98455	-	-	G	A	C	C	A	G	T	T	G
a	0.00865	18	3	G	**G**	C	C	A	G	T	T	G
b	0.0032	7	2	G	A	C	C	A	G	T	T	■
c	0.0009	2	2	G	A	**T**	C	A	G	T	T	G
d	0.00045	1	1	G	A	C	■	A	G	T	T	G
g	0.00045	1	1	**C**	A	C	C	A	G	T	T	G
e	0.00045	1	1	G	A	C	C	A	G	**C**	T	G
h	0.00045	1	1	G	A	C	C	**G**	G	T	T	G
f	0.00045	1	1	G	A	C	C	A	G	T	**C**	G
i	0.00045	1	1	G	A	C	C	A	■	T	T	G

Table 13.145 - IGLV9-49

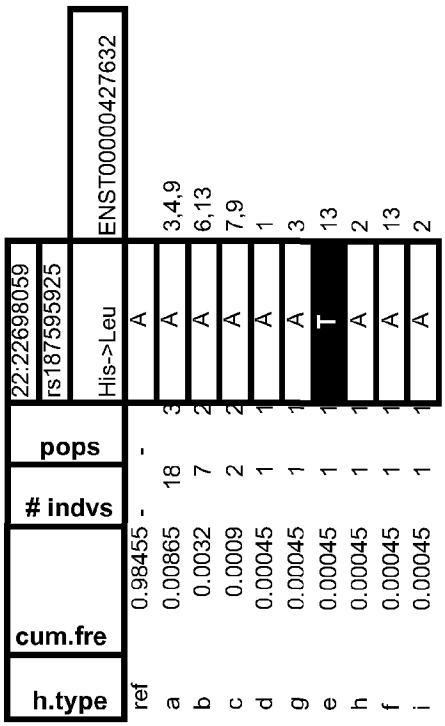


Table 14A: Variants Having a Cumulative Frequency Greater than 5% or 10%

GENE SEGMENT	VARIANT NAME	CUMULATIVE FREQUENCY	CUMULATIVE FREQUENCY >10%	NO UNIQUE POPULATIONS
Heavy C				
IGHA	IGHA1-a	0.31385	Yes	14
IGHA	IGHA2-a	0.35655	Yes	14
IGHA	IGHA2-b	0.1647	Yes	14
IGHG	IGHG2-a	0.28675	Yes	12
IGHG	IGHG3-a	0.2813	Yes	14
IGHG	IGHG3-b	0.19955	Yes	11
IGHG	IGHG4-a	0.19825	Yes	12
IGHM	IGHM-a	0.7527	Yes	14
IGHM	IGHM-b	0.15245	Yes	14
Heavy J				
IGHJ	IGHJ6-a	0.6845	Yes	14
Heavy V				
IGHV	IGHV1-2-a	0.3268	Yes	14
IGHV	IGHV1-2-b	0.1304	Yes	14
IGHV	IGHV1-2-c	0.0797	No	14
IGHV	IGHV1-3-a	0.3564	Yes	14

IGHV	IGHV1-3-b	0.0724	No	13
IGHV	IGHV1-45-a	0.1286	Yes	12
IGHV	IGHV1-45-b	0.1126	Yes	12
IGHV	IGHV1-58-a	0.4503	Yes	14
IGHV	IGHV1-58-b	0.3144	Yes	14
IGHV	IGHV1-69-a	0.1689	Yes	14
IGHV	IGHV1-69-b	0.1543	Yes	14
IGHV	IGHV1-69-c	0.0561	No	13
IGHV	IGHV2-26-a	0.06135	No	13
IGHV	IGHV2-5-a	0.6657	Yes	14
IGHV	IGHV2-70-a	0.31	Yes	14
IGHV	IGHV2-70-b	0.16595	Yes	14
IGHV	IGHV2-70-d	0.06035	No	13
IGHV	IGHV3-13-a	0.15245	Yes	13
IGHV	IGHV3-15-a	0.0572	No	11
IGHV	IGHV3-16-a	0.11075	Yes	13
IGHV	IGHV3-20-a	0.2829	Yes	14
IGHV	IGHV3-20-b	0.1209	Yes	13
IGHV	IGHV3-23-a	0.09825	No	12
IGHV	IGHV3-33-a	0.145	Yes	12
IGHV	IGHV3-35-a	0.11765	Yes	14
IGHV	IGHV3-35-b	0.09695	No	10
IGHV	IGHV3-38-a	0.0723	No	14
IGHV	IGHV3-43-a	0.2428	Yes	14
IGHV	IGHV3-43-b	0.08365	No	13
IGHV	IGHV3-48-a	0.3355	Yes	14

IGHV	IGHV3-48-b	0.20455	Yes	14
IGHV	IGHV3-49-a	0.38705	Yes	14
IGHV	IGHV3-49-b	0.28555	Yes	13
IGHV	IGHV3-53-a	0.2535	Yes	14
IGHV	IGHV3-53-b	0.2264	Yes	13
IGHV	IGHV3-64-a	0.47965	Yes	14
IGHV	IGHV3-64-b	0.06225	No	13
IGHV	IGHV3-66-a	0.27865	Yes	14
IGHV	IGHV3-66-b	0.12575	Yes	14
IGHV	IGHV3-66-c	0.0649	No	13
IGHV	IGHV4-28-a	0.41435	Yes	14
IGHV	IGHV4-28-b	0.13715	Yes	14
IGHV	IGHV4-28-c	0.07385	No	14
IGHV	IGHV4-28-d	0.06485	No	13
IGHV	IGHV4-31-a	0.5475	Yes	14
IGHV	IGHV4-31-b	0.2035	Yes	14
IGHV	IGHV4-31-c	0.1554	Yes	14
IGHV	IGHV4-39-a	0.10295	Yes	14
IGHV	IGHV4-39-b	0.0918	No	14
IGHV	IGHV4-4-a	0.1136	Yes	14
IGHV	IGHV4-4-d	0.0817	No	14
IGHV	IGHV4-4-f	0.06675	No	14
IGHV	IGHV4-4-b	0.06385	No	13
IGHV	IGHV4-4-c	0.0599	No	14
IGHV	IGHV4-4-e	0.05215	No	12
IGHV	IGHV4-61-a	0.5701	Yes	14

IGHV	IGHV4-61-b	0.18115	Yes	14
IGHV	IGHV4-61-c	0.13535	Yes	13
IGHV	IGHV4-61-f	0.05855	No	14
Kappa C IGKC	IGKC-a	0.2427	Yes	14
Kappa V IGKV	IGKV1-17-a	0.08655	No	13
IGKV	IGKV1-17-b	0.0545	No	12
IGKV	IGKV1-8-a	0.20315	Yes	14
IGKV	IGKV1-8-b	0.1334	Yes	14
IGKV	IGKV1-9-a	0.1854	Yes	14
IGKV	IGKV1D-42-a	0.22345	Yes	14
IGKV	IGKV1D-8-a	0.1872	Yes	14
IGKV	IGKV2-24-a	0.1708	Yes	13
IGKV	IGKV2-30-a	0.24085	Yes	13
IGKV	IGKV2D-26-a	0.325	Yes	14
IGKV	IGKV2D-29-a	0.30175	Yes	14
IGKV	IGKV3-7-a	0.2018	Yes	14
IGKV	IGKV3D-11-a	0.35105	Yes	14
IGKV	IGKV3D-15-a	0.2061	Yes	14
IGKV	IGKV5-2-a	0.32315	Yes	14
Lambda C IGLC	IGLC7-a	0.0751	No	14

Lambda J				
IGLJ	IGLJ2-a	0.0741	No	13
IGLJ	IGLJ7-a	0.93225	Yes	14
Lambda V				
IGLV	IGLV10-54-a	0.50725	Yes	14
IGLV	IGLV10-54-b	0.09425	No	14
IGLV	IGLV11-55-a	0.4463	Yes	14
IGLV	IGLV1-36-a	0.0942	No	14
IGLV	IGLV1-47-a	0.8413	Yes	14
IGLV	IGLV1-50-a	0.19325	Yes	14
IGLV	IGLV1-51-a	0.1177	Yes	11
IGLV	IGLV2-14-a	0.48365	Yes	14
IGLV	IGLV2-14-c	0.08855	No	14
IGLV	IGLV2-14-b	0.0804	No	10
IGLV	IGLV2-18-a	0.763	Yes	14
IGLV	IGLV2-23-a	0.34655	Yes	14
IGLV	IGLV2-8-a	0.1451	Yes	11
IGLV	IGLV3-12-a	0.38715	Yes	14
IGLV	IGLV3-12-b	0.1116	Yes	14
IGLV	IGLV3-12-c	0.0774	No	14
IGLV	IGLV3-19-a	0.10015	Yes	12
IGLV	IGLV3-21-a	0.31195	Yes	14
IGLV	IGLV3-21-b	0.178	Yes	14
IGLV	IGLV3-21-c	0.1613	Yes	14

IGLV	IGLV3-22-a	0.2962	Yes	14
IGLV	IGLV3-25-a	0.6581	Yes	14
IGLV	IGLV3-25-b	0.19965	Yes	14
IGLV	IGLV4-60-a	0.4868	Yes	14
IGLV	IGLV4-60-b	0.3607	Yes	14
IGLV	IGLV5-37-a	0.1266	Yes	13
IGLV	IGLV5-45-a	0.1706	Yes	12
IGLV	IGLV5-48-a	0.0873	No	14
IGLV	IGLV6-57-a	0.3734	Yes	14
IGLV	IGLV6-57-b	0.2802	Yes	14
IGLV	IGLV7-46-a	0.9166	Yes	14
IGLV	IGLV8-61-a	0.2193	Yes	12

Table 14B: Variants Having a Cumulative Frequency Greater than 20%

GENE SEGMENT	VARIANT NAME	CUMULATIAVE FREQUENCY	NO UNIQUE POPULATIONS
Heavy C			
IGHA	IGHA1-a	0.31385	14
IGHA	IGHA2-a	0.35655	14
IGHG	IGHG2-a	0.28675	12
IGHG	IGHG3-a	0.2813	14
IGHM	IGHM-a	0.7527	14
Heavy J			
IGHJ	IGHJ6-a	0.6845	14

Heavy V			
IGHV	IGHV1-2-a	0.3268	14
IGHV	IGHV1-3-a	0.3564	14
IGHV	IGHV1-58-a	0.4503	14
IGHV	IGHV1-58-b	0.3144	14
IGHV	IGHV2-5-a	0.6657	14
IGHV	IGHV2-70-a	0.31	14
IGHV	IGHV3-20-a	0.2829	14
IGHV	IGHV3-43-a	0.2428	14
IGHV	IGHV3-48-a	0.3355	14
IGHV	IGHV3-48-b	0.20455	14
IGHV	IGHV3-49-a	0.38705	14
IGHV	IGHV3-49-b	0.28555	13
IGHV	IGHV3-53-a	0.2535	14
IGHV	IGHV3-53-b	0.2264	13
IGHV	IGHV3-64-a	0.47965	14
IGHV	IGHV3-66-a	0.27865	14
IGHV	IGHV4-28-a	0.41435	14
IGHV	IGHV4-31-a	0.5475	14
IGHV	IGHV4-31-b	0.2035	14
IGHV	IGHV4-61-a	0.5701	14
Kappa C			
IGKC	IGKC-a	0.2427	14
	IGKV1-8-a	0.20315	14

Kappa V IGKV			
IGKV	IGKV1D-42-a	0.22345	14
IGKV	IGKV2-30-a	0.24085	13
IGKV	IGKV2D-26-a	0.325	14
IGKV	IGKV2D-29-a	0.30175	14
IGKV	IGKV3-7-a	0.2018	14
IGKV	IGKV3D-11-a	0.35105	14
IGKV	IGKV3D-15-a	0.2061	14
IGKV	IGKV5-2-a	0.32315	14
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V IGLV	IGLV10-54-a	0.50725	14
IGLV	IGLV11-55-a	0.4463	14
IGLV	IGLV1-47-a	0.8413	14
IGLV	IGLV2-14-a	0.48365	14
IGLV	IGLV2-18-a	0.763	14
IGLV	IGLV2-23-a	0.34655	14
IGLV	IGLV3-12-a	0.38715	14
IGLV	IGLV3-21-a	0.31195	14
IGLV	IGLV3-22-a	0.2962	14
IGLV	IGLV3-25-a	0.6581	14
IGLV	IGLV4-60-a	0.4868	14

IGLV	IGLV4-60-b	0.3607	14
IGLV	IGLV6-57-a	0.3734	14
IGLV	IGLV6-57-b	0.2802	14
IGLV	IGLV7-46-a	0.9166	14
IGLV	IGLV8-61-a	0.2193	12

Table 14C: Variants Having a Cumulative Frequency Greater than 30%

GENE SEGMENT	VARIANT NAME	CUMULATIVE FREQUENCY	NO UNIQUE POPULATIONS
Heavy C			
IGHA	IGHA1-a	0.31385	14
IGHA	IGHA2-a	0.35655	14
IGHM	IGHM-a	0.7527	14
Heavy J			
IGHJ	IGHJ6-a	0.6845	14
Heavy V			
IGHV	IGHV1-2-a	0.3268	14
IGHV	IGHV1-3-a	0.3564	14
IGHV	IGHV1-58-a	0.4503	14
IGHV	IGHV1-58-b	0.3144	14
IGHV	IGHV2-5-a	0.6657	14
IGHV	IGHV2-70-a	0.31	14
IGHV	IGHV3-48-a	0.3355	14

IGHV	IGHV3-49-a	0.38705	14
IGHV	IGHV3-64-a	0.47965	14
IGHV	IGHV4-28-a	0.41435	14
IGHV	IGHV4-31-a	0.5475	14
IGHV	IGHV4-61-a	0.5701	14
IGKV	IGKV2D-26-a	0.325	14
IGKV	IGKV2D-29-a	0.30175	14
IGKV	IGKV3D-11-a	0.35105	14
IGKV	IGKV5-2-a	0.32315	14
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V IGLV	IGLV10-54-a	0.50725	14
IGLV	IGLV11-55-a	0.4463	14
IGLV	IGLV1-47-a	0.8413	14
IGLV	IGLV2-14-a	0.48365	14
IGLV	IGLV2-18-a	0.763	14
IGLV	IGLV2-23-a	0.34655	14
IGLV	IGLV3-12-a	0.38715	14
IGLV	IGLV3-21-a	0.31195	14
IGLV	IGLV3-25-a	0.6581	14
IGLV	IGLV4-60-a	0.4868	14
IGLV	IGLV4-60-b	0.3607	14
IGLV	IGLV6-57-a	0.3734	14

IGLV	IGLV7-46-a	0.9166	14

Table 14D: Variants Having a Cumulative Frequency Greater than 40%

GENE SEGMENT	VARIANT NAME	CUMULATIAVE FREQUENCY	NO UNIQUE POPULATIONS
Heavy C IGHM	IGHM-a	0.7527	14
Heavy J IGHJ	IGHJ6-a	0.6845	14
Heavy V IGHV	IGHV1-58-a	0.4503	14
IGHV	IGHV2-5-a	0.6657	14
IGHV	IGHV3-64-a	0.47965	14
IGHV	IGHV4-28-a	0.41435	14
IGHV	IGHV4-31-a	0.5475	14
IGHV	IGHV4-61-a	0.5701	14
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V IGLV	IGLV10-54-a	0.50725	14
IGLV	IGLV11-55-a	0.4463	14

IGLV	IGLV1-47-a	0.8413	14
IGLV	IGLV2-14-a	0.48365	14
IGLV	IGLV2-18-a	0.763	14
IGLV	IGLV3-25-a	0.6581	14
IGLV	IGLV4-60-a	0.4868	14
IGLV	IGLV7-46-a	0.9166	14

Table 14E: Variants Having a Cumulative Frequency Greater than 50%

GENE SEGMENT	VARIANT NAME	CUMULATIVE FREQUENCY	NO UNIQUE POPULATIONS
Heavy C IGHM	IGHM-a	0.7527	14
Heavy J IGHJ	IGHJ6-a	0.6845	14
Heavy V IGHV	IGHV2-5-a	0.6657	14
IGHV	IGHV4-31-a	0.5475	14
IGHV	IGHV4-61-a	0.5701	14
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V	IGLV10-54-a	0.50725	14

IGLV			
IGLV	IGLV1-47-a	0.8413	14
IGLV	IGLV2-18-a	0.763	14
IGLV	IGLV3-25-a	0.6581	14
IGLV	IGLV7-46-a	0.9166	14

Table 14F: Variants Having a Cumulative Frequency Greater than 60%

GENE SEGMENT	VARIANT NAME	CUMULATIVE FREQUENCY	NO UNIQUE POPULATIONS
Heavy C IGHM	IGHM-a	0.7527	14
Heavy J IGHJ	IGHJ6-a	0.6845	14
Heavy V IGHV	IGHV2-5-a	0.6657	14
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V IGLV	IGLV1-47-a	0.8413	14
IGLV	IGLV2-18-a	0.763	14
IGLV	IGLV3-25-a	0.6581	14

IGLV	IGLV7-46-a	0.9166	14

Table 14G: Variants Having a Cumulative Frequency Greater than 70%

GENE SEGMENT	VARIANT NAME	CUMULATIVE FREQUENCY	NO UNIQUE POPULATIONS
Heavy C IGHM	IGHM-a	0.7527	14
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V IGLV	IGLV1-47-a	0.8413	14
IGLV	IGLV2-18-a	0.763	14
IGLV	IGLV7-46-a	0.9166	14

Table 14H: Variants Having a Cumulative Frequency Greater than 80%

GENE SEGMENT	VARIANT NAME	CUMULATIVE FREQUENCY	NO UNIQUE POPULATIONS
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V	IGLV1-47-a	0.8413	14

IGLV			
IGLV	IGLV7-46-a	0.9166	14

Table 14I: Variants Having a Cumulative Frequency Greater than 90%

GENE SEGMENT	VARIANT NAME	CUMULATIVE FREQUENCY	NO UNIQUE POPULATIONS
Lambda J IGLJ	IGLJ7-a	0.93225	14
Lambda V			
IGLV	IGLV7-46-a	0.9166	14

Table 15: NAÏVE HEAVY CHAIN REPERTOIRES

TABLE 15A: All Naïve – J Usage			
J	Average HCDR3 Length	Count	%
IGHJ1*01	11	12	0.90%
IGHJ2*01	12	47	3.51%
IGHJ3*02	12	205	15.30%
IGHJ4*02	11	689	51.42%
IGHJ5*02	13	88	6.57%
IGHJ6*02	**16**	**299**	**22.31%**

Naive HCDR3>=20			
J	Average HCDR3 Length	Count	%

IGHJ3*02	20	1	2.22%
IGHJ4*02	20	1	2.22%
IGHJ5*02	20	2	4.44%
IGHJ6*02	**21**	**41**	**91.11%**

Naïve – Long HCDR3 using IGHJ6*02

HCDR3Length	Count
20	19
21	10
22	7
23	3
24	1
26	1

TABLE 15B: All Naïve – J & D Usage			
J	D	Average HCDR3 Length	Count
IGHJ6*02	IGHD3-9*01	19	21
IGHJ6*02	IGHD4-17*01	18	7
IGHJ6*02	IGHD3-10*01	17	98
IGHJ6*02	IGHD2-2*02	17	4
IGHJ6*02	IGHD5-24*01	17	1
IGHJ6*02	IGHD6-19*01	16	30
IGHJ6*02	IGHD3-22*01	16	5
IGHJ6*02	IGHD6-13*01	16	33
IGHJ6*02	IGHD5-12*01	15	7
IGHJ6*02	IGHD1-26*01	15	25
IGHJ6*02	IGHD1-20*01	14	6
IGHJ6*02	IGHD5-18*01	14	3
IGHJ6*02	IGHD3-16*02	14	6
IGHJ6*02	IGHD2-21*02	13	5
IGHJ6*02	IGHD1-14*01	13	4
IGHJ6*02	IGHD7-27*02	12	8
IGHJ6*02	IGHD1-1*01	12	8
IGHJ6*02	IGHD6-25*01	12	3
IGHJ6*02	IGHD4-23*01	11	2

Naive HCDR3>=20			
J	D	Average HCDR3 Length	Count
IGHJ6*02	IGHD6-19*01	23	1
IGHJ6*02	IGHD4-17*01	22	2
IGHJ6*02	IGHD3-9*01	21	11
IGHJ6*02	IGHD3-10*01	21	25
IGHJ6*02	IGHD6-13*01	20	1
IGHJ6*02	IGHD3-22*01	20	1

TABLE 15C: All Naïve – V & J Usage			
V	J	Average HCDR3 Length	Count
IGHV3-21*03	IGHJ6*02	20	5
IGHV1-18*01	IGHJ6*02	20	10
IGHV1-2*02	IGHJ6*02	19	5
IGHV3-13*01	IGHJ6*02	19	3
IGHV3-7*01	IGHJ6*02	19	6
IGHV1-8*01	IGHJ6*02	19	21
IGHV7-4-1*01	IGHJ6*02	19	9
IGHV3-9*01	IGHJ6*02	18	13
IGHV4-61*01	IGHJ6*02	17	20
IGHV3-23*04	IGHJ6*02	17	6
IGHV4-4*02	IGHJ6*02	16	34
IGHV1-3*01	IGHJ6*02	16	39
IGHV3-20*d01	IGHJ6*02	16	1
IGHV6-1*01	IGHJ6*02	14	69
IGHV2-5*10	IGHJ6*02	14	8
IGHV3-11*01	IGHJ6*02	14	3
IGHV3-66*03	IGHJ6*02	13	21
IGHV3-15*01	IGHJ6*02	12	3

Naive HCDR3>=20		

V	J	Average HCDR3 Length	Count
IGHV3-21*03	IGHJ6*02	23	3
IGHV3-13*01	IGHJ6*02	22	1
IGHV3-7*01	IGHJ6*02	22	3
IGHV6-1*01	IGHJ6*02	21	3
IGHV1-8*01	IGHJ6*02	21	10
IGHV1-2*02	IGHJ6*02	21	3
IGHV7-4-1*01	IGHJ6*02	21	3
IGHV1-3*01	IGHJ6*02	21	4
IGHV1-18*01	IGHJ6*02	21	5
IGHV4-4*02	IGHJ6*02	20	3
IGHV3-9*01	IGHJ6*02	20	2
IGHV3-23*04	IGHJ6*02	20	1

TABLE 16: IMMUNISED HEAVY CHAIN REPERTOIRES

TABLE 16A: All Immunised – J Usage			
J	HCDR3 Length	Count	%
IGHJ1*01	11	2	0.78%
IGHJ2*01	14	7	2.73%
IGHJ3*02	15	12	4.69%
IGHJ4*02	15	120	46.88%
IGHJ5*02	15	19	7.42%
IGHJ6*02	**17**	**96**	**37.50%**

HCDR3>20			
J	HCDR3 Length	Count	%
IGHJ4*02	22	2	22.22%
IGHJ5*02	25	1	11.11%
IGHJ6*02	**21**	**6**	**66.67%**

Immunised – Long HCDR3 using IGHJ6*02

HCDR3Length	Count
20	4

21	1
24	1

TABLE 16B:
All Immunised – J & D Usage

J	D	HCDR3 Length	Count
IGHJ6*02	IGHD4-17*01	18	1
IGHJ6*02	IGHD1-26*01	18	5
IGHJ6*02	IGHD6-19*01	17	27
IGHJ6*02	IGHD3-10*01	17	42
IGHJ6*02	IGHD6-13*01	16	2
IGHJ6*02	IGHD5-18*01	15	2
IGHJ6*02	IGHD4-23*01	15	5
IGHJ6*02	IGHD5-12*01	14	3
IGHJ6*02	IGHD3-16*02	14	3
IGHJ6*02	IGHD3-22*01	13	3
IGHJ6*02	IGHD3-9*01	12	1
IGHJ6*02	IGHD1-20*01	11	1

HCDR3>20			
J	D	HCDR3 Length	Count
IGHJ6*02	IGHD3-10*01	21	4
IGHJ6*02	IGHD6-19*01	20	1
IGHJ6*02	IGHD1-26*01	20	1

TABLE 16C:
All Immunised – V & J Usage

V	J	HCDR3 Length	Count
IGHV3-7*01	IGHJ6*02	18	33
IGHV1-8*01	IGHJ6*02	17	3
IGHV3-9*01	IGHJ6*02	17	22
IGHV3-13*01	IGHJ6*02	17	5
IGHV1-2*02	IGHJ6*02	16	8
IGHV3-11*01	IGHJ6*02	16	3
IGHV4-4*02	IGHJ6*02	15	9
IGHV6-1*01	IGHJ6*02	15	6
IGHV1-3*01	IGHJ6*02	15	6

HCDR3>20			
V	J	HCDR3 Length	Count
IGHV3-7*01	IGHJ6*02	21	4
IGHV3-11*01	IGHJ6*02	21	1
IGHV4-4*02	IGHJ6*02	20	1

TABLE 17: ANTIGEN-SPECIFIC HEAVY CHAIN REPERTOIRES

TABLE 17A: All Antigen-Specific – J Usage			
J	HCDR3 Length	Count	%
IGHJ1*01	12	2	1.68%
IGHJ3*02	17	4	3.36%
IGHJ4*02	13	64	53.78%
IGHJ5*02	19	6	5.04%
IGHJ6*02	**17**	**43**	**36.13%**
HCDR3>20			
J	HCDR3 Length	Count	%
IGHJ4*02	22	1	9.09%
IGHJ5*02	21	3	27.27%
IGHJ6*02	**21**	**7**	**63.64%**

Immunised – Long HCDR3 using IGHJ6*02

HCDR3Length	Count
20	4
21	1
22	1
24	1

TABLE 17B: All Antigen-Specific – J & D Usage			
J	D	HCDR3 Length	Count
IGHJ6*02	IGHD3-9*01	19	6
IGHJ6*02	IGHD3-10*01	19	12
IGHJ6*02	IGHD6-19*01	17	13
IGHJ6*02	IGHD5-12*01	16	2
IGHJ6*02	IGHD2-15*01	16	1
IGHJ6*02	IGHD1-26*01	15	7
IGHJ6*02	IGHD3-16*02	12	1
IGHJ6*02	IGHD5-18*01	11	1
HCDR3>20			
J	D	HCDR3 Length	Count
IGHJ6*02	IGHD3-9*01	21	1
IGHJ6*02	IGHD3-10*01	21	6

TABLE 17C: All Antigen-Specific – V & J Usage			
V	J	HCDR3 Length	Count
IGHV1-3*01	IGHJ6*02	18	1
IGHV3-7*01	IGHJ6*02	18	6
IGHV3-9*01	IGHJ6*02	18	6
IGHV1-8*01	IGHJ6*02	17	6
IGHV3-20*d01	IGHJ6*02	17	3
IGHV3-11*01	IGHJ6*02	17	8
IGHV4-4*02	IGHJ6*02	17	5
IGHV3-15*01	IGHJ6*02	15	5
IGHV3-13*01	IGHJ6*02	15	3
HCDR3>20			
V	J	HCDR3 Length	Count
IGHV4-4*02	IGHJ6*02	20	2

IGHV1-8*01	IGHJ6*02	22	2
IGHV3-9*01	IGHJ6*02	20	1
IGHV3-11*01	IGHJ6*02	21	1
IGHV3-20*d01	IGHJ6*02	22	1

Table 18: Sequence Correlation Table

SEQ ID NO:	DESCRIPTION:
1	JH5 reference (nucleotide sequence)
2	JH6*03=JH6 reference (nucleotide sequence)
3	JH2 reference (nucleotide sequence)
4	JH6*02 (nucleotide sequence)
5	JH5 reference (nucleotide sequence)
6	JH5 reference (amino acid sequence)
7	JH6*03=JH6 reference (nucleotide sequence)
8	JH6*03=JH6 reference (amino acid sequence)
9	JH2 reference (nucleotide sequence)
10	JH2 reference (amino acid sequence)
11	IGHV1-2*01
12	IGHV1-3*01
13	IGHV1-8*01
14	IGHV1-24*01
15	IGHV1-45*01
16	IGHV1-46*01
17	IGHV1-58*01
18	IGHV1-69*01
19	IGHV1-c*01

20	IGHV1-f*01
21	IGHV2-5*01
22	IGHV2-26*01
23	IGHV2-70*01
24	IGHV3-7*01
25	IGHV3-9*01
26	IGHV3-11*01
27	IGHV3-13*01
28	IGHV3-15*01
29	IGHV3-16*01
30	IGHV3-19*01
31	IGHV3-20*01
32	IGHV3-21*01
33	IGHV3-23*01
34	IGHV3-30*01

35	IGHV3-30-3*01
36	IGHV3-33*01
37	IGHV3-35*01
38	IGHV3-38*01
39	IGHV3-43*01
40	IGHV3-47*01
41	IGHV3-48*01
42	IGHV3-49*01
43	IGHV3-53*01
44	IGHV3-64*01
45	IGHV3-66*01
46	IGHV3-72*01
47	IGHV3-73*01
48	IGHV3-74*01
49	IGHV3-d*01

50	IGHV3-h*01
51	IGHV4-4*01
52	IGHV4-28*01
53	IGHV4-30-2*01
54	IGHV4-30-4*01
55	IGHV4-31*01
56	IGHV4-34*01
57	IGHV4-39*01
58	IGHV4-55*01
59	IGHV4-59*01
60	IGHV4-61*01
61	IGHV4-b*01
62	IGHV5-51*01
63	IGHV5-a*01

64	IGHV6-1*01
65	IGHV7-4-1*01
66	IGHV7-81*01
67	IGHD1-1*01
68	IGHD1-7*01
69	IGHD1-14*01
70	IGHD1-20*01
71	IGHD1-26*01
72	IGHD2-2*01
73	IGHD2-8*01
74	IGHD2-15*01
75	IGHD2-21*01
76	IGHD3-3*01
77	IGHD3-9*01
78	IGHD3-10*01
79	IGHD3-16*01
80	IGHD3-22*01
81	IGHD4-4*01
82	IGHD4-11*01
83	IGHD4-17*01
84	IGHD4-23*01
85	IGHD5-5*01
86	IGHD5-12*01

87	IGHD5-18*01
88	IGHD5-24*01
89	IGHD6-6*01
90	IGHD6-13*01
91	IGHD6-19*01
92	IGHD6-25*01
93	IGHD6-25*01
94	JH1*01
95	JH2*01
96	JH3*01
97	JH4*01
98	JH5*01
99	JH6*01
100	IGKV1-5*01
101	IGKV1-6*01
102	IGKV1-8*01
103	IGKV1D-8*01
104	IGKV1-9*01
105	IGKV1-12*01
106	IGKV1-13*01

107	IGKV1-16*01
108	IGKV1D-16*01
109	IGKV1-17*01
110	IGKV1D-17*01
111	IGKV1-27*01
112	IGKV1-33*01
113	IGKV1D-33*01
114	IGKV1-37*01
115	IGKV1D-37*01
116	IGKV1-39*01
117	IGKV1D-39
118	IGKV1D-42*01
119	IGKV1D-43*01
120	IGKV2-24*01
121	IGKV2D-24*01

122	IGKV2-28*01
123	IGKV2D-28*01
124	IGKV2-29*01
125	IGKV2D-29*01
126	IGKV2-30*01
127	IGKV2D-30*01
128	IGKV2-40*01
129	IGKV2D-40*01
130	IGKV3-7*01
131	IGKV3-7*03
132	IGKV3D-7*01
133	IGKV3-11*01
134	IGKV3D-11*01
135	IGKV3-15*01
136	IGKV3D-15*01

137	IGKV3-20*01
138	IGKV3D-20*01
139	IGKV4-1*01
140	IGKV5-2*01
141	IGKV6-21*01
142	IGKV6D-21*01
143	IGKV6D-41*01
144	IGLV1-36*01
145	IGLV1-40*01
146	IGLV1-41*01
147	IGLV1-44*01
148	IGLV1-47*01
149	IGLV1-50*01
150	IGLV1-51*01
151	IGLV2-8*01

152	IGLV2-11*01
153	IGLV2-14*01
154	IGLV2-18*01
155	IGLV2-23*01
156	IGLV2-33*01
157	IGLV3-1*01
158	IGLV3-10*01
159	IGLV3-12*01
160	IGLV3-16*01
161	IGLV3-19*01
162	IGLV3-21*01
163	IGLV3-22*01
164	IGLV3-25*01
165	IGLV3-27*01
166	IGLV3-32*01

167	IGLV4-3*01
168	IGLV4-60*01
169	IGLV4-69*01
170	IGLV5-37*01
171	IGLV5-45*01
172	IGLV5-48*01
173	IGLV5-52*01
174	IGLV6-57*01
175	IGLV7-43*01
176	IGLV7-46*01
177	IGLV8-61*01
178	IGLV9-49*01
179	IGLV10-54*01

180	IGLV11-55*01
181	IGKJ1*01
182	IGKJ2*01
183	IGKJ3*01
184	IGKJ4*01
185	IGKJ5*01
186	IGLJ1*01
187	IGLJ2*01
188	IGLJ3*01
189	IGLJ4*01
190	IGLJ5*01
191	IGLJ6*01
192	IGLJ7*01
193	IGKV2D-26*01
194	IGLV1-36*01
195	IGLV1-47*01
196	IGLV1-50*01
197	IGLV1-51*01
198	IGLV10-54*01

199	IGLV11-55*01
200	IGLV2-14*01
201	IGLV2-18*01
202	IGLV2-23*01
203	IGLV2-8*01
204	IGLV3-1*01
205	IGLV3-12*01
206	IGLV3-19*01
207	IGLV3-21*01
208	IGLV3-22*01
209	IGLV3-25*01
210	IGLV4-60*01
211	IGLV5-37*01
212	IGLV5-45*01

213	IGLV5-48*01
214	IGLV6-57*01
215	IGLV7-46*01
216	IGLV8-61*01
217	IGHG3
218	IGHA2
219	IGHG2

220	IGHE
221	IGHE
222	IGHA1
223	IGHG1
224	IGHG1

225	IGHM
226	IGHG1
227	IGHD
228	IGHG4
229	IGHE
230	IGKC
231	IGLC2
232	IGLC1
233	IGLC3
234	IGLC7
235	JH6*02 (nucleotide sequence)
236	JH6*01 & JH6*02(amino acid sequence)
237	Human RSS-JH6*02
238	Human RSS
239	IGHJ1 ref

240	IGHJ2 ref
241	IGHJ3 ref
242	IGHJ4 ref
243	IGHJ5 ref
244	IGHJ6 ref
245	IGHV1-18 ref
246	IGHV1-2 ref
247	IGHV1-24 ref
248	IGHV1-3 ref
249	IGHV1-45 ref
250	IGHV1-46 ref
251	IGHV1-58 ref
252	IGHV1-69 ref
253	IGHV1-8 ref
254	IGHV2-26 ref
255	IGHV2-5 ref
256	IGHV2-70 ref
257	IGHV3-11 ref
258	IGHV3-13 ref
259	IGHV3-15 ref

260	IGHV3-16 ref
261	IGHV3-20 ref
262	IGHV3-21 ref
263	IGHV3-23 ref
264	IGHV3-30 ref
265	IGHV3-33 ref
266	IGHV3-35 ref
267	IGHV3-38 ref
268	IGHV3-43 ref
269	IGHV3-48 ref
270	IGHV3-49 ref
271	IGHV3-53 ref
272	IGHV3-64 ref
273	IGHV3-66 ref
274	IGHV3-7 ref

275	IGHV3-72 ref
276	IGHV3-73 ref
277	IGHV3-74 ref
278	IGHV3-9 ref
279	IGHV4-28 ref
280	IGHV4-31 ref
281	IGHV4-34 ref
282	IGHV4-39 ref
283	IGHV4-4 ref
284	IGHV4-59 ref
285	IGHV4-61 ref
286	IGHV5-51 ref
287	IGHV6-1 ref
288	IGHV7-81 ref
289	IGKJ2 ref
290	IGKJ3 ref
291	IGKJ4 ref

292	IGKJ5 ref
293	IGKV1-16 ref
294	IGKV1-17 ref
295	IGKV1-5 ref
296	IGKV1-6 ref
297	IGKV1-8 ref
298	IGKV1-9 ref
299	IGKV1D-12 ref
300	IGKV1D-16 ref
301	IGKV1D-17 ref
302	IGKV1D-42 ref
303	IGKV1D-43 ref
304	IGKV1D-8 ref
305	IGKV2-24 ref
306	IGKV2-30 ref
307	IGKV2D-24 ref

308	IGKV2D-26 ref
309	IGKV2D-29 ref
310	IGKV2D-30 ref
311	IGKV3-11 ref
312	IGKV3-20 ref
313	IGKV3-7 ref
314	IGKV3D-11 ref
315	IGKV3D-15 ref
316	IGKV3D-20 ref
317	IGKV4-1 ref
318	IGKV5-2 ref
319	IGKV6-21 ref
320	IGKV6D-21 ref
321	IGKV6D-41 ref
322	IGLJ1 ref

323	IGLJ2 ref
324	IGLJ3 ref
325	IGLJ5 ref
326	IGLJ6 ref
327	IGLJ7 ref
328	IGLV10-54 ref
329	IGLV11-55 ref
330	IGLV1-36 ref
331	IGLV1-40 ref
332	IGLV1-44 ref
333	IGLV1-47 ref
334	IGLV1-50 ref
335	IGLV1-51 ref
336	IGLV2-11 ref
337	IGLV2-14 ref
338	IGLV2-18 ref
339	IGLV2-23 ref
340	IGLV2-33 ref

341	IGLV2-8 ref
342	IGLV3-1 ref
343	IGLV3-10 ref
344	IGLV3-12 ref
345	IGLV3-16 ref
346	IGLV3-19 ref
347	IGLV3-21 ref
348	IGLV3-22 ref
349	IGLV3-25 ref
350	IGLV3-27 ref
351	IGLV3-32 ref
352	IGLV3-9 ref
353	IGLV4-3 ref
354	IGLV4-60 ref

355	IGLV4-69 ref
356	IGLV5-37 ref
357	IGLV5-45 ref
358	IGLV5-48 ref
359	IGLV5-52 ref
360	IGLV6-57 ref
361	IGLV7-43 ref
362	IGLV7-46 ref
363	IGLV8-61 ref
364	IGLV9-49 ref

365	IGHG1 ref GENOMIC
366	IGHG1 ref - CDS (ensembl transcript ID ENST00000390542)
367	IGHG1 ref - CDS (Ensembl transcript ID ENST00000390549)
368	IGHG1 ref - CDS (ensembl transcript ID ENST00000390548)

369	IGHG2 ref - GENOMIC
370	IGHG2 ref - CDS (ensembl transcript ID ENST00000390545)
371	IGHG2-a CDS
372	IGHG2-a GENOMIC
373	IGHG3 ref - GENOMIC
374	IGHG3 ref - CDS (ensembl transcript ID ENST00000390551)
375	IGHG3-a CDS

376	IGHG3-a GENOMIC
377	IGHG3-b CDS
378	IGHG3-b GENOMIC
379	IGHG4 ref - GENOMIC
380	IGHG4 ref - CDS (ensembl transcript ID ENST00000390543)
381	IGHG4-a CDS
382	IGHG4-a GENOMIC
383	IGHA1 ref GENOMIC
384	IGHA1 ref- CDS (ensembl transcript ID ENST00000390547)
385	IGHA1-a GENOMIC
386	IGHA1-a CDS
387	IGHA2 ref GENOMIC
388	IGHA2 ref- CDS (ensembl transcript ID ENST00000390539)

389	IGHA2-a GENOMIC
390	IGHA2-a CDS
391	IGHA2-b GENOMIC
392	IGHA2-b CDS

393	IGHD ref GENOMIC
394	IGHD ref- CDS (ensembl transcript ID ENST00000390556)

395	IGHE ref GENOMIC
396	IGHE ref- CDS (ensembl transcript ID ENST00000390541)
397	IGHE ref- CDS (ensembl transcript ID ENST00000576077)
398	IGHE ref- CDS (ensembl transcript ID ENST00000577108)
399	IGHM ref GENOMIC
400	IGHM ref- CDS (ensembl transcript ID ENST00000390559)

401	IGHM-a GENOMIC
402	IGHM-a CDS
403	IGHM-b GENOMIC
404	IGHM-b CDS
405	IGHD7-27 GENOMIC
406	IGHD2-15 GENOMIC
407	IGHD3-16 GENOMIC
408	IGHD6-6 ref GENOMIC
409	IGHD5-18 ref GENOMIC
410	IGHD2-2 ref GENOMIC
411	IGHD4-11 ref GENOMIC
412	IGHD5-12 ref

	GENOMIC
413	IGHD3-3 ref GENOMIC
414	IGHD2-8 ref GENOMIC
415	IGHD4-4 ref GENOMIC
416	IGHD4-23 ref GENOMIC
417	IGHD1-14 ref GENOMIC
418	IGHD3-10 ref GENOMIC
419	IGHD1-26 ref GENOMIC
420	IGHD3-9 ref GENOMIC
421	IGHD1-1 ref GENOMIC
422	IGHD6-25 ref GENOMIC
423	IGHD5-24 ref GENOMIC
424	IGHD2-21 ref GENOMIC
425	IGHD1-20 ref GENOMIC
426	IGHD6-13 ref GENOMIC
427	IGHD4-17 ref GENOMIC
428	IGHD3-22 ref GENOMIC
429	IGHD5-5 ref GENOMIC
430	IGHD6-19 ref GENOMIC
431	IGHD1-7 ref GENOMIC
432	Rabbit JH6 (amino acid)
433	Rabbit JH6 (nucleotide)
434	Sheep JH6 (amino acid)
435	Sheep JH6 (nucleotide)
436	Bovine JH6 (amino acid)
437	Bovine JH6 (nucleotide)
438	Dog JH3 (amino acid)
439	Dog JH3 (nucleotide)
440	Human JH6*02 (amino acid)
441	Human JH6*02 (nucleotide)
442	Leader Sequence (nucleotide)
443	Leader Sequence (amino acid)
444	PRIMER
445	PRIMER

Table 19

	IgG1	IgG2	IgG3	IgG4
Complement activation				
Classical pathway	+++	+	+++	–
Alternative pathway	–	+	–	–
Fc receptor recognition				
FcγRI	+++	–	+++	++
FcγRIIa, 131R/R	++	–	++	–
FcγRIIa, 131H/H	+	+	++	–
FcγRIIb	++	–	++	+
FcγRIII	+	+/–	+	+/–

Table 20

Table 20A: Variants Found In No More Than 3 Human Populations

TYPE	VARIANT	C.FREQ	C.FREQ (%)	#INDIVIDUALS	#POPULATIONS
IGKV	IGKV1D-8-b	0.01785	1.79%	36	3
IGHV	IGHV3-11-a	0.01695	1.70%	36	3
IGHV	IGHV5-51-a	0.0164	1.64%	35	3
IGHV	IGHV2-5-b	0.01055	1.06%	18	3
IGLV	IGLV2-14-g	0.0104	1.04%	22	3
IGHV	IGHV7-81-b	0.01005	1.01%	21	3
IGKV	IGKV1D-42-b	0.01005	1.01%	21	3
IGKV	IGKV3-7-b	0.01	1.00%	19	3
IGHV	IGHV1-69-d	0.00995	1.00%	22	3
IGLV	IGLV2-14-e	0.00995	1.00%	19	3
IGLV	IGLV9-49-a	0.00865	0.87%	18	3
IGLV	IGLV1-44-b	0.00775	0.78%	16	2
IGLV	IGLV2-11-a	0.00775	0.78%	17	3
IGKV	IGKV3-11-d	0.0073	0.73%	14	3
IGKV	IGKV1D-43-b	0.00685	0.69%	14	3
IGHV	IGHV3-38-d	0.0064	0.64%	13	3

IGLJ	IGLJ6-a	0.00595	0.60%	13	3
IGHV	IGHV5-51-b	0.0055	0.55%	11	3
IGLV	IGLV5-45-c	0.00545	0.55%	12	2
IGHV	IGHV4-31-i	0.0054	0.54%	10	3
IGKV	IGKV2D-26-b	0.00455	0.46%	9	3
IGLV	IGLV1-36-c	0.00455	0.46%	10	2
IGHV	IGHV1-24-c	0.00415	0.42%	9	3
IGHV	IGHV1-8-c	0.00405	0.41%	8	2
IGLV	IGLV1-47-d	0.00405	0.41%	9	3
IGHV	IGHV3-43-i	0.00365	0.37%	8	3
IGHV	IGHV3-53-h	0.0036	0.36%	8	2
IGLV	IGLV10-54-e	0.0036	0.36%	8	3
IGLV	IGLV1-36-f	0.0036	0.36%	8	2
IGLV	IGLV5-45-d	0.0036	0.36%	8	3
IGHV	IGHV1-46-b	0.0032	0.32%	7	3
IGHV	IGHV2-26-c	0.0032	0.32%	7	2
IGLJ	IGLJ6-b	0.0032	0.32%	7	3
IGLV	IGLV3-32-a	0.0032	0.32%	7	3
IGLV	IGLV9-49-b	0.0032	0.32%	7	2
IGHJ	IGHJ6-b	0.00315	0.32%	7	2
IGHV	IGHV3-16-d	0.00315	0.32%	7	3
IGHV	IGHV4-28-k	0.00315	0.32%	7	3
IGKJ	IGKJ2-d	0.00315	0.32%	6	3
IGLV	IGLV10-54-f	0.00315	0.32%	7	2
IGLV	IGLV2-8-c	0.00315	0.32%	7	3
IGLV	IGLV3-10-c	0.00315	0.32%	6	2
IGLV	IGLV1-40-c	0.00275	0.28%	6	2
IGLV	IGLV1-51-b	0.00275	0.28%	6	3
IGHV	IGHV1-69-m	0.0027	0.27%	6	3
IGHV	IGHV3-33-e	0.0027	0.27%	6	3
IGHV	IGHV3-49-c	0.0027	0.27%	6	3
IGHV	IGHV3-66-g	0.0027	0.27%	6	3
IGHV	IGHV4-4-u	0.0027	0.27%	6	3
IGKV	IGKV1D-42-c	0.0027	0.27%	6	1
IGKV	IGKV1D-43-c	0.0027	0.27%	5	3
IGLV	IGLV10-54-h	0.0027	0.27%	6	2
IGLV	IGLV10-54-i	0.0027	0.27%	6	2
IGLV	IGLV1-36-g	0.0027	0.27%	6	3
IGLV	IGLV2-8-d	0.0027	0.27%	6	3
IGLV	IGLV3-22-e	0.0027	0.27%	6	1
IGLV	IGLV3-22-d	0.0027	0.27%	6	3
IGHJ	IGHJ2-b	0.00225	0.23%	4	3
IGHV	IGHV1-3-u	0.00225	0.23%	5	2
IGHV	IGHV1-3-v	0.00225	0.23%	5	2
IGHV	IGHV1-3-ap	0.00225	0.23%	4	3
IGHV	IGHV1-46-c	0.00225	0.23%	5	3

IGHV	IGHV1-58-f	0.00225	0.23%	5	3
IGHV	IGHV1-69-j	0.00225	0.23%	5	3
IGHV	IGHV3-13-e	0.00225	0.23%	5	3
IGHV	IGHV3-23-f	0.00225	0.23%	5	3
IGHV	IGHV3-23-g	0.00225	0.23%	5	3
IGHV	IGHV4-28-p	0.00225	0.23%	5	3
IGHV	IGHV4-31-j	0.00225	0.23%	5	3
IGKV	IGKV1-17-d	0.00225	0.23%	4	1
IGKV	IGKV3-20-a	0.00225	0.23%	5	3
IGKV	IGKV3-7-c	0.00225	0.23%	5	3
IGKV	IGKV3-7-d	0.00225	0.23%	5	3
IGKV	IGKV3D-11-b	0.00225	0.23%	5	3
IGLJ	IGLJ3-a	0.00225	0.23%	4	2
IGLV	IGLV11-55-g	0.00225	0.23%	5	3
IGLV	IGLV1-50-c	0.00225	0.23%	5	3
IGLV	IGLV2-11-g	0.00225	0.23%	5	3
IGLV	IGLV2-14-n	0.00225	0.23%	5	3
IGLV	IGLV2-23-f	0.00225	0.23%	5	2
IGLV	IGLV3-19-c	0.00225	0.23%	5	2
IGLV	IGLV5-37-c	0.00225	0.23%	4	2
IGLV	IGLV5-48-k	0.00225	0.23%	5	3
IGLV	IGLV2-14-i	0.00185	0.19%	4	2
IGHJ	IGHJ1-a	0.0018	0.18%	3	3
IGHV	IGHV1-2-g	0.0018	0.18%	4	3
IGHV	IGHV1-3-y	0.0018	0.18%	4	3
IGHV	IGHV1-3-ar	0.0018	0.18%	4	3
IGHV	IGHV1-3-ai	0.0018	0.18%	4	3
IGHV	IGHV1-69-s	0.0018	0.18%	4	3
IGHV	IGHV1-8-d	0.0018	0.18%	4	2
IGHV	IGHV2-5-d	0.0018	0.18%	4	2
IGHV	IGHV3-13-j	0.0018	0.18%	4	2
IGHV	IGHV3-21-d	0.0018	0.18%	4	3
IGHV	IGHV3-23-n	0.0018	0.18%	4	3
IGHV	IGHV3-33-d	0.0018	0.18%	4	3
IGHV	IGHV3-64-e	0.0018	0.18%	4	2
IGHV	IGHV3-9-g	0.0018	0.18%	4	3
IGHV	IGHV4-39-y	0.0018	0.18%	4	3
IGHV	IGHV4-39-n	0.0018	0.18%	4	3
IGHV	IGHV4-39-o	0.0018	0.18%	4	3
IGHV	IGHV7-81-c	0.0018	0.18%	4	3
IGKJ	IGKJ2-e	0.0018	0.18%	4	3
IGKV	IGKV2D-26-c	0.0018	0.18%	4	1
IGKV	IGKV3D-11-g	0.0018	0.18%	3	2
IGKV	IGKV3D-11-f	0.0018	0.18%	4	3
IGKV	IGKV4-1-w	0.0018	0.18%	3	3
IGLJ	IGLJ1-a	0.0018	0.18%	4	2

IGLV	IGLV10-54-k	0.0018	0.18%	4	2
IGLV	IGLV11-55-e	0.0018	0.18%	4	3
IGLV	IGLV1-36-e	0.0018	0.18%	4	3
IGLV	IGLV1-40-f	0.0018	0.18%	4	3
IGLV	IGLV1-44-d	0.0018	0.18%	4	3
IGLV	IGLV1-44-e	0.0018	0.18%	4	3
IGLV	IGLV1-47-g	0.0018	0.18%	4	2
IGLV	IGLV1-50-d	0.0018	0.18%	4	2
IGLV	IGLV1-51-d	0.0018	0.18%	4	2
IGLV	IGLV1-51-c	0.0018	0.18%	4	3
IGLV	IGLV2-14-r	0.0018	0.18%	4	1
IGLV	IGLV2-33-b	0.0018	0.18%	4	3
IGLV	IGLV2-8-e	0.0018	0.18%	4	2
IGLV	IGLV3-16-h	0.0018	0.18%	4	3
IGLV	IGLV3-21-j	0.0018	0.18%	4	3
IGLV	IGLV3-22-g	0.0018	0.18%	4	3
IGLV	IGLV3-22-h	0.0018	0.18%	4	2
IGLV	IGLV5-48-j	0.0018	0.18%	4	3
IGHJ	IGHJ4-d	0.00135	0.14%	3	3
IGHV	IGHV1-18-d	0.00135	0.14%	3	3
IGHV	IGHV1-18-e	0.00135	0.14%	3	3
IGHV	IGHV1-2-h	0.00135	0.14%	3	2
IGHV	IGHV1-2-f	0.00135	0.14%	3	3
IGHV	IGHV1-3-aa	0.00135	0.14%	3	1
IGHV	IGHV1-3-ak	0.00135	0.14%	3	3
IGHV	IGHV1-3-w	0.00135	0.14%	3	1
IGHV	IGHV1-3-at	0.00135	0.14%	3	2
IGHV	IGHV1-69-k	0.00135	0.14%	3	1
IGHV	IGHV1-69-ac	0.00135	0.14%	3	3
IGHV	IGHV1-8-e	0.00135	0.14%	3	2
IGHV	IGHV1-8-f	0.00135	0.14%	3	2
IGHV	IGHV2-70-y	0.00135	0.14%	3	3
IGHV	IGHV2-70-z	0.00135	0.14%	3	3
IGHV	IGHV2-70-af	0.00135	0.14%	3	2
IGHV	IGHV2-70-h	0.00135	0.14%	3	3
IGHV	IGHV3-13-g	0.00135	0.14%	3	2
IGHV	IGHV3-16-e	0.00135	0.14%	3	2
IGHV	IGHV3-20-f	0.00135	0.14%	3	3
IGHV	IGHV3-30-d	0.00135	0.14%	3	2
IGHV	IGHV3-30-c	0.00135	0.14%	3	3
IGHV	IGHV3-33-h	0.00135	0.14%	3	3
IGHV	IGHV3-43-k	0.00135	0.14%	3	2
IGHV	IGHV3-43-m	0.00135	0.14%	3	3
IGHV	IGHV3-53-n	0.00135	0.14%	3	2
IGHV	IGHV3-53-m	0.00135	0.14%	3	3
IGHV	IGHV3-64-f	0.00135	0.14%	3	1

IGHV	IGHV3-73-c	0.00135	0.14%	3	3
IGHV	IGHV3-9-h	0.00135	0.14%	2	1
IGHV	IGHV4-31-h	0.00135	0.14%	3	2
IGHV	IGHV4-39-u	0.00135	0.14%	3	3
IGHV	IGHV4-39-t	0.00135	0.14%	3	2
IGHV	IGHV4-39-q	0.00135	0.14%	3	3
IGHV	IGHV4-4-ao	0.00135	0.14%	3	3
IGHV	IGHV4-4-ac	0.00135	0.14%	3	1
IGHV	IGHV4-4-ae	0.00135	0.14%	3	2
IGHV	IGHV5-51-g	0.00135	0.14%	3	3
IGHV	IGHV7-81-e	0.00135	0.14%	3	1
IGKJ	IGKJ2-f	0.00135	0.14%	3	2
IGKJ	IGKJ3-d	0.00135	0.14%	3	3
IGKJ	IGKJ4-g	0.00135	0.14%	3	3
IGKJ	IGKJ5-c	0.00135	0.14%	3	3
IGKJ	IGKJ5-b	0.00135	0.14%	3	3
IGKV	IGKV1-5-c	0.00135	0.14%	3	3
IGKV	IGKV1-9-j	0.00135	0.14%	3	3
IGKV	IGKV1D-17-c	0.00135	0.14%	3	2
IGKV	IGKV1D-17-b	0.00135	0.14%	3	2
IGKV	IGKV1D-42-e	0.00135	0.14%	3	1
IGKV	IGKV1D-42-d	0.00135	0.14%	3	2
IGKV	IGKV2D-24-b	0.00135	0.14%	3	2
IGKV	IGKV2D-30-c	0.00135	0.14%	3	2
IGKV	IGKV3-20-d	0.00135	0.14%	3	3
IGKV	IGKV3-20-f	0.00135	0.14%	3	2
IGKV	IGKV3D-11-d	0.00135	0.14%	3	2
IGKV	IGKV4-1-r	0.00135	0.14%	3	3
IGKV	IGKV4-1-x	0.00135	0.14%	3	2
IGKV	IGKV4-1-e	0.00135	0.14%	3	3
IGKV	IGKV4-1-o	0.00135	0.14%	3	2
IGLJ	IGLJ6-c	0.00135	0.14%	3	2
IGLJ	IGLJ7-b	0.00135	0.14%	3	3
IGLV	IGLV10-54-g	0.00135	0.14%	3	1
IGLV	IGLV11-55-h	0.00135	0.14%	3	2
IGLV	IGLV11-55-f	0.00135	0.14%	3	2
IGLV	IGLV1-40-g	0.00135	0.14%	3	3
IGLV	IGLV1-44-f	0.00135	0.14%	3	2
IGLV	IGLV1-47-f	0.00135	0.14%	3	2
IGLV	IGLV2-23-e	0.00135	0.14%	3	3
IGLV	IGLV3-12-d	0.00135	0.14%	3	3
IGLV	IGLV3-16-d	0.00135	0.14%	3	3
IGLV	IGLV3-19-d	0.00135	0.14%	3	1
IGLV	IGLV3-19-l	0.00135	0.14%	3	2
IGLV	IGLV3-1-ag	0.00135	0.14%	3	3
IGLV	IGLV3-1-y	0.00135	0.14%	3	3

IGLV	IGLV3-1-s	0.00135	0.14%	3	3
IGLV	IGLV3-1-x	0.00135	0.14%	3	3
IGLV	IGLV3-1-ad	0.00135	0.14%	3	2
IGLV	IGLV3-1-l	0.00135	0.14%	3	3
IGLV	IGLV3-1-p	0.00135	0.14%	3	2
IGLV	IGLV3-21-t	0.00135	0.14%	3	2
IGLV	IGLV3-22-f	0.00135	0.14%	3	3
IGLV	IGLV3-22-i	0.00135	0.14%	3	1
IGLV	IGLV3-25-e	0.00135	0.14%	3	3
IGLV	IGLV5-48-t	0.00135	0.14%	3	2
IGLV	IGLV5-48-s	0.00135	0.14%	3	3
IGLV	IGLV5-52-d	0.00135	0.14%	3	2
IGLV	IGLV5-52-c	0.00135	0.14%	3	1
IGLV	IGLV7-43-c	0.00135	0.14%	3	2
IGHJ	IGHJ5-d	0.0009	0.09%	2	2
IGHJ	IGHJ6-e	0.0009	0.09%	2	2
IGHJ	IGHJ6-h	0.0009	0.09%	2	2
IGHJ	IGHJ6-f	0.0009	0.09%	2	2
IGHV	IGHV1-18-f	0.0009	0.09%	2	2
IGHV	IGHV1-24-g	0.0009	0.09%	2	1
IGHV	IGHV1-2-i	0.0009	0.09%	2	2
IGHV	IGHV1-2-m	0.0009	0.09%	2	2
IGHV	IGHV1-2-l	0.0009	0.09%	2	2
IGHV	IGHV1-3-ag	0.0009	0.09%	2	2
IGHV	IGHV1-3-aj	0.0009	0.09%	2	2
IGHV	IGHV1-3-aq	0.0009	0.09%	2	1
IGHV	IGHV1-3-s	0.0009	0.09%	2	2
IGHV	IGHV1-3-q	0.0009	0.09%	2	1
IGHV	IGHV1-3-z	0.0009	0.09%	2	1
IGHV	IGHV1-3-ad	0.0009	0.09%	2	2
IGHV	IGHV1-3-aw	0.0009	0.09%	2	2
IGHV	IGHV1-3-ae	0.0009	0.09%	2	2
IGHV	IGHV1-3-an	0.0009	0.09%	2	2
IGHV	IGHV1-46-k	0.0009	0.09%	2	2
IGHV	IGHV1-46-e	0.0009	0.09%	2	2
IGHV	IGHV1-46-n	0.0009	0.09%	2	2
IGHV	IGHV1-58-g	0.0009	0.09%	2	1
IGHV	IGHV1-69-ah	0.0009	0.09%	2	2
IGHV	IGHV1-69-t	0.0009	0.09%	2	2
IGHV	IGHV1-69-ao	0.0009	0.09%	2	1
IGHV	IGHV1-69-af	0.0009	0.09%	2	2
IGHV	IGHV1-69-ad	0.0009	0.09%	2	2
IGHV	IGHV1-69-n	0.0009	0.09%	2	2
IGHV	IGHV1-69-an	0.0009	0.09%	2	2
IGHV	IGHV1-69-o	0.0009	0.09%	2	2
IGHV	IGHV1-8-h	0.0009	0.09%	2	2

IGHV	IGHV2-26-f	0.0009	0.09%	2	2
IGHV	IGHV2-5-c	0.0009	0.09%	2	2
IGHV	IGHV2-70-g	0.0009	0.09%	2	2
IGHV	IGHV2-70-s	0.0009	0.09%	2	2
IGHV	IGHV2-70-ab	0.0009	0.09%	2	2
IGHV	IGHV3-11-b	0.0009	0.09%	2	2
IGHV	IGHV3-13-f	0.0009	0.09%	2	2
IGHV	IGHV3-20-d	0.0009	0.09%	2	2
IGHV	IGHV3-21-c	0.0009	0.09%	2	2
IGHV	IGHV3-21-h	0.0009	0.09%	2	2
IGHV	IGHV3-23-j	0.0009	0.09%	2	1
IGHV	IGHV3-23-h	0.0009	0.09%	2	2
IGHV	IGHV3-33-f	0.0009	0.09%	2	2
IGHV	IGHV3-43-p	0.0009	0.09%	2	2
IGHV	IGHV3-64-h	0.0009	0.09%	2	1
IGHV	IGHV3-66-h	0.0009	0.09%	2	2
IGHV	IGHV3-74-c	0.0009	0.09%	2	2
IGHV	IGHV3-9-o	0.0009	0.09%	2	1
IGHV	IGHV4-28-s	0.0009	0.09%	2	2
IGHV	IGHV4-28-n	0.0009	0.09%	2	2
IGHV	IGHV4-31-t	0.0009	0.09%	2	2
IGHV	IGHV4-31-m	0.0009	0.09%	2	2
IGHV	IGHV4-31-l	0.0009	0.09%	2	2
IGHV	IGHV4-34-d	0.0009	0.09%	2	2
IGHV	IGHV4-34-g	0.0009	0.09%	2	2
IGHV	IGHV4-34-b	0.0009	0.09%	2	2
IGHV	IGHV4-39-aa	0.0009	0.09%	2	2
IGHV	IGHV4-39-s	0.0009	0.09%	2	2
IGHV	IGHV4-39-x	0.0009	0.09%	2	2
IGHV	IGHV4-39-p	0.0009	0.09%	2	2
IGHV	IGHV4-4-aa	0.0009	0.09%	2	2
IGHV	IGHV4-4-af	0.0009	0.09%	2	2
IGHV	IGHV4-4-x	0.0009	0.09%	2	2
IGHV	IGHV4-4-ad	0.0009	0.09%	2	2
IGHV	IGHV4-4-ai	0.0009	0.09%	2	1
IGHV	IGHV5-51-d	0.0009	0.09%	2	2
IGHV	IGHV5-51-k	0.0009	0.09%	2	2
IGHV	IGHV5-51-c	0.0009	0.09%	2	2
IGHV	IGHV6-1-a	0.0009	0.09%	2	2
IGHV	IGHV6-1-c	0.0009	0.09%	2	1
IGHV	IGHV7-81-d	0.0009	0.09%	2	2
IGKJ	IGKJ2-g	0.0009	0.09%	2	2
IGKJ	IGKJ3-b	0.0009	0.09%	2	2
IGKJ	IGKJ4-j	0.0009	0.09%	2	1
IGKJ	IGKJ4-c	0.0009	0.09%	2	2
IGKJ	IGKJ4-k	0.0009	0.09%	2	2

IGKJ	IGKJ4-h	0.0009	0.09%	2	2
IGKV	IGKV1-16-d	0.0009	0.09%	2	2
IGKV	IGKV1-17-e	0.0009	0.09%	2	2
IGKV	IGKV1-5-d	0.0009	0.09%	2	1
IGKV	IGKV1-8-h	0.0009	0.09%	2	2
IGKV	IGKV1-9-i	0.0009	0.09%	2	2
IGKV	IGKV1D-16-c	0.0009	0.09%	2	1
IGKV	IGKV1D-16-b	0.0009	0.09%	2	1
IGKV	IGKV1D-42-g	0.0009	0.09%	2	2
IGKV	IGKV1D-42-f	0.0009	0.09%	2	1
IGKV	IGKV1D-43-e	0.0009	0.09%	2	1
IGKV	IGKV1D-43-d	0.0009	0.09%	2	2
IGKV	IGKV1D-8-d	0.0009	0.09%	2	1
IGKV	IGKV1D-8-f	0.0009	0.09%	2	2
IGKV	IGKV2D-26-d	0.0009	0.09%	2	2
IGKV	IGKV2D-26-f	0.0009	0.09%	2	2
IGKV	IGKV2D-29-b	0.0009	0.09%	2	1
IGKV	IGKV3-11-g	0.0009	0.09%	2	1
IGKV	IGKV3-11-f	0.0009	0.09%	2	2
IGKV	IGKV3-20-j	0.0009	0.09%	2	2
IGKV	IGKV3-20-i	0.0009	0.09%	2	2
IGKV	IGKV3-20-e	0.0009	0.09%	2	2
IGKV	IGKV3-20-c	0.0009	0.09%	2	2
IGKV	IGKV3D-11-h	0.0009	0.09%	2	1
IGKV	IGKV3D-11-i	0.0009	0.09%	2	1
IGKV	IGKV3D-20-d	0.0009	0.09%	2	2
IGKV	IGKV4-1-y	0.0009	0.09%	2	2
IGKV	IGKV4-1-g	0.0009	0.09%	2	2
IGKV	IGKV4-1-t	0.0009	0.09%	2	2
IGKV	IGKV4-1-i	0.0009	0.09%	2	2
IGKV	IGKV4-1-n	0.0009	0.09%	2	2
IGKV	IGKV4-1-l	0.0009	0.09%	2	2
IGLJ	IGLJ2-e	0.0009	0.09%	2	2
IGLJ	IGLJ2-c	0.0009	0.09%	2	2
IGLJ	IGLJ5-b	0.0009	0.09%	2	1
IGLV	IGLV10-54-n	0.0009	0.09%	2	2
IGLV	IGLV10-54-j	0.0009	0.09%	2	2
IGLV	IGLV1-40-h	0.0009	0.09%	2	2
IGLV	IGLV1-40-e	0.0009	0.09%	2	2
IGLV	IGLV1-47-e	0.0009	0.09%	2	2
IGLV	IGLV1-51-g	0.0009	0.09%	2	2
IGLV	IGLV2-11-j	0.0009	0.09%	2	2
IGLV	IGLV2-11-f	0.0009	0.09%	2	2
IGLV	IGLV2-11-i	0.0009	0.09%	2	2
IGLV	IGLV2-11-e	0.0009	0.09%	2	1
IGLV	IGLV2-14-p	0.0009	0.09%	2	2

IGLV	IGLV2-14-q	0.0009	0.09%	2	1
IGLV	IGLV2-14-o	0.0009	0.09%	2	1
IGLV	IGLV2-23-k	0.0009	0.09%	2	2
IGLV	IGLV2-23-o	0.0009	0.09%	2	2
IGLV	IGLV2-33-c	0.0009	0.09%	2	1
IGLV	IGLV3-19-j	0.0009	0.09%	2	2
IGLV	IGLV3-19-k	0.0009	0.09%	2	2
IGLV	IGLV3-19-h	0.0009	0.09%	2	1
IGLV	IGLV3-19-i	0.0009	0.09%	2	1
IGLV	IGLV3-19-p	0.0009	0.09%	2	2
IGLV	IGLV3-1-ah	0.0009	0.09%	2	2
IGLV	IGLV3-1-d	0.0009	0.09%	2	2
IGLV	IGLV3-1-j	0.0009	0.09%	2	1
IGLV	IGLV3-1-g	0.0009	0.09%	2	2
IGLV	IGLV3-1-z	0.0009	0.09%	2	2
IGLV	IGLV3-1-ac	0.0009	0.09%	2	2
IGLV	IGLV3-1-ae	0.0009	0.09%	2	2
IGLV	IGLV3-21-k	0.0009	0.09%	2	2
IGLV	IGLV3-21-r	0.0009	0.09%	2	1
IGLV	IGLV3-21-i	0.0009	0.09%	2	2
IGLV	IGLV3-22-j	0.0009	0.09%	2	1
IGLV	IGLV3-22-n	0.0009	0.09%	2	2
IGLV	IGLV3-22-m	0.0009	0.09%	2	2
IGLV	IGLV3-22-o	0.0009	0.09%	2	2
IGLV	IGLV3-25-g	0.0009	0.09%	2	1
IGLV	IGLV3-25-l	0.0009	0.09%	2	2
IGLV	IGLV3-27-a	0.0009	0.09%	2	2
IGLV	IGLV3-32-c	0.0009	0.09%	2	2
IGLV	IGLV3-9-a	0.0009	0.09%	2	2
IGLV	IGLV3-9-b	0.0009	0.09%	2	2
IGLV	IGLV4-3-d	0.0009	0.09%	2	2
IGLV	IGLV4-60-g	0.0009	0.09%	2	2
IGLV	IGLV4-60-n	0.0009	0.09%	2	2
IGLV	IGLV4-60-o	0.0009	0.09%	2	2
IGLV	IGLV5-45-g	0.0009	0.09%	2	2
IGLV	IGLV5-48-u	0.0009	0.09%	2	2
IGLV	IGLV5-48-q	0.0009	0.09%	2	2
IGLV	IGLV5-48-r	0.0009	0.09%	2	2
IGLV	IGLV5-48-n	0.0009	0.09%	2	2
IGLV	IGLV5-48-m	0.0009	0.09%	2	1
IGLV	IGLV5-48-l	0.0009	0.09%	2	2
IGLV	IGLV5-52-b	0.0009	0.09%	2	1
IGLV	IGLV7-46-c	0.0009	0.09%	2	2
IGLV	IGLV9-49-c	0.0009	0.09%	2	2

Table 20B: Variants Found In No More Than 3 Human Populations & In At Least 10 Individuals

TYPE	VARIANT	C.FREQ	C.FREQ (%)	#INDIVIDUALS	#POPULATIONS
IGKV	IGKV1D-8-b	0.01785	1.79%	36	3
IGHV	IGHV3-11-a	0.01695	1.70%	36	3
IGHV	IGHV5-51-a	0.0164	1.64%	35	3
IGHV	IGHV2-5-b	0.01055	1.06%	18	3
IGLV	IGLV2-14-g	0.0104	1.04%	22	3
IGHV	IGHV7-81-b	0.01005	1.01%	21	3
IGKV	IGKV1D-42-b	0.01005	1.01%	21	3
IGKV	IGKV3-7-b	0.01	1.00%	19	3
IGHV	IGHV1-69-d	0.00995	1.00%	22	3
IGLV	IGLV2-14-e	0.00995	1.00%	19	3
IGLV	IGLV9-49-a	0.00865	0.87%	18	3
IGLV	IGLV1-44-b	0.00775	0.78%	16	2
IGLV	IGLV2-11-a	0.00775	0.78%	17	3
IGKV	IGKV3-11-d	0.0073	0.73%	14	3
IGKV	IGKV1D-43-b	0.00685	0.69%	14	3
IGHV	IGHV3-38-d	0.0064	0.64%	13	3
IGLJ	IGLJ6-a	0.00595	0.60%	13	3
IGHV	IGHV5-51-b	0.0055	0.55%	11	3
IGLV	IGLV5-45-c	0.00545	0.55%	12	2
IGHV	IGHV4-31-i	0.0054	0.54%	10	3
IGLV	IGLV1-36-c	0.00455	0.46%	10	2

Table 20C: Variants Found In No More Than 3 Human Populations & In At Least 20 Individuals

TYPE	VARIANT	C.FREQ	C.FREQ (%)	#INDIVIDUALS	#POPULATIONS
IGKV	IGKV1D-8-b	0.01785	1.79%	36	3
IGHV	IGHV3-11-a	0.01695	1.70%	36	3
IGHV	IGHV5-51-a	0.0164	1.64%	35	3
IGLV	IGLV2-14-g	0.0104	1.04%	22	3
IGHV	IGHV7-81-b	0.01005	1.01%	21	3
IGKV	IGKV1D-42-b	0.01005	1.01%	21	3
IGHV	IGHV1-69-d	0.00995	1.00%	22	3

Table 20D: Variants Found In No More Than 3 Human Populations & In At Least 30 Individuals

TYPE	VARIANT	C.FREQ	C.FREQ (%)	#INDIVIDUALS	#POPULATIONS
IGKV	IGKV1D-8-b	0.01785	1.79%	36	3
IGHV	IGHV3-11-a	0.01695	1.70%	36	3
IGHV	IGHV5-51-a	0.0164	1.64%	35	3

Table 21

Seq ID #	Associated Gene Name	Ensembl Gene ID	Gene Location [Chromosome Name:Gene Chr Start (bp)-GeneChr End (bp)]	Ensembl Transcript ID	Exon Positions [Chromosome Name:Exon Chr Start (bp)-Exon Chr End (bp)]
239	IGHJ1	ENSG00000211905	14:106331617-106331668	ENST00000390565	14:106331617-106331668
240	IGHJ2	ENSG00000211904	14:106331409-106331460	ENST00000390564	14:106331409-106331460
241	IGHJ3	ENSG00000242887	14:106330797-106330845	ENST00000463911	14:106330797-106330845
242	IGHJ4	ENSG00000240041	14:106330425-106330470	ENST00000461719	14:106330425-106330470
243	IGHJ5	ENSG00000242472	14:106330024-106330072	ENST00000488476	14:106330024-106330072
244	IGHJ6	ENSG00000211900	14:106329408-106329468	ENST00000390560	14:106329408-106329468
245	IGHV1-18	ENSG00000211945	14:106641563-106642056	ENST00000390605	14:106641952-106642056 14:106641563-106641867
246	IGHV1-2	ENSG00000211934	14:106452671-106453170	ENST00000390594	14:106453061-106453170 14:106452671-106452975
247	IGHV1-24	ENSG00000211950	14:106733144-106733639	ENST00000390610	14:106733534-106733639 14:106733144-106733448
248	IGHV1-3	ENSG00000211935	14:106471246-106471723	ENST00000390595	14:106471636-106471723 14:106471246-106471550
249	IGHV1-45	ENSG00000211961	14:106962931-106963424	ENST00000390621	14:106963321-106963424 14:106962931-106963235
250	IGHV1-46	ENSG00000211962	14:106967049-106967788	ENST00000390622	14:106967439-106967788 14:106967049-106967353
251	IGHV1-58	ENSG00000211968	14:107078373-107078869	ENST00000390628	14:107078763-107078869 14:107078373-107078677

252	IGHV1-69	ENSG00000211973	14:107169931-107170428	ENST00000390633	14:107170322-107170428 14:107169931-107170235
253	IGHV1-8	ENSG00000211939	14:106539079-106539577	ENST00000390599	14:106539470-106539577 14:106539079-106539383
254	IGHV2-26	ENSG00000211951	14:106757650-106758116	ENST00000390611	14:106758047-106758116 14:106757650-106757960
255	IGHV2-5	ENSG00000211937	14:106494135-106494597	ENST00000390597	14:106494532-106494597 14:106494135-106494445
256	IGHV2-70	ENSG00000211974	14:107178820-107179338	ENST00000390634	14:107179217-107179338 14:107178820-107179130
257	IGHV3-11	ENSG00000211941	14:106573233-106573800	ENST00000390601	14:106573635-106573800 14:106573233-106573537
258	IGHV3-13	ENSG00000211942	14:106586137-106586667	ENST00000390602	14:106586542-106586667 14:106586137-106586438
259	IGHV3-15	ENSG00000211943	14:106610313-106610852	ENST00000390603	14:106610727-106610852 14:106610313-106610623
260	IGHV3-16	ENSG00000211944	14:106621894-106622419	ENST00000390604	14:106622300-106622419 14:106621894-106622198
261	IGHV3-20	ENSG00000211946	14:106667581-106668095	ENST00000390606	14:106667988-106668095 14:106667581-106667885
262	IGHV3-21	ENSG00000211947	14:106691673-106692203	ENST00000390607	14:106692079-106692203 14:106691673-106691977
263	IGHV3-23	ENSG00000211949	14:106725201-106725733	ENST00000390609	14:106725609-106725733 14:106725201-106725505
264	IGHV3-30	ENSG00000211953	14:106791005-106791536	ENST00000390613	14:106791411-106791536 14:106791005-106791309
265	IGHV3-33	ENSG00000211955	14:106815722-106816253	ENST00000390615	14:106816128-106816253 14:106815722-106816026
266	IGHV3-35	ENSG00000211957	14:106845323-106845789	ENST00000390617	14:106845729-106845789 14:106845323-106845627
267	IGHV3-38	ENSG00000211958	14:106866406-106866934	ENST00000390618	14:106866811-106866934 14:106866406-106866707
268	IGHV3-43	ENSG00000232216	14:106926188-106926724	ENST00000434710	14:106926599-106926724 14:106926188-106926495

269	IGHV3-48	ENSG00000211964	14:106993814-106994346	ENST00000390624	14:106994222-106994346 14:106993814-106994118
270	IGHV3-49	ENSG00000211965	14:107012938-107013477	ENST00000390625	14:107013352-107013477 14:107012938-107013248
271	IGHV3-53	ENSG00000211967	14:107048672-107049341	ENST00000390627	14:107049075-107049341 14:107048672-107048973
272	IGHV3-64	ENSG00000223648	14:107113741-107114274	ENST00000454421	14:107114149-107114274 14:107113741-107114045
273	IGHV3-66	ENSG00000211972	14:107131033-107131560	ENST00000390632	14:107131436-107131560 14:107131033-107131334
274	IGHV3-7	ENSG00000211938	14:106518400-106518932	ENST00000390598	14:106518808-106518932 14:106518400-106518704
275	IGHV3-72	ENSG00000225698	14:107198932-107199471	ENST00000433072	14:107199346-107199471 14:107198932-107199242
276	IGHV3-73	ENSG00000211976	14:107210932-107211471	ENST00000390636	14:107211346-107211471 14:107210932-107211242
277	IGHV3-74	ENSG00000224650	14:107218676-107219365	ENST00000424969	14:107219084-107219365 14:107218676-107218980
278	IGHV3-9	ENSG00000211940	14:106552285-106552809	ENST00000390600	14:106552684-106552809 14:106552285-106552592
279	IGHV4-28	ENSG00000211952	14:106780513-106781017	ENST00000390612	14:106780900-106781017 14:106780513-106780817
280	IGHV4-31	ENSG00000231475	14:106805209-106805716	ENST00000438142	14:106805599-106805716 14:106805209-106805516
281	IGHV4-34	ENSG00000211956	14:106829594-106830076	ENST00000390616	14:106829979-106830076 14:106829594-106829895
282	IGHV4-39	ENSG00000211959	14:106877619-106878126	ENST00000390619	14:106878010-106878126 14:106877619-106877926
283	IGHV4-4	ENSG00000211936	14:106478110-106478603	ENST00000390596	14:106478494-106478603 14:106478110-106478411
284	IGHV4-59	ENSG00000224373	14:107081806-107083830	ENST00000390629	14:107083640-107083830 14:107083256-107083557
284	IGHV4-59	ENSG00000224373	14:107081806-107083830	ENST00000455737	14:107083640-107083725 14:107083240-107083557

					14:107081806-107082728
285	IGHV4-61	ENSG00000211970	14:107095126-107095662	ENST00000390630	14:107095516-107095662 14:107095126-107095433
286	IGHV5-51	ENSG00000211966	14:107034729-107035221	ENST00000390626	14:107035117-107035221 14:107034729-107035033
287	IGHV6-1	ENSG00000211933	14:106405611-106406108	ENST00000390593	14:106406008-106406108 14:106405611-106405924
288	IGHV7-81	ENSG00000211979	14:107282792-107283280	ENST00000390639	14:107283181-107283280 14:107282792-107283096
289	IGKJ2	ENSG00000211596	2:89161037-89161074	ENST00000390241	2:89161037-89161074
290	IGKJ3	ENSG00000211595	2:89160733-89160770	ENST00000390240	2:89160733-89160770
291	IGKJ4	ENSG00000211594	2:89160398-89160434	ENST00000390239	2:89160398-89160434
292	IGKJ5	ENSG00000211593	2:89160080-89160117	ENST00000390238	2:89160080-89160117
293	IGKV1-16	ENSG00000240864	2:89399352-89399854	ENST00000479981	2:89399773-89399854 2:89399352-89399647
294	IGKV1-17	ENSG00000240382	2:89416833-89417335	ENST00000490686	2:89417254-89417335 2:89416833-89417128
295	IGKV1-5	ENSG00000243466	2:89246819-89247475	ENST00000496168	2:89247240-89247475 2:89246819-89247114
296	IGKV1-6	ENSG00000239855	2:89265781-89266286	ENST00000464162	2:89266203-89266286 2:89265781-89266076
297	IGKV1-8	ENSG00000240671	2:89291928-89292450	ENST00000495489	2:89292349-89292450 2:89291928-89292223
298	IGKV1-9	ENSG00000241755	2:89309479-89310012	ENST00000493819	2:89309900-89310012 2:89309479-89309774
299	IGKV1D-12	ENSG00000240834	2:90198535-90199190	ENST00000390276	2:90198535-90198770 2:90198895-90199190
300	IGKV1D-16	ENSG00000241244	2:90139078-90139580	ENST00000492446	2:90139078-90139159 2:90139285-90139580

301	IGKV1D-17	ENSG00000242766	2:90121477-90122133	ENST00000483379	2:90121477-90121712 2:90121838-90122133
302	IGKV1D-42	ENSG00000211633	2:90229045-90229531	ENST00000390278	2:90229045-90229119 2:90229236-90229531
303	IGKV1D-43	ENSG00000242580	2:90248739-90249395	ENST00000468879	2:90248739-90248974 2:90249100-90249395
304	IGKV1D-8	ENSG00000239819	2:90259593-90260248	ENST00000471857	2:90259593-90259828 2:90259953-90260248
305	IGKV2-24	ENSG00000241294	2:89475812-89476644	ENST00000484817	2:89476566-89476644 2:89475812-89476122
306	IGKV2-30	ENSG00000243238	2:89544264-89545079	ENST00000468494	2:89545001-89545079 2:89544264-89544574
307	IGKV2D-24	ENSG00000241566	2:90043607-90044439	ENST00000462693	2:90043607-90043685 2:90044129-90044439
308	IGKV2D-26	ENSG00000211623	2:90024732-90025512	ENST00000390268	2:90024732-90024810 2:90025202-90025512
309	IGKV2D-29	ENSG00000243264	2:89986322-89987079	ENST00000491977	2:89986322-89986400 2:89986769-89987079
310	IGKV2D-30	ENSG00000239571	2:89975669-89976489	ENST00000474213	2:89975669-89975752 2:89976179-89976489
311	IGKV3-11	ENSG00000241351	2:89326668-89327228	ENST00000483158	2:89327133-89327228 2:89326668-89326963
312	IGKV3-20	ENSG00000239951	2:89442057-89442643	ENST00000492167	2:89442543-89442643 2:89442057-89442355
313	IGKV3-7	ENSG00000243063	2:89277987-89278600	ENST00000390247	2:89278455-89278600 2:89277987-89278285
314	IGKV3D-11	ENSG00000211632	2:90211643-90212253	ENST00000390277	2:90211643-90211788 2:90211958-90212253
315	IGKV3D-15	ENSG00000224041	2:90153696-90154258	ENST00000417279	2:90153696-90153793 2:90153963-90154258
316	IGKV3D-20	ENSG00000211625	2:90077680-90078311	ENST00000390270	2:90077680-90077825 2:90078013-90078311
317	IGKV4-1	ENSG00000211598	2:89184913-89185669	ENST00000390243	2:89184913-89185136 2:89185356-89185669

318	IGKV5-2	ENSG00000211599	2:89196748-89197300	ENST00000390244	2:89196748-89196859 2:89197005-89197300
319	IGKV6-21	ENSG00000211611	2:89459235-89459850	ENST00000390256	2:89459741-89459850 2:89459235-89459530
320	IGKV6D-21	ENSG00000225523	2:90060377-90060995	ENST00000436451	2:90060377-90060489 2:90060700-90060995
321	IGKV6D-41	ENSG00000211626	2:90108504-90109080	ENST00000390271	2:90108504-90108578 2:90108785-90109080
322	IGLJ1	ENSG00000211674	22:23235872-23235998	ENST00000390320	22:23235872-23235998
323	IGLJ2	ENSG00000211676	22:23241661-23241835	ENST00000390322	22:23241661-23241835
324	IGLJ3	ENSG00000211678	22:23247030-23247205	ENST00000390324	22:23247030-23247205
325	IGLJ5	ENSG00000211681	22:23256408-23256479	ENST00000390327	22:23256408-23256479
326	IGLJ6	ENSG00000211682	22:23260304-23260373	ENST00000390328	22:23260304-23260373
327	IGLJ7	ENSG00000211684	22:23263562-23263607	ENST00000390330	22:23263562-23263607
328	IGLV10-54	ENSG00000211642	22:22569184-22569660	ENST00000390287	22:22569184-22569242 22:22569355-22569660
329	IGLV11-55	ENSG00000211641	22:22556057-22556600	ENST00000390286	22:22556057-22556114 22:22556233-22556600
330	IGLV1-36	ENSG00000211655	22:22786296-22786802	ENST00000390301	22:22786296-22786381 22:22786497-22786802
331	IGLV1-40	ENSG00000211653	22:22764098-22764614	ENST00000390299	22:22764098-22764194 22:22764305-22764614
332	IGLV1-44	ENSG00000211651	22:22735135-22735715	ENST00000390297	22:22735135-22735294 22:22735410-22735715
333	IGLV1-47	ENSG00000211648	22:22712087-22712608	ENST00000390294	22:22712087-22712188 22:22712304-22712608
334	IGLV1-50	ENSG00000211645	22:22681658-22682172	ENST00000390291	22:22681658-22681754 22:22681865-22682172

335	IGLV1-51	ENSG00000211644	22:22676828-22677336	ENST00000390290	22:22676828-22676909 22:22677019-22677336
336	IGLV2-11	ENSG00000211668	22:23134980-23135496	ENST00000390314	22:23134980-23135067 22:23135185-23135496
337	IGLV2-14	ENSG00000211666	22:23101189-23101707	ENST00000390312	22:23101189-23101275 22:23101393-23101707
338	IGLV2-18	ENSG00000211664	22:23077095-23077584	ENST00000390310	22:23077095-23077153 22:23077270-23077584
339	IGLV2-23	ENSG00000211660	22:23040274-23040892	ENST00000390306	22:23040274-23040481 22:23040599-23040892
340	IGLV2-33	ENSG00000211656	22:22930626-22931145	ENST00000390302	22:22930626-22930729 22:22930852-22931145
341	IGLV2-8	ENSG00000211671	22:23165153-23165787	ENST00000390317	22:23165153-23165360 22:23165476-23165787
342	IGLV3-1	ENSG00000211673	22:23222886-23223576	ENST00000390319	22:23222886-23222983 22:23223277-23223576
343	IGLV3-10	ENSG00000211669	22:23154244-23154782	ENST00000390315	22:23154244-23154324 22:23154478-23154782
344	IGLV3-12	ENSG00000211667	22:23114317-23115079	ENST00000390313	22:23114317-23114371 22:23114775-23115079
345	IGLV3-16	ENSG00000211665	22:23089870-23090398	ENST00000390311	22:23089870-23089953 22:23090108-23090398
346	IGLV3-19	ENSG00000211663	22:23063108-23063630	ENST00000390309	22:23063108-23063193 22:23063340-23063630
347	IGLV3-21	ENSG00000211662	22:23054174-23055688	ENST00000390308	22:23054174-23054290 22:23054770-23054949 22:23055384-23055688
348	IGLV3-22	ENSG00000211661	22:23046750-23047307	ENST00000390307	22:23046750-23046859 22:23047009-23047307
349	IGLV3-25	ENSG00000211659	22:23029190-23029735	ENST00000390305	22:23029190-23029278 22:23029445-23029735
350	IGLV3-27	ENSG00000211658	22:23010758-23011276	ENST00000390304	22:23010758-23010841 22:23010984-23011276

351	IGLV3-32	ENSG00000211657	22:22936998-22937501	ENST00000390303	22:22936998-22937079 22:22937226-22937501
352	IGLV3-9	ENSG00000211670	22:23161507-23162253	ENST00000390316	22:23161507-23161667 22:23161954-23162253
353	IGLV4-3	ENSG00000211672	22:23213686-23214214	ENST00000390318	22:23213686-23213763 22:23213890-23214214
354	IGLV4-60	ENSG00000211639	22:22516592-22517074	ENST00000390284	22:22516592-22516643 22:22516765-22517074
355	IGLV4-69	ENSG00000211637	22:22385332-22385870	ENST00000390282	22:22385332-22385440 22:22385561-22385870
356	IGLV5-37	ENSG00000211654	22:22781876-22782371	ENST00000390300	22:22781876-22781926 22:22782049-22782371
357	IGLV5-45	ENSG00000211650	22:22730355-22730874	ENST00000390296	22:22730355-22730428 22:22730552-22730874
358	IGLV5-48	ENSG00000211647	22:22707289-22707781	ENST00000390293	22:22707289-22707334 22:22707459-22707781
359	IGLV5-52	ENSG00000211643	22:22673082-22673581	ENST00000390289	22:22673082-22673132 22:22673254-22673581
360	IGLV6-57	ENSG00000211640	22:22550113-22550686	ENST00000390285	22:22550113-22550244 22:22550370-22550686
361	IGLV7-43	ENSG00000211652	22:22749356-22749827	ENST00000390298	22:22749356-22749435 22:22749523-22749827
362	IGLV7-46	ENSG00000211649	22:22723982-22724454	ENST00000390295	22:22723982-22724061 22:22724151-22724454
363	IGLV8-61	ENSG00000211638	22:22453110-22453622	ENST00000390283	22:22453110-22453216 22:22453316-22453622
364	IGLV9-49	ENSG00000223350	22:22697539-22698084	ENST00000427632	22:22697539-22697621 22:22697759-22698084

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 idiotype and the relation between idiotype 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.

***

CLAIMS:

1. A non-human vertebrate (optionally a mouse or a rat) or vertebrate cell whose genome comprises an immunoglobulin heavy chain locus comprising human gene segment JH6*02, one or more VH gene segments and one or more D gene segments upstream of a constant region; wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof so that the mouse is capable of producing an antibody heavy chain produced by recombination of the human JH6*02 with a D segment and a VH segment.

2. The vertebrate of claim 1, wherein the vertebrate has been immunised with a target antigen and wherein the variable domain of the heavy chain is the product of recombination between a VH, D and JH6*02 and wherein the HCDR3 length is at least 20 amino acids.

3. A non-human vertebrate cell (optionally a mouse cell or a rat cell) whose genome comprises an immunoglobulin heavy chain locus comprising human gene segment JH6*02, one or more VH gene segments and one or more D gene segments upstream of a constant region; wherein the gene segments in the heavy chain locus are operably linked to the constant region thereof for producing (eg, in a subsequent progeny cell) an antibody heavy chain produced by recombination of the human JH6*02 with a D segment and a VH segment.

4. The cell of claim 3, which is an ES cell capable of differentiation into a progeny antibody-producing cell that expresses said heavy chain.

5. The vertebrate or cell of any preceding claim, wherein the heavy chain locus comprises a human JH6*02 recombination signal sequence (RSS) operably connected 5' to the JH6*02 gene segment.

6. The vertebrate or cell of claim 5, wherein the RSS is SEQ ID NO: 238 or a sequence having an identical 9mer and 7mer sequence flanking a sequence that is at least 70% identical to the 22mer sequence of SEQ ID NO: 238.

7. The vertebrate or cell of claim 6, wherein the RSS and JH6*02 are provided as SEQ ID NO: 237.

8. The vertebrate or cell of any preceding claim, wherein the JH6*02 is the only JH6-type gene segment in the genome.

9. The vertebrate or cell of any preceding claim, wherein the JH6*02 is the closest JH gene segment to the constant region in the locus.

10. The vertebrate or cell of any preceding claim, wherein the locus comprises one, more or all human D gene segments D3-9; D4-17; D3-10; D2-2; D5-24; D6-19; D3-22; D6-13; D5-12; D1-26; D1-20; D5-18; D3-16; D2-21; D1-14; D7-27; D1-1; D6-25; D2-14; and D4-23.

11. The vertebrate or cell of claim 10, wherein the locus comprises one, more or all human D gene segments D3-9, D3-10, D6-19, D4-17, D6-13, D3-22, D2-2, D2-25 and D3-3.

12. The vertebrate or cell of any preceding claim, wherein the locus comprises a plurality of human D gene segments and the JH6*02 is in human germline configuration with respect to the 3'-most human D gene segment.

13. The vertebrate or cell of any preceding claim, wherein the locus comprises one, more or all of IGHV gene segments selected from V3-21, V3-13, V3-7, V6-1, V1-8, V1-2, V7-4-1, V1-3, V1-18, V4-4, V3-9, V3-23, V3-11 and V3-20.

14. The vertebrate or cell of any preceding claim, wherein the locus comprises one, more or all of human D3-9*01, D3-10*01, D6-19*01, D6-13*01, D1-26*01, IGHV1-8*01, IGHV4-61*01, IGHV6-1*01, IGHV4-4*02, IGHV1-3*01, IGHV3-66*03, IGHV3-7*01 and IGHV3-9*01.

15. An antibody-producing cell (eg, a B-cell) that is a progeny of the cell of any one of claims 3 to 14, wherein the antibody-producing cell comprises a heavy chain locus comprising a rearranged variable region produced by recombination of human JH6*02 with a D segment and a VH segment.

16. The cell of claim 15, which is a B-cell or hybridoma that expresses a target antigen-specific antibody comprising a heavy chain that comprises a rearranged variable region produced by recombination of human JH6*02 with a D segment and a VH segment.

17. The vertebrate or cell of any preceding claim, wherein the antibody heavy chain specifically binds a target antigen.

18. The vertebrate or cell of any preceding claim, wherein the antibody heavy chain has a HCDR3 length of at least 20 amino acids.

19. The vertebrate or cell of any preceding claim, wherein the antibody heavy chain is a product of the recombination of JH6*02 with a human VH gene segment recited in claim 13 or 14 and/or a D gene segment recited in claim 10, 11 or 14.

20. The vertebrate or cell of any preceding claim, wherein all endogenous non-human vertebrate heavy chain variable region gene segments have been inactivated in the genome.

21. The vertebrate or cell of any preceding claim, wherein the genome is homozygous for said heavy chain locus.

22. A heavy chain (eg, comprised by an antibody) isolated from a vertebrate of any one of claims 1, 2, 5 to 14 and 17 to 21 wherein the heavy chain comprises a HCDR3 of at least 20 amino acids.

23. The heavy chain of claim 22, wherein the HCDR3 is the product of recombination of human JH6*02 with a human VH gene segment recited in claim 13 or 14 and/or a D gene segment recited in claim 10, 11 or 14.

24. A heavy chain (eg, comprised by an antibody) whose VH variable domain is identical to the VH variable domain of the heavy chain of claim 22 or 23, and which comprises a human constant region or a human-mouse chimaeric constant region.

25. The heavy chain of claim 22, 23 or 24, whose VH variable domain is specific for a target antigen.

26. A method for producing a heavy chain, VH domain or an antibody specific to a target antigen, the method comprising immunizing a non-human vertebrate according to any one

of claims 1, 2, 5 to 14 and 17 to 21 with the antigen and isolating the heavy chain, VH domain or an antibody specific to a target antigen or a cell producing the heavy chain, VH domain or an antibody, wherein the heavy chain, VH domain or an antibody comprises a HCDR3 that is derived from the recombination of human JH6*02 with a VH gene segment and a D gene segment.

27. A method for producing a human heavy chain or antibody comprising carrying out the method of claim 26, wherein the constant region of the locus is a non-human vertebrate (eg, mouse or rat) constant region, and then replacing the non-human constant region of the isolated heavy chain or antibody with a human constant region.

28. A heavy chain, VH domain or an antibody produced by the method of claim 26 or 27.

29. A B-cell or hybridoma expressing a heavy chain VH domain that is identical to the VH domain of the heavy chain of claim 22, 23 or 28.

30. A nucleic acid encoding the VH domain of the heavy chain of claim 22, 23 or 28, or encoding the heavy chain of claim 22, 23, 24, 25 or 28.

31. A vector (eg, a CHO cell or HEK293 cell vector) comprising the nucleic acid of claim 30; optionally wherein the vector is in a host cell (eg, a CHO cell or HEK293 cell).

32. A pharmaceutical composition comprising the antibody, heavy chain or VH domain (eg, comprised by an antibody) of any one of claims 22 to 25 and 28, together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, heavy chain or antibody).

33. The antibody, heavy chain or VH domain (eg, comprised by an antibody) of any one of claims 22 to 25 and 28 for use in medicine.

34. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 3 human variable region gene segments of the same type (eg, at least 3 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 3 human Vκ1-39 gene segments, at least 3 human D2-2 gene segments or

at least 3 human Jκ1 gene segments), wherein at least two of the human gene segments are variants that are not identical to each other.

35. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *cis* at the same Ig locus.

36. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *trans* at the same Ig locus; and optionally a third human gene segment of the same type, wherein the third gene segment is *cis* with one of said 2 different gene segments.

37. A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human variable region gene segments, wherein the plurality comprises at least 2 human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), a first of said different gene segments is provided in the genome of a first vertebrate of the population, and a second of said different gene segments being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second gene segment.

38. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

39. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 3 human variable region gene segments of the same type (eg, at

least 3 human VH6-1 gene segments, at least 3 human JH6 gene segments, at least 3 human Vκ1-39 gene segments, at least 3 human D2-2 gene segments or at least 3 human Jκ1 gene segments), wherein at least two of the human gene segments are variants that are not identical to each other.

40. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *cis* at the same Ig locus.

41. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments) *trans* at the same Ig locus; and optionally a third human gene segment of the same type, wherein the third gene segment is *cis* with one of said 2 different gene segments.

42. A method of providing an enhanced human immunoglobulin variable region gene segment repertoire, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human variable region gene segments, wherein the method comprises providing at least 2 different human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein a first of said different gene segments is provided in the genome of a first vertebrate of the population, and a second of said different gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second gene segment.

43. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

44. The vertebrate, cell or method of any one of claims 34 to 43, wherein at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

45. The vertebrate, cell or method of any one of claims 34 to 44, wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

46. The vertebrate, cell or method of any one of claims 34 to 45, wherein the gene segments are derived from the genome sequence of two or more different human individuals; optionally wherein the different human individuals are from different human populations.

47. The vertebrate, cell or method of claim 46, wherein the individuals are not genetically related.

48. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human variable region gene segments of the same type (eg, at least 2 human VH6-1 gene segments, at least 2 human JH6 gene segments, at least 2 human Vκ1-39 gene segments, at least 2 human D2-2 gene segments or at least 2 human Jκ1 gene segments), wherein the gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same Ig locus in said genome.

49. The method of claim 48, wherein the different human individuals are from different human populations.

50. The method of claim 48, wherein the individuals are not genetically related.

51. The vertebrate, cell or method of preceding claim, wherein at least one of the different segments is a synthetic mutant of a human germline gene segment.

52. The vertebrate, cell or method of any one of claims 34 to 51, wherein each of said gene segments occurs in 10 or more different human populations.

53. The vertebrate, cell or method of preceding claim, wherein each of said gene segments has a human frequency of 5% or greater.

54. The vertebrate, cell or method of claim 53, wherein each of said gene segments occurs in 10 or more different human populations.

55. The vertebrate, cell or method of any one of claims 34 to 54, wherein each of said gene segments occurs in the 1000 Genomes database in more than 50 individuals.

56. The vertebrate, cell or method of preceding claim, wherein each of said gene segments (i) has a human frequency of 5% or greater; and (ii) occurs in 10 or more different human populations.

57. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) having a genome comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence (eg, IGHJ6 ref; SEQ ID NO: 244).

58. The vertebrate or cell of claim 57, wherein the genome comprises a third human gene segment of said type, the third gene segment being different from the first and second gene segments.

59. The vertebrate or cell of claim 57 or 58, wherein the first and second gene segments are *cis* on the same chromosome; and optionally the third gene segment is also *cis* on said chromosome.

60. The vertebrate or cell of claim 59, wherein the gene segments are targeted insertions into an endogenous non-human Ig locus.

61. The vertebrate or cell of claim 57 or 58, wherein the first and second gene segments are *trans* on different chromosomes.

62. The vertebrate or cell of any one of claims 57 to 61, wherein the first gene segment is a gene segment selected any one of Tables 1 to 7 and 9 to 14 and the second gene segment is the corresponding reference sequence.

63. A population of non-human vertebrates (eg, mice or rats) comprising first and second human Ig locus gene segments of the same type (eg, first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence (eg, SEQ ID NO: 244), wherein the first gene segment is provided in the genome of a first vertebrate of the population, and the second gene segment is provided in the genome of a second vertebrate of the population.

64. The population of claim 63, wherein the genome of the first vertebrate does not comprise the second gene segment.

65. The population of claim 63 or 64, wherein the population comprises a third human gene segment of said type, the third gene segment being different from the first and second gene segments and optionally wherein the first and third gene segments are present in the genome of the first vertebrate.

66. The population of claim 63, 64 or 65, wherein the gene segments are targeted insertions into an endogenous non-human Ig locus in the respective genome.

67. The population of any one of claims 63 to 66, wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 and the second gene segment is the corresponding reference sequence.

68. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising first and second human Ig locus gene segments of the same type (eg,

first and second human JH6 gene segments; or first and second IgG2 gene segments; or first and second human Jλ7 gene segments), wherein the first gene segment is a gene segment selected from any one of Tables 1 to 7 and 9 to 14 (eg, IGHJ6-a) and the second gene segment is the corresponding reference sequence (eg, SEQ ID NO: 244).

69. A method of providing an enhanced human immunoglobulin gene segment repertoire, the method comprising providing a population according to any one of claims 63 to 66.

70. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 34, having a genome comprising at least 3 human D gene segments of the same type (eg, D2-2 gene segments), wherein at least two of the human D gene segments are variants that are not identical to each other (eg, D2-2ref and D2-2a).

71. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 35, having a genome comprising at least 2 different non-endogenous D gene segments of the same type type (eg, D2-2ref and D2-2a) *cis* at the same Ig locus.

72. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 36, having a genome comprising at least 2 different human D gene segments of the same type (eg, D2-2ref and D2-2a) *trans* at the same Ig locus; and optionally a third human D gene segment (eg, (eg, D2-2ref, D2-2a or D2-2b) of the same type, wherein the third D is *cis* with one of said 2 different D gene segments.

73. A population of non-human vertebrates (eg, mice or rats) according to claim 37, comprising a repertoire of human D gene segments, wherein the plurality comprises at least 2 different human D gene segments of the same type (eg, D2-2 gene segments), a first of said different D gene segments (eg, D2-2ref) is provided in the genome of a first vertebrate of the population, and a second of said different D gene segment (eg, D2-2a) being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second D gene segment.

74. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 38, having a genome comprising at least 2 different non-endogenous D gene segments of the same type (eg, human D2-2 gene segments), wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

75. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 39, the method comprising providing the vertebrate with

a genome comprising at least 3 human D gene segments of the same type (eg, D2-2 gene segments), wherein at least two of the human D gene segments are variants that are not identical to each other (eg, D2-2ref and D2-2a).

76. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 40, the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous D gene segments of the same type (eg, human D2-2 gene segments) *cis* at the same Ig locus.

77. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 41, the method comprising providing the vertebrate with a genome comprising at least 2 different human D gene segments of the same type (eg, D2-2ref and D2-2a) *trans* at the same Ig locus; and optionally a third human D gene segment (eg, D2-2ref, D2-2a or D2-2b) of the same type, wherein the third D is *cis* with one of said 2 different D gene segments.

78. A method of providing an enhanced human immunoglobulin D gene segment repertoire according to claim 42, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human D gene segments, wherein the method comprises providing at least 2 different human D gene segments of the same type (eg, D2-2ref and D2-2a), wherein a first of said different D gene segments is provided in the genome of a first vertebrate of the population, and a second of said different D gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second D gene segment.

79. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 43, the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous D gene segments of the same type (eg, D2-2ref and D2-2a), wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

80. The vertebrate or cell of claim 70, 71 or 74, or the method of claim 75, 77 or 79, wherein at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

81. The vertebrate or cell of claim 70, 71 or 71, or the method of any one of claims 75 to 78 and 80, wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

82. The vertebrate or cell of any one of claims 70 to 73, or the method of any one of claims 75 to 78 and 81, wherein the D gene segments are derived from the genome sequence of two or more

different human individuals; optionally wherein the different human individuals are from different human populations.

83. The vertebrate, cell or method of claim 82, wherein the individuals are not genetically related.

84. The vertebrate, cell or method of any one of claims 70 to 83, wherein at least one of the different D segments is a synthetic mutant of a human germline D gene segment.

85. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 48, the method comprising providing the vertebrate with a genome comprising at least 2 human D gene segments (eg, D2-2ref and D2-2a), wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same IgH locus in said genome.

86. The vertebrate or cell of any one of claims 70 to 74 and 80 to 84, wherein the genome comprises a substantially complete functional repertoire of human D gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

87. The population of claim 73, wherein the population comprises a substantially complete functional repertoire of human D gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

88. A non-human vertebrate (eg, a mouse or rat) or a non-human cell (eg, an ES cell or a B-cell) having a genome comprising a substantially complete functional repertoire of human D gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

89. A population of non-human vertebrates (eg, mice or rats) comprising a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more variant human D gene segments, wherein said substantially complete functional repertoire and the supplementary D gene segments are not found together in the germline genome of a human individual.

90. The vertebrate or cell of claim 88 or the population of claim 89, comprising first and second D gene segments selected from
D2-2ref and D2-2a; or
D2-21ref and D2-21a; or
D3-10ref and D3-10a; or
D3-16ref and D3-16a; or
D2-8ref and D2-8a; or
D3-3ref and D3-3a; or
D4-23ref and D4-23a; or
D6-13ref and D6-13a; or
D3-9ref and D3-9a; or
D4-4ref and D4-4a; or
D7-27ref and D7-27a;

optionally wherein the first and/or second D gene segment is present in two or more copies.

91. The vertebrate, cell or population of claim 90, comprising human gene segments D2-2ref and D2-2a; and D3-3ref and D3-3a; and optionally also D2-15.

92. A non-human vertebrate or vertebrate cell according to claim 71, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the D gene repertoire comprising one or more of

a plurality of D2-2 gene segments provided by at least 2 different D2-2 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D2-21 gene segments provided by at least 2 different D2-21 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D3-10 gene segments provided by at least 2 different D3-10 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D3-16 gene segments provided by at least 2 different D3-16 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D2-8 gene segments provided by at least 2 different D2-8 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D3-3 gene segments provided by at least 2 different D3-3 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D4-23 gene segments provided by at least 2 different D4-23 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D6-13 gene segments provided by at least 2 different D6-13 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D3-9 gene segments provided by at least 2 different D3-9 gene segments in *cis* at the same Ig locus in said genome;
a plurality of D4-4 gene segments provided by at least 2 different D4-4 gene segments in *cis* at

the same Ig locus in said genome; and
a plurality of D7-27 gene segments provided by at least 2 different D7-27 gene segments in *cis* at the same Ig locus in said genome;

optionally wherein the D gene segments are derived from the genome sequence of two or more different human individuals.

93. A non-human vertebrate or vertebrate cell according to claim 71, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the D gene repertoire comprising one or more of

a plurality of D2-2 gene segments provided by at least 2 different D2-2 gene segments in *trans* in said genome;
a plurality of D2-21 gene segments provided by at least 2 different D2-21 gene segments in *trans* in said genome;
a plurality of D3-10 gene segments provided by at least 2 different D3-10 gene segments in *trans* in said genome;
a plurality of D3-16 gene segments provided by at least 2 different D3-16 gene segments in *trans* in said genome;
a plurality of D2-8 gene segments provided by at least 2 different D2-8 gene segments in *trans* in said genome;
a plurality of D3-3 gene segments provided by at least 2 different D3-3 gene segments in *trans* in said genome;
a plurality of D4-23 gene segments provided by at least 2 different D4-23 gene segments in *trans* in said genome;
a plurality of D6-13 gene segments provided by at least 2 different D6-13 gene segments in *trans* in said genome;
a plurality of D3-9 gene segments provided by at least 2 different D3-9 gene segments in *trans* in said genome;
a plurality of D4-4 gene segments provided by at least 2 different D4-4 gene segments in *trans* in said genome; and
a plurality of D7-27 gene segments provided by at least 2 different D7-27 gene segments in *trans* in said genome;

optionally wherein the D gene segments are derived from the genome sequence of two or more different human individuals.

94. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell), according to claim 70, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the D gene repertoire comprising one or more of

a plurality of D2-2 gene segments provided by at least 3 different D2-2 gene segments;

a plurality of D2-21 gene segments provided by at least 3 different D2-21 gene segments;
a plurality of D3-10 gene segments provided by at least 3 different D3-10 gene segments;
a plurality of D3-16 gene segments provided by at least 3 different D3-16 gene segments;
a plurality of D2-8 gene segments provided by at least 3 different D2-8 gene segments;
a plurality of D3-3 gene segments provided by at least 3 different D3-3 gene segments;
a plurality of D4-23 gene segments provided by at least 3 different D4-23 gene segments;
a plurality of D6-13 gene segments provided by at least 3 different D6-13 gene segments;
a plurality of D3-9 gene segments provided by at least 3 different D3-9 gene segments;
a plurality of D4-4 gene segments provided by at least 3 different D4-4 gene segments; and
a plurality of D7-27 gene segments provided by at least 3 different D7-27 gene segments;

optionally wherein the D gene segments are derived from the genome sequence of two or three different human individuals;

optionally wherein at least 2 or 3 of said different gene segments are provided in *cis* at the same Ig locus in said genome.

95. The vertebrate or cell of claim 92, 93 or 94, wherein the different human individuals are from different human populations.

96. The vertebrate or cell of any one of claims 92 to 95, wherein the individuals are not genetically related.

97. The vertebrate or cell of any one of claims 92 to 96, wherein at least one of the different D segments is a synthetic mutant of a human germline D gene segment.

98. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell) according to claim 74, comprising a genome comprising human VH, D and JH gene repertoires, the D gene repertoire comprising of one or more of

a plurality of D2-2 gene segments provided by at least 2 different D2-2 gene ; optionally in *cis* in said genome;
a plurality of D2-21 gene segments provided by at least 2 different D2-21 gene ; optionally in *cis* in said genome;
a plurality of D3-10 gene segments provided by at least 2 different D3-10 gene ; optionally in *cis* in said genome;
a plurality of D3-16 gene segments provided by at least 2 different D3-16 gene ; optionally in *cis* in said genome;
a plurality of D2-8 gene segments provided by at least 2 different D2-8 gene ; optionally in *cis* in said genome;
a plurality of D3-3 gene segments provided by at least 2 different D3-3 gene ; optionally in *cis* in said genome;

a plurality of D4-23 gene segments provided by at least 2 different D4-23 gene ; optionally in *cis* in said genome;
a plurality of D6-13 gene segments provided by at least 2 different D6-13 gene ; optionally in *cis* in said genome;
a plurality of D3-9 gene segments provided by at least 2 different D3-9 gene ; optionally in *cis* in said genome;
a plurality of D4-4 gene segments provided by at least 2 different D4-4 gene ; optionally in *cis* in said genome; and
a plurality of D7-27 gene segments provided by at least 2 different D7-27 gene ; optionally in *cis* in said genome;

wherein the D gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

99. The vertebrate or cell of claim 98, wherein the human individuals are from different human populations.

100. The vertebrate, cell or population of any one of claims 70 to 99, wherein one or more of the D gene segments is a variant of a human germline D gene segment, wherein the variant gene segment encodes an amino acid sequence that differs by 1, 2 or 3 amino acids from the corresponding amino acid sequence encoded by the human germline D gene segment, provided in that said amino acid sequence encoded by the variant does not include a stop codon when said corresponding amino acid sequence does not include a stop codon.

101. The vertebrate, cell or population of claim 100, wherein said corresponding amino acid sequence encoded by the human germline D gene segment is a hydrophilic or hydrophobic sequence (according to J Mol Biol. 1997 Jul 25;270(4):587-97; Corbett SJ *et al*; Table 2).

102. The vertebrate, cell or population of claim 100 or 101, comprising said variant and said germline human D gene segments; optionally wherein the variant and germline human D gene segments are *cis* on the same chromosome.

103. The vertebrate, cell or population of any one of claims 100 to 102, wherein germline human D gene segment is a D2, D3, D5 or D6 family gene segment; optionally a D2-2, D2-15, D3-3, D3-9, D3-10, D3-22, D5-5, D5-18, D6-6, D6-13, D6-19 gene segment.

104. The vertebrate, cell or population of any one of claims 70 to 103, comprising a plurality of D2-2 gene segments, wherein the plurality comprises D2-2 gene segments that vary from each other at one or more nucleotide positions corresponding to positions

106,382,687 and

106,382,711
on human chromosome 14.

105. The vertebrate, cell or population of claim 104, wherein the plurality comprises a human D2-2 gene segment comprising a thymine at a position corresponding to position 106,382,687 on human chromosome 14; and optionally no further mutation from the sequence of D2-2ref.

106. The vertebrate, cell or population of claim 104 or 105, wherein the plurality comprises a human D2-2 gene segment comprising a cytosine at a position corresponding to position 106,382,687 on human chromosome 14; and optionally no further mutation from the sequence of D2-2a.

107. The vertebrate, cell or population of any one of claims 104 to 106, wherein the plurality comprises a human D2-2 gene segment comprising an adenine at a position corresponding to position 106,382,711 on human chromosome 14; and optionally no further mutation from the sequence of D2-2b.

108. The vertebrate, cell or population of any one of claims 104 to 107, wherein the plurality comprises a human D2-2 gene segment comprising an thymine at a position corresponding to position 106,382,711 on human chromosome 14; and optionally no further mutation from the sequence of D2-2ref.

109. The vertebrate, cell or population of any one of claims 70 to 108, comprising a plurality of D7-27 gene segments, wherein the plurality comprises D7-27 gene segments that vary from each other at a nucleotide position corresponding to position 106,331,767 on human chromosome 14.

110. The vertebrate, cell or population of claim 109, wherein the plurality comprises a human D7-27 gene segment comprising a cytosine at a position corresponding to position 106,331,767 on human chromosome 14; and optionally no further mutation from the sequence of D7-27ref.

111. The vertebrate, cell or population of claim 109 or 110, wherein the plurality comprises a human D7-27 gene segment comprising a guanine at a position corresponding to position 106,331,767 on human chromosome 14; and optionally no further mutation from the sequence of D7-27a.

112. The vertebrate, cell or population of any one of claims 70 to 111, comprising a plurality of D4-23 gene segments, wherein the plurality comprises D4-23 gene segments that vary from each other at a nucleotide position corresponding to position 106,350,740 on human chromosome 14.

113. The vertebrate, cell or population of claim 112, wherein the plurality comprises a human D4-23 gene segment comprising an adenine at a position corresponding to position 106,350,740 on human chromosome 14; and optionally no further mutation from the sequence of D4-23ref.

114. The vertebrate, cell or population of claim 112 or 113, wherein the plurality comprises a human D4-23 gene segment comprising an guanine at a position corresponding to position 106,350,740 on human chromosome 14; and optionally no further mutation from the sequence of D4-23a.

115. The vertebrate, cell or population of any one of claims 70 to 113, comprising a plurality of D2-21 gene segments, wherein the plurality comprises D2-21 gene segments that vary from each other at a nucleotide position corresponding to position 106,354,418 on human chromosome 14.

116. The vertebrate, cell or population of claim 115, wherein the plurality comprises a human D2-21 gene segment comprising an adenine at a position corresponding to position 106,354,418 on human chromosome 14; and optionally no further mutation from the sequence of D2-21ref.

117. The vertebrate, cell or population of claim 115 or 116, wherein the plurality comprises a human D2-21 gene segment comprising a guanine at a position corresponding to position 106,354,418 on human chromosome 14; and optionally no further mutation from the sequence of D2-21a.

118. The vertebrate, cell or population of any one of claims 70 to 117, comprising a plurality of D3-16 gene segments, wherein the plurality comprises D3-16 gene segments that vary from each other at a nucleotide position corresponding to position 106,354,418 on human chromosome 14.

119. The vertebrate, cell or population of claim 118, wherein the plurality comprises a human D3-16 gene segment comprising a thymine at a position corresponding to position 106,361,515 on human chromosome 14; and optionally no further mutation from the sequence of D3-16ref.

120. The vertebrate, cell or population of claim 118 or 119, wherein the plurality comprises a human D3-16 gene segment comprising a cytosine at a position corresponding to position 106,361,515 on human chromosome 14; and optionally no further mutation from the sequence of D3-16a.

121. The vertebrate, cell or population of any one of claims 70 to 120, comprising a plurality of D6-13 gene segments, wherein the plurality comprises D6-13 gene segments that vary from each other at a nucleotide position corresponding to position 106,367,013 on human chromosome 14.

122. The vertebrate, cell or population of claim 121, wherein the plurality comprises a human D6-13 gene segment comprising a thymine at a position corresponding to position 106,367,013 on human chromosome 14; and optionally no further mutation from the sequence of D6-13ref.

123. The vertebrate, cell or population of claim 121 or 122, wherein the plurality comprises a human D6-13 gene segment comprising a cytosine at a position corresponding to position 106,367,013 on human chromosome 14; and optionally no further mutation from the sequence of D6-13a.

124. The vertebrate, cell or population of any one of claims 70 to 123, comprising a plurality of D3-10 gene segments, wherein the plurality comprises D3-10 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,370,370 and
106,370,371
on human chromosome 14.

125. The vertebrate, cell or population of claim 124, wherein the plurality comprises a human D3-10 gene segment comprising a thymine at a position corresponding to position 106,370,370 on human chromosome 14; and optionally no further mutation from the sequence of D3-10ref.

126. The vertebrate, cell or population of claim 124 or 125, wherein the plurality comprises a human D3-10 gene segment comprising a cytosine at a position corresponding to position 106,370,370 on human chromosome 14; and optionally no further mutation from the sequence of D3-10a.

127. The vertebrate, cell or population of claim 124, 125 or 126 wherein the plurality comprises a human D3-10 gene segment comprising an adenine at a position corresponding to position 106,370,371 on human chromosome 14; and optionally no further mutation from the sequence of D3-10ref.

128. The vertebrate, cell or population of any one of claims 124 to 127, wherein the plurality comprises a human D3-10 gene segment comprising a guanine at a position corresponding to position 106,370,371 on human chromosome 14; and optionally no further mutation from the

sequence of D3-10b.

129. The vertebrate, cell or population of any one of claims 70 to 128, comprising a plurality of D3-9 gene segments, wherein the plurality comprises D3-9 gene segments that vary from each other at a nucleotide position corresponding to position 106,370,567 on human chromosome 14.

130. The vertebrate, cell or population of claim 129, wherein the plurality comprises a human D3-9 gene segment comprising an adenine at a position corresponding to position 106,370,567 on human chromosome 14; and optionally no further mutation from the sequence of D3-9ref.

131. The vertebrate, cell or population of claim 129 or 130, wherein the plurality comprises a human D3-9 gene segment comprising a thymine at a position corresponding to position 106,370,567 on human chromosome 14; and optionally no further mutation from the sequence of D3-9a.

132. The vertebrate, cell or population of any one of claims 70 to 131, comprising a plurality of D2-8 gene segments, wherein the plurality comprises D2-8 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,373,085;
106,373,086 and
106,373,089
on human chromosome 14.

133. The vertebrate, cell or population of claim 132, wherein the plurality comprises a human D2-8 gene segment comprising a cytosine at a position corresponding to position 106,373,085 on human chromosome 14.

134. The vertebrate, cell or population of claim 132 or 133, wherein the plurality comprises a human D2-8 gene segment comprising a thymine at a position corresponding to position 106,373,085 on human chromosome 14; and optionally no further mutation from the sequence of D2-8b.

135. The vertebrate, cell or population of claim 132, 133 or 134 wherein the plurality comprises a human D2-8 gene segment comprising a cytosine at a position corresponding to position 106,373,086 on human chromosome 14; and optionally no further mutation from the sequence of D2-8ref.

136. The vertebrate, cell or population of any one of claims 132 to 135, wherein the plurality comprises a human D2-8 gene segment comprising a thymine at a position corresponding to position 106,373,086 on human chromosome 14; and optionally no further mutation from the sequence of D2-8ref.

137. The vertebrate, cell or population of any one of claims 70 to 136, comprising a plurality of D4-4 gene segments, wherein the plurality comprises D4-4 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,379,086; and
106,379,089
on human chromosome 14.

138. The vertebrate, cell or population of claim 137, wherein the plurality comprises a D4-4 gene segment comprising a cytosine at a position corresponding to position 106,379,086 on human chromosome 14; and optionally no further mutation from the sequence of D4-4ref.

139. The vertebrate, cell or population of claim 137 or 138, wherein the plurality comprises a human D4-4 gene segment comprising a thymine at a position corresponding to position 106,379,086 on human chromosome 14; and optionally no further mutation from the sequence of D4-4a.

140. The vertebrate, cell or population of claim 137, 138 or 139 wherein the plurality comprises a human D4-4 gene segment comprising a cytosine at a position corresponding to position 106,379,089 on human chromosome 14; and optionally no further mutation from the sequence of D4-4ref or a cytosine at a position corresponding to position 106,379,086 on human chromosome 14.

141. The vertebrate, cell or population of any one of claims 137 to 140, wherein the plurality comprises a human D4-4 gene segment comprising a thymine at a position corresponding to position 106,373,089 on human chromosome 14; and optionally no further mutation from the sequence of D4-4a.

142. The vertebrate, cell or population of any one of claims 70 to 141, comprising a plurality of D3-3 gene segments, wherein the plurality comprises D3-3 gene segments that vary from each other at one or more nucleotide positions corresponding to positions
106,380,241; and
106,380,246
on human chromosome 14.

143. The vertebrate, cell or population of claim 142, wherein the plurality comprises a D3-3 gene segment comprising a thymine at a position corresponding to position 106,380,241 on human chromosome 14; and optionally no further mutation from the sequence of D3-3ref.

144. The vertebrate, cell or population of claim 142 or 143, wherein the plurality comprises a human D3-3 gene segment comprising a cytosine at a position corresponding to position 106,380,241 on human chromosome 14; and optionally no further mutation from the sequence of D3-3a.

145. The vertebrate, cell or population of claim 142, 143 or 144 wherein the plurality comprises a human D3-3 gene segment comprising an adenine at a position corresponding to position 106,380,246 on human chromosome 14; and optionally no further mutation from the sequence of D3-3ref.

146. The vertebrate, cell or population of any one of claims 142 to 145, wherein the plurality comprises a human D3-3 gene segment comprising a thymine at a position corresponding to position 106,380,246 on human chromosome 14; and optionally no further mutation from the sequence of D3-3a.

147. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 34, having a genome comprising at least 3 human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein at least two of the human JH gene segments are variants that are not identical to each other.

148. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 35, having a genome comprising at least 2 different non-endogenous JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *cis* at the same Ig locus.

149. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 36, having a genome comprising at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *trans* at the same Ig locus; and optionally a third human JH gene segments of the same type, wherein the third JH is *cis* with one of said 2 different JH gene segments.

150. A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human JH gene segments according to claim 37, wherein the plurality comprises at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), a first of said different JH gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JH gene segments being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JH gene segment.

151. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 38, having a genome comprising at least 2 different non-

endogenous JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

152. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 39, the method comprising providing the vertebrate with a genome comprising at least 3 human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein at least two of the human JH gene segments are variants that are not identical to each other.

153. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 40, the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *cis* at the same Ig locus.

154. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 41, the method comprising providing the vertebrate with a genome comprising at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6) *trans* at the same Ig locus; and optionally a third human JH gene segments of the same type, wherein the third JH is *cis* with one of said 2 different JH gene segments.

155. A method of providing an enhanced human immunoglobulin JH gene segment repertoire according to claim 42, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human JH gene segments, wherein the method comprises providing at least 2 different human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein a first of said different JH gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JH gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JH gene segment.

156. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 43, the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

157. The vertebrate or cell of claim 147, 149 or 151, or the method of claim 152, 154, 155 or 156, wherein at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

158. The vertebrate or cell of claim 147, 148 or 149, or the method of claim 152, 153, 154 or 155, wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

159. The vertebrate or cell of claim 147, 148 or 149, or the method of claim 152, 153, 154 or 155, wherein the JH gene segments are derived from the genome sequence of two or more different human individuals; optionally wherein the different human individuals are from different human populations.

160. The vertebrate, cell or method of claim 159, wherein the individuals are not genetically related.

161. The vertebrate, cell or method of any one of claims 147 to 160, wherein at least one of the different JH segments is a synthetic mutant of a human germline JH gene segment.

162. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human JH gene segments of the same type (JH1, JH2, JH3, JH4, JH5 or JH6), wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same IgH locus in said genome.

163. The vertebrate or cell of any one of claims 147 to 149 and 151, wherein the genome comprises a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

164. The population of claim 150, wherein the population comprises a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

165. A non-human vertebrate (eg, a mouse or rat) or a non-human cell (eg, an ES cell or a B-cell) having a genome comprising a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

166. A population of non-human vertebrates (eg, mice or rats) comprising a substantially complete functional repertoire of human JH gene segment types supplemented with one, two or more human JH gene segments, wherein said substantially complete functional repertoire and the supplementary JH gene segments are not found together in the germline genome of a human individual.

167. The vertebrate of claim 165 or the population of claim 166, wherein at least one of said JH gene segments is SEQ ID NO: 1, 2, 3 or 4.

168. The vertebrate or population of claim 167, wherein at least one of said JH gene segments is SEQ ID NO: 1 and at least one, two or more of said supplementary JH gene segments is a variant according to any one of claims 178 to 196.

169. The vertebrate or population of claim 167 or 168, wherein at least one of said JH gene segments is SEQ ID NO: 2 and at least one, two or more of said supplementary JH gene segments is a variant according to any one of claims 198 to 209.

170. The vertebrate or population of claim 167, 168 or 169, wherein at least one of said JH gene segments is SEQ ID NO: 2 and at least one, two or more of said supplementary JH gene segments

is a variant according to any one of claims 212 to 217.

171. A non-human vertebrate or vertebrate cell according to claim 148, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the JH gene repertoire comprising

a plurality of JH1 gene segments provided by at least 2 different JH1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH2 gene segments provided by at least 2 different JH2 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH3 gene segments provided by at least 2 different JH3 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH4 gene segments provided by at least 2 different JH4 gene segments in *cis* at the same Ig locus in said genome;
a plurality of JH5 gene segments provided by at least 2 different JH5 gene segments in *cis* at the same Ig locus in said genome; or
a plurality of JH6 gene segments provided by at least 2 different JH6 gene segments in *cis* at the same Ig locus in said genome;

optionally wherein the JH gene segments are derived from the genome sequence of two or more different human individuals.

172. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell), according to claim 147, comprising a genome that comprises VH, D and JH gene repertoires comprising human gene segments, the JH gene repertoire comprising

a plurality of JH1 gene segments provided by at least 3 different JH1 gene segments;
a plurality of JH2 gene segments provided by at least 3 different JH2 gene segments;
a plurality of JH3 gene segments provided by at least 3 different JH3 gene segments;
a plurality of JH4 gene segments provided by at least 3 different JH4 gene segments;
a plurality of JH5 gene segments provided by at least 3 different JH5 gene segments; or
a plurality of JH6 gene segments provided by at least 3 different JH6 gene segments;

optionally wherein the JH gene segments are derived from the genome sequence of two or three different human individuals;

optionally wherein at least 2 or 3 of said different gene segments are provided in *cis* at the same Ig locus in said genome.

173. The vertebrate or cell of claim 171 or 172, wherein the different human individuals are from different human populations.

174. The vertebrate or cell of any one of claims 171 to 173, wherein the individuals are not genetically related.

175. The vertebrate or cell of any one of claims 171 to 174, wherein at least one of the different JH segments is a synthetic mutant of a human germline JH gene segment.

176. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell) according to claim 151, comprising a genome comprising human VH, D and JH gene repertoires, the JH gene repertoire comprising

a plurality of JH1 gene segments provided by at least 2 different human JH1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH2 gene segments provided by at least 2 different human JH2 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH3 gene segments provided by at least 2 different human JH3 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH4 gene segments provided by at least 2 different human JH4 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of JH5 gene segments provided by at least 2 different human JH5 gene segments, optionally in *cis* at the same Ig locus in said genome; or
a plurality of JH6 gene segments provided by at least 2 different human JH6 gene segments, optionally in *cis* at the same Ig locus in said genome;

wherein the JH gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

177. The vertebrate or cell of claim 176, wherein the human individuals are from different human populations.

178. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 177, comprising a plurality of JH5 gene segments, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a nucleotide mutation at one or more positions corresponding to positions

106,330,024
106,330,027
106,330,032
106,330,041
106,330,044
106,330,045
106,330,062
106,330,063
106,330,065
106,330,066
106,330,067
106,330,068 and
106,330,071

on human chromosome 14.

179. The vertebrate, cell or population of claim 178, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a guanine at a position corresponding to position 106,330,067 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

180. The vertebrate, cell or population of claim 179, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,071 on human chromosome 14 (optionally the additional mutation being a guanine); (ii) position 106,330,066 on human

chromosome 14 (optionally the additional mutation being a guanine); and/or (iii) position 106,330,068 on human chromosome 14 (optionally the additional mutation being a thymine).

181. The vertebrate, cell or population of claim 178, 179 or 180, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a guanine at a position corresponding to position 106,330,071 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

182. The vertebrate, cell or population of claim 181, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,063 on human chromosome 14 (optionally the additional mutation being an adenine); and/or (ii) position 106,330,067 on human chromosome 14 (optionally the additional mutation being a guanine).

183. The vertebrate, cell or population of claim 178 to 182, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a cytosine at a position corresponding to position 106,330,045 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

184. The vertebrate, cell or population of claim 178 to 183, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises an adenine at a position corresponding to position 106,330,044 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

185. The vertebrate, cell or population of claim 184, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,066 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,330,068 on human chromosome 14 (optionally the additional mutation being a thymine).

186. The vertebrate, cell or population of claim 178 to 185, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a guanine at a position corresponding to position 106,330,066 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

187. The vertebrate, cell or population of claim 186, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,067 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,330,068 on human chromosome 14 (optionally the additional mutation being a thymine).

188. The vertebrate, cell or population of claim 178 to 185, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a thymine at a position corresponding to position 106,330,068 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

189. The vertebrate, cell or population of claim 188, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,330,067 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,330,066 on human chromosome 14 (optionally the additional mutation being a guanine).

190. The vertebrate, cell or population of claim 178 to 189, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a cytosine at a position corresponding to position 106,330,027 on human chromosome 14; and optionally no further

mutation from the sequence of SEQ ID NO: 1.

191. The vertebrate, cell or population of claim 178 to 190, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises an adenine at a position corresponding to position 106,330,024 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

192. The vertebrate, cell or population of claim 178 to 191, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a thymine at a position corresponding to position 106,330,032 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

193. The vertebrate, cell or population of claim 178 to 192, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises a thymine at a position corresponding to position 106,330,041 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

194. The vertebrate, cell or population of claim 178 to 193, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1, wherein the variant comprises an adenine or thymine at a position corresponding to position 106,330,063 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

195. The vertebrate, cell or population of claim 194, wherein the variant comprises additionally a mutation at a position corresponding to position 106,330,071 on human chromosome 14 (optionally the additional mutation being a guanine).

196. The vertebrate, cell or population of claim 178 to 195, wherein the plurality comprises a human JH5 gene variant of SEQ ID NO: 1,wherein the variant comprises a cytosine at a position corresponding to position 106,330,062 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 1.

197. The vertebrate, cell or population of claim 178 to 196, wherein the genome comprises SEQ ID NO:1; optionally in *cis* at the same Ig locus as one, two or more of the variants.

198. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 96 to 130, comprising a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a nucleotide mutation at one or more positions corresponding to positions

106,329,411
106,329,413
106,329,414
106,329,417
106,329,419
106,329,426
106,329,434
106,329,435, and
106,329,468

on human chromosome 14.

199. The vertebrate, cell or population of claim 198, comprising a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a guanine at a position corresponding to position 106,329,435 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

200. The vertebrate, cell or population of claim 199, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,329,468 on human chromosome 14 (optionally the additional mutation being a guanine); (ii) position 106,329,419 on human chromosome 14 (optionally the additional mutation being an adenine); (iii) position 106,329,434 on human chromosome 14 (optionally the additional mutation being a cytosine) and/or position 106,329,414 on human chromosome 14 (optionally the additional mutation being a guanine); (iv) position 106,329,426 on human chromosome 14 (optionally the additional mutation being an adenine); (v) position 106,329,413 on human chromosome 14 (optionally the additional mutation being an adenine); (vi) position 106,329,417 on human chromosome 14 (optionally the additional mutation being a thymine); (vii) position 106,329,411 on human chromosome 14 (optionally the additional mutation being a thymine); (viii) position 106,329,451 on human chromosome 14 (optionally the additional mutation being an adenine); (ix) position 106,329,452 on human chromosome 14 (optionally the additional mutation being a cytosine); and/or (x) position 106,329,453 on human chromosome 14 (optionally the additional mutation being a cytosine).

201. The vertebrate, cell or population of claim 200, wherein the variant comprises additionally mutations at positions corresponding to position 106,329,451 on human chromosome 14, the additional mutation being an adenine; position 106,329,452 on human chromosome 14, the additional mutation being a cytosine; and position 106,329,453 on human chromosome 14, the additional mutation being a cytosine.

202. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 201, comprising a plurality of JH6 gene segments,, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a guanine at a position corresponding to position 106,329,468 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

203. The vertebrate, cell or population of claim 202, wherein the variant comprises additionally a mutation at a position corresponding to position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

204. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 203, comprising a plurality of JH6 gene segments,, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a thymine at a position corresponding to position 106,329,417 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

205. The vertebrate, cell or population of claim 204, wherein the variant comprises additionally a mutation at a position corresponding to position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

206. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 205, comprising a plurality of JH6 gene segments,, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a cytosine at a position corresponding to position 106,329,434 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

207. The vertebrate, cell or population of claim 206, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,329,414 on human chromosome 14 (optionally the additional mutation being a guanine); and/or (ii) position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

208. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 207, comprising a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant of SEQ ID NO: 2, wherein the variant comprises a thymine at a position corresponding to position 106,329,411 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 2.

209. The vertebrate, cell or population of claim 208, wherein the variant comprises additionally a mutation at a position corresponding to position 106,329,435 on human chromosome 14 (optionally the additional mutation being a guanine).

210. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 209, comprising a plurality of JH6 gene segments, wherein the plurality comprises a human JH6 gene variant that is an antisense sequence of the variant of any one of claims 107 to 118.

211. The vertebrate, cell or population of any one of claims 198 to 210, wherein the genome comprises SEQ ID NO:2; optionally *cis* at the same Ig locus as one, two or more of the JH6 variants.

212. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 211, comprising a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises a nucleotide mutation at one or more positions corresponding to positions

106,331,455
106,331,453, and
106,331,409

on human chromosome 14.

213. The vertebrate, cell or population of claim 212, comprising said plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises a guanine at a position corresponding to position 106,331,455 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 3.

214. The vertebrate, cell or population of claim 213, wherein the variant comprises additionally a mutation at a position corresponding to (i) position 106,331,453 on human chromosome 14 (optionally the additional mutation being an adenine); and/or (ii) position 106,331,409 on human chromosome 14 (optionally the additional mutation being an adenine); (iii) position 106,329,434 on human chromosome 14 (optionally the additional mutation being an adenine).

215. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 214, comprising a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises an adenine at a position corresponding to position 106,331,453 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 3.

216. The vertebrate, cell or population of claim 215, wherein the variant comprises additionally a mutation at a position corresponding to position 106,331,409 on human chromosome 14 (optionally the additional mutation being an adenine).

217. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 216, comprising a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant of SEQ ID NO: 3, wherein the variant comprises an adenine at a position corresponding to position 106,331,409 on human chromosome 14; and optionally no further mutation from the sequence of SEQ ID NO: 3.

218. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 217, comprising a plurality of JH2 gene segments, wherein the plurality comprises a human JH2 gene variant that is an antisense sequence of the variant of any one of claims 188 to 193.

219. The vertebrate, cell or population of any one of claims 212 to 218, wherein the genome comprises SEQ ID NO:3; optionally *cis* at the same Ig locus as one, two or more of the JH2 variants.

220. The vertebrate, cell or population of any one of claims 147 to 151, 157 to 161 and 163 to 219, wherein the genome comprises two or more different JH gene segments selected from SEQ ID NOs: 1 to 3 and variants according to any one of claims 178 to 219; optionally wherein said JH gene segments are *cis* at the same immunoglobulin Ig locus.

221. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 34, having a genome comprising at least 3 human JL gene segments of the same type (eg, Jκ1), wherein at least two of the human JL gene segments are variants that are not identical to each other.

222. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 35, having a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1) *cis* at the same Ig locus.

223. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 36, having a genome comprising at least 2 different human JL gene segments of the same type (eg, Jκ1) *trans* at the same Ig locus; and optionally a third human JL gene segment of the same type, wherein the third JL is *cis* with one of said 2 different JL gene segments.

224. A population of non-human vertebrates (eg, mice or rats) comprising a repertoire of human JL gene segments according to claim 37, wherein the plurality comprises at least 2 different human JL gene segments of the same type (eg, Jκ1), a first of said different JL gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JL gene segments being provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JL gene segment.

225. A non-human vertebrate (eg, a mouse or rat) or a non-human vertebrate cell (eg, an ES cell or a B-cell) according to claim 38, having a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1), wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

226. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 39, the method comprising providing the vertebrate with a genome comprising at least 3 human JL gene segments of the same type (eg, Jκ1), wherein at least two of the human JL gene segments are variants that are not identical to each other.

227. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 40, the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1) *cis* at the same Ig locus.

228. A method of enhancing the immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 41, the method comprising providing the vertebrate with a genome comprising at least 2 different human JL gene segments of the same type(eg, Jκ1) *trans* at the same Ig locus; and optionally a third human JL gene segment of the same type, wherein the third JL is *cis* with one of said 2 different JL gene segments.

229. A method of providing an enhanced human immunoglobulin JL gene segment repertoire according to claim 42, the method comprising providing a population of non-human vertebrates (eg, a mouse or rat) comprising a repertoire of human JL gene segments, wherein the method comprises providing at least 2 different human JL gene segments of the same type (eg, Jκ1), wherein a first of said different JL gene segments is provided in the genome of a first vertebrate of the population, and a second of said different JL gene segments is provided in the genome of a second vertebrate of the population, wherein the genome of the first vertebrate does not comprise the second JL gene segment.

230. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat) according to claim 43, the method comprising providing the vertebrate with a genome comprising at least 2 different non-endogenous JL gene segments of the same type (eg, Jκ1), wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

231. The vertebrate or cell of claim 221, 223 or 225, or the method of claim 226, 228, 229 or 230, wherein at least 2 or 3 of said different gene segments are provided *cis* at the same Ig locus in said genome.

232. The vertebrate or cell of claim 221, 222 or 223, or the method of claim 226, 227, 228 or 229, wherein the JL gene segments are derived from the genome sequence of different human

individuals that are not genetically related over at least 3 generations.

233. The vertebrate or cell of claim 221, 222 or 223, or the method of claim 226, 227, 228 or 229, wherein the JL gene segments are derived from the genome sequence of two or more different human individuals; optionally wherein the different human individuals are from different human populations.

234. The vertebrate, cell or method of claim 233, wherein the individuals are not genetically related.

235. The vertebrate, cell or method of any one of claims 221 to 234, wherein at least one of the different JL segments is a synthetic mutant of a human germline JL gene segment.

236. A method of enhancing the human immunoglobulin gene diversity of a non-human vertebrate (eg, a mouse or rat), the method comprising providing the vertebrate with a genome comprising at least 2 human JL gene segments of the same type (eg, Jκ1), wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations; optionally wherein at least 2 or 3 of said different gene segments are provided at the same IgL locus in said genome.

237. The vertebrate or cell of any one of claims claim 221 to 223 and 225, wherein the genome comprises a substantially complete functional repertoire of human Jκ and/or Jλ gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

238. The population of claim 224, wherein the population comprises a substantially complete functional repertoire of human JL gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

239. A non-human vertebrate (eg, a mouse or rat) or a non-human cell (eg, an ES cell or a B-cell) having a genome comprising a substantially complete functional repertoire of human Jκ and/or Jλ gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

240. A population of non-human vertebrates (eg, mice or rats) comprising a substantially complete functional repertoire of human Jκ and/or Jλ gene segment types supplemented with one, two or more human Jκ and/or Jλ gene segments respectively, wherein said substantially complete functional repertoire and the supplementary gene segments are not found together in the germline genome of a human individual.

241. A non-human vertebrate or vertebrate cell according to claim 222, comprising a genome that comprises VL and JL gene repertoires comprising human gene segments, the JL gene repertoire comprising

a plurality of human Jκ1 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ2 gene segments provided by at least 2 different human Jκ1 gene

segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ3 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ4 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jκ5 gene segments provided by at least 2 different human Jκ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ1 gene segments provided by at least 2 different human Jλ1 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ2 gene segments provided by at least 2 different human Jλ2 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ3 gene segments provided by at least 2 different human Jλ3 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ4 gene segments provided by at least 2 different human Jλ4 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ5 gene segments provided by at least 2 different human Jλ5 gene segments in *cis* at the same Ig locus in said genome;
a plurality of human Jλ6 gene segments provided by at least 2 different human Jλ6 gene segments in *cis* at the same Ig locus in said genome; or
a plurality of human Jλ7 gene segments provided by at least 2 different human Jλ7 gene segments in *cis* at the same Ig locus in said genome;

optionally wherein the JL gene segments are derived from the genome sequence of two or more different human individuals.

242. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell), according to claim 221, comprising a genome that comprises VL and JL gene repertoires comprising human gene segments, the JL gene repertoire comprising

a plurality of human Jκ1 gene segments provided by at least 3 different human Jκ1 gene segments;
a plurality of human Jκ2 gene segments provided by at least 3 different human Jκ1 gene segments;
a plurality of human Jκ3 gene segments provided by at least 3 different human Jκ1 gene segments;
a plurality of human Jκ4 gene segments provided by at least 3 different human Jκ1 gene segments;
a plurality of human Jκ5 gene segments provided by at least 3 different human Jκ1 gene segments;
a plurality of human Jλ1 gene segments provided by at least 3 different human Jλ1 gene segments;
a plurality of human Jλ2 gene segments provided by at least 3 different human Jλ2 gene segments;
a plurality of human Jλ3 gene segments provided by at least 3 different human Jλ3 gene segments;
a plurality of human Jλ4 gene segments provided by at least 3 different human Jλ4 gene segments;
a plurality of human Jλ5 gene segments provided by at least 3 different human Jλ5 gene segments;
a plurality of human Jλ6 gene segments provided by at least 3 different human Jλ6 gene segments; or

a plurality of human Jλ7 gene segments provided by at least 3 different human Jλ7 gene segments;

optionally wherein the JL gene segments are derived from the genome sequence of two or three different human individuals;

optionally wherein at least 2 or 3 of said different gene segments are provided in *cis* at the same Ig locus in said genome.

243. The vertebrate or cell of claim 241 or 242, wherein the different human individuals are from different human populations.

244. The vertebrate or cell of any one of claims 241 to 243, wherein the individuals are not genetically related.

245. The vertebrate or cell of any one of claims 241 to 244, wherein at least one of the different JL segments is a synthetic mutant of a human germline JL gene segment.

246. A non-human vertebrate or vertebrate cell (optionally an ES cell or B-cell) according to claim 225, comprising a genome comprising human VL and JL gene repertoires, the JL gene repertoire comprising

a plurality of human Jκ1 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ2 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ3 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ4 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jκ5 gene segments provided by at least 2 different human Jκ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ1 gene segments provided by at least 2 different human Jλ1 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ2 gene segments provided by at least 2 different human Jλ2 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ3 gene segments provided by at least 2 different human Jλ3 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ4 gene segments provided by at least 2 different human Jλ4 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ5 gene segments provided by at least 2 different human Jλ5 gene segments, optionally in *cis* at the same Ig locus in said genome;
a plurality of human Jλ6 gene segments provided by at least 2 different human Jλ6 gene segments, optionally in *cis* at the same Ig locus in said genome; or
a plurality of human Jλ7 gene segments provided by at least 2 different human Jλ7 gene segments, optionally in *cis* at the same Ig locus in said genome;

wherein the JL gene segments are derived from the genome sequence of different human individuals that are not genetically related over at least 3 generations.

247. The vertebrate or cell of claim 246, wherein the human individuals are from different human populations.

248. A method of producing an antibody heavy chain, the method comprising

(a) providing an antigen-specific heavy chain variable domain; and
(b) combining the variable domain with a human heavy chain constant region to produce an antibody heavy chain comprising (in N- to C-terminal direction) the variable domain and the constant region;

wherein

the human heavy chain constant region is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region.

249. The method of claim 248, wherein the variable domain is a human variable domain.

250. The method of any claim 248 or 249, wherein the variable domain has previously been selected from a non-human vertebrate that has been immunised with the antigen.

251. The method of any one of claims 248 to 250, comprising providing an expression vector comprising a nucleotide sequence encoding the constant region; inserting a nucleotide sequence encoding the variable domain into the vector 5' of the constant region sequence; inserting the vector into a host cell and expressing the heavy chain by the host cell; the method further comprising isolating a heavy chain (eg, as part of an antibody) comprising the variable domain and the human constant region.

252. The method of any one of claims 248 to 251, further comprising obtaining a nucleotide sequence encoding the heavy chain.

253. An antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region.

254. A polypeptide comprising (in N- to C- terminal direction) a leader sequence, a human variable domain that is specific for an antigen and a human constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region; wherein (i) the leader sequence is not the native human variable domain leader sequence; and/or (ii) the variable domain comprises mouse AID-pattern somatic mutations or mouse terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations.

255. A nucleotide sequence encoding (in 5' to 3' direction) a leader sequence and a human antibody heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region; and the leader sequence being operable for expression of the heavy chain and wherein the leader sequence is not the native human variable domain leader sequence.

256. A nucleotide sequence encoding (in 5' to 3' direction) a promoter and a human antibody heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region that is an IGHAref, IGHA1a, IGHA2a, IGHA2b, IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref, IGHG4a, IGHDref, IGHEref, IGHMref, IGHMa or IGHMb constant region; and the promoter being operable for expression of the heavy chain and wherein the promoter is not the native human promoter.

257. The antibody, polypeptide or nucleotide sequence of any one of claims 253 to 256, wherein the variable domain comprises mouse AID-pattern somatic mutations or mouse terminal deoxynucleotidyl transferase (TdT)- pattern junctional mutations.

258. A vector (eg, a CHO cell or HEK293 cell vector) comprising the nucleic acid of claim 255, 256 or 257; optionally wherein the vector is in a host cell (eg, a CHO cell or HEK293 cell).

259. A pharmaceutical composition comprising the antibody or polypeptide of any one of claims 253, 254 and 257, together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

260. The antibody or polypeptide of any one of claims 253, 254 and 257 for use in treating and/or preventing a medical condition in a human patient.

261. Use of the antibody or polypeptide of any one of claims 253, 254 and 257 for the manufacture of a medicament for treating and/or preventing a medical condition in a human patient.

262. The antibody, polypeptide or use of claim 260 or 261, wherein the human is a member of a human population selected from population numbers 1-14, wherein the populations are numbered as follows (population labels being according to 1000 Genomes Project nomenclature)
1= ASW;
2= CEU;
3=CHB;
4=CHS;
5=CLM;

6=FIN;
7=GBR;
8=IBS;
9=JPT;
10=LWK;
11=MXL;
12=PUR;
13=TSI;
14=YRI.

263. The antibody, polypeptide or use of claim 262, wherein the constant region is a

(i) IGHA1a constant region and the human population is selected from any population number 1-14;
(ii) IGHA2a constant region and the human population is selected from any population number 1-14;
(iII) IGHA2b constant region and the human population is selected from any population number 1-14;
(iV) IGHG2a constant region and the human population is selected from any population number 1-9 and 11-13;
(v) IGHG3a constant region and the human population is selected from any population number 1-14;
(vi) IGHG3b constant region and the human population is selected from any population number 1-8 and 11-13;
(vii) IGHG4a constant region and the human population is selected from any population number 1-9 and 11-13;
(viii) IGHMa constant region and the human population is selected from any population number 1-14; or
(ix) IGHMb constant region and the human population is selected from any population number 1-14;

Wherein the populations are numbered as follows (population labels being according to 1000 Genomes Project nomenclature)
1= ASW;
2= CEU;
3=CHB;
4=CHS;
5=CLM;
6=FIN;
7=GBR;
8=IBS;
9=JPT;
10=LWK;
11=MXL;
12=PUR;

13=TSI;
14=YRI.

265. A vector (eg, a CHO cell or HEK293 cell vector) comprising a IGHG1ref, IGHG2ref, IGHG2a, IGHG3ref, IGHG3a, IGHG3b, IGHG4ref or IGHG4a constant region nucleotide sequence that is 3' of a cloning site for the insertion of a human antibody heavy chain variable domain nucleotide sequence, such that upon insertion of such a variable domain sequence the vector comprises (in 5' to 3' direction) a promoter, a leader sequence, the variable domain sequence and the constant region sequence so that the vector is capable of expressing a human antibody heavy chain when present in a host cell.

A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) comprising a genome having a superhuman immunoglobulin heavy chain human VH and/or D and/or J gene repertoire.

266. The vertebrate or cell of claim 265, wherein the genome comprises a transgenic immunoglobulin heavy chain locus comprising a plurality of human immunoglobulin VH gene segments, one or more human D gene segments and one or more human J gene segments, wherein the plurality of VH gene segments consists of more than the natural human repertoire of functional VH gene segments; optionally wherein the genome is homozygous for said transgenic heavy chain locus.

267. The vertebrate or cell of claim 266, wherein VH gene repertoire consists of a plurality of VH gene segments derived from the genome sequence of a first human individual, supplemented with one or more different VH gene segments derived from the genome sequence of a second, different human individual; optionally wherein the individuals are not related; optionally wherein the D and J segments are derived from the genome sequence of the first human individual; and optionally wherein the VH gene segments from the genome sequence of the second individual are selected from the VH gene segments listed in Table 1.

268. The vertebrate or cell of claim 267, wherein the transgenic locus comprises more than 26 functional human VH gene segments; optionally wherein the locus comprises at least 27or28functional human VH gene segments (eg, wherein the locus comprises the full functional VH repertoire of said first individual supplemented with one or more VH gene segments derived from the genome sequence of the second human individual and optionally with one or more VH gene segments derived from the genome sequence of a third human individual).

269. The vertebrate of cell of claim 267 or 268, wherein the transgenic locus comprises a first VH gene segment derived from the genome sequence of the first individual and a second VH gene segment derived from the genome sequence of the second individual, wherein the second VH gene segment is a polymorphic variant of the first VH gene segment; optionally wherein the locus comprises a further polymorphic variant of the first VH gene segment (eg, a variant derived from the genome sequence of a third human individual).

270. The vertebrate or cell of any one of claims 265 to 269, wherein the genome comprises a (or said) transgenic immunoglobulin heavy chain locus comprising a plurality of human immunoglobulin VH gene segments, a plurality of human D gene segments and one or more human J gene segments, wherein the plurality of D gene segments consists of more than the natural human repertoire of functional D gene segments; optionally wherein the genome is homozygous for said transgenic heavy chain locus.

271. The vertebrate or cell of claim 270, wherein D gene repertoire consists of a plurality of D gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more different D gene segments derived from the genome sequence of a (or said) second, different human individual; optionally wherein the individuals are not related; optionally wherein J segments are derived from the genome sequence of the first human individual; and optionally wherein the D gene segments from the genome sequence of the second individual are selected from the D gene segments listed in Table 2.

272. The vertebrate or cell of claim 271, wherein the transgenic locus comprises more than 26 functional human D gene segments; optionally wherein the locus comprises at least 27or 28functional human D gene segments (eg, wherein the locus comprises the full functional D repertoire of said first individual supplemented with one or more D gene segments derived from the genome sequence of the second human individual and optionally with one or more D gene segments derived from the genome sequence of a third human individual).

273. The vertebrate of cell of claim 271 or 272, wherein the transgenic locus comprises a first D gene segment derived from the genome sequence of the first individual and a second D gene segment derived from the genome sequence of the second individual, wherein the second D gene segment is a polymorphic variant of the first D gene segment; optionally wherein the locus comprises a further polymorphic variant of the first D gene segment (eg, a variant derived from the genome sequence of a third human individual).

274. The vertebrate or cell of any one of claims 265 to 273, wherein the genome comprises a (or said) transgenic immunoglobulin heavy chain locus comprising a plurality of human immunoglobulin VH gene segments, one or more human D gene segments and a plurality of human JH gene segments, wherein the plurality of J gene segments consists of more than the natural human repertoire of functional J gene segments; optionally wherein the genome is homozygous for said transgenic heavy chain locus.

275. The vertebrate or cell of claim 274, wherein the J gene repertoire consists of a plurality of J gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more different J gene segments derived from the genome sequence of a (or said) second, different human individual; optionally wherein the individuals are not related; optionally wherein D segments are derived from the genome sequence of the first human individual; and optionally wherein the J gene segments from the genome sequence of the second individual are selected from the J gene segments listed in Table 3.

276. The vertebrate or cell of claim 275, wherein the transgenic locus comprises more than 26 functional human J gene segments; optionally wherein the locus comprises at least 27 or 28 functional human J gene segments (eg, wherein the locus comprises the full functional J repertoire of said first individual supplemented with one or more J gene segments derived from the genome sequence of the second human individual and optionally with one or more J gene segments derived from the genome sequence of a third human individual).

277. The vertebrate of cell of claim 275 or 276, wherein the transgenic locus comprises a first J gene segment derived from the genome sequence of the first individual and a second J gene segment derived from the genome sequence of the second individual, wherein the second J gene segment is a polymorphic variant of the first J gene segment; optionally wherein the locus comprises a further polymorphic variant of the first J gene segment (eg, a variant derived from the genome sequence of a third human individual).

278. A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) comprising a genome having a superhuman immunoglobulin light chain human VL gene repertoire; optionally wherein the vertebrate or cell is according to any one of claims 265 to 277.

279. The vertebrate or cell of claim 278, wherein the genome comprises

(i) a transgenic immunoglobulin kappa light chain locus comprising a plurality of human immunoglobulin Vκ gene segments and one or more human J gene segments, wherein the plurality of Vκ gene segments consists of more than the natural human repertoire of functional Vκ gene segments; optionally wherein the genome is homozygous for said transgenic kappa light chain locus; and/or

(ii) a transgenic immunoglobulin lambda light chain locus comprising a plurality of human immunoglobulin Vλ gene segments and one or more human J gene segments, wherein the plurality of Vλ gene segments consists of more than the natural human repertoire of functional Vλ gene segments; optionally wherein the genome is homozygous for said transgenic lambda light chain locus.

280. The vertebrate or cell of claim 279, wherein

(i) the Vκ gene repertoire consists of a plurality of Vκ gene segments derived from the genome sequence of a first human individual, supplemented with one or more Vκ gene segments derived from the genome sequence of a second, different human individual; optionally wherein the individuals are not related; optionally wherein the J segments are derived from the genome sequence of the first human individual; and optionally wherein the Vκ gene segments from the genome sequence of the second individual are selected from the Vκ gene segments listed in Table 4; and

(i) the Vλ gene repertoire consists of a plurality of Vλ gene segments derived from the genome sequence of a first human individual, supplemented with one or more Vλ gene segments derived from the genome sequence of a second, different human individual; optionally wherein the

individuals are not related; optionally wherein the J segments are derived from the genome sequence of the first human individual; and optionally wherein the Vλ gene segments from the genome sequence of the second individual are selected from the Vλ gene segments listed in Table 5.

281. The vertebrate or cell of claim 280, wherein

the kappa light transgenic locus comprises more than 26 functional human Vκ gene segments; optionally wherein the locus comprises at least 27 or 28 functional human Vκ gene segments (eg, wherein the locus comprises the full functional Vκ repertoire of said first individual supplemented with one or more Vκ gene segments derived from the genome sequence of the second human individual and optionally with one or more Vκ gene segments derived from the genome sequence of a third human individual); and

the lambda light transgenic locus comprises more than 26 functional human Vλ gene segments; optionally wherein the locus comprises at least 27 or 28 functional human Vλ gene segments (eg, wherein the locus comprises the full functional Vλ repertoire of said first individual supplemented with one or more Vλ gene segments derived from the genome sequence of the second human individual and optionally with one or more Vλ gene segments derived from the genome sequence of a third human individual);

282. The vertebrate of cell of claim 280 or 281, wherein

the kappa light transgenic locus comprises a first Vκ gene segment derived from the genome sequence of the first individual and a second Vκ gene segment derived from the genome sequence of the second individual, wherein the second Vκ gene segment is a polymorphic variant of the first Vκ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Vκ gene segment (eg, a variant derived from the genome sequence of a third human individual); and

the lambda light transgenic locus comprises a first Vλ gene segment derived from the genome sequence of the first individual and a second Vλ gene segment derived from the genome sequence of the second individual, wherein the second Vλ gene segment is a polymorphic variant of the first Vλ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Vλ gene segment (eg, a variant derived from the genome sequence of a third human individual).

283. The vertebrate or cell of any one of claims 278 to 282, wherein the genome comprises a (or said) transgenic immunoglobulin light chain locus comprising a plurality of human immunoglobulin VL gene segments and a plurality of human JL gene segments, wherein the plurality of J gene segments consists of more than the natural human repertoire of functional J gene segments; optionally wherein the genome is homozygous for said transgenic heavy chain locus.

284. The vertebrate or cell of claim 283, wherein

(i) the Jκ gene repertoire consists of a plurality of Jκ gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more Jκ gene segments derived from the genome sequence of a (or said) second, different human individual; optionally wherein the individuals are not related; optionally wherein the Vκ segments are derived from the genome sequence of the first human individual; optionally wherein the Jκ gene segments from the genome sequence of the second individual are selected from the Jκ gene segments listed in Table 6; and

(ii) the Jκ gene repertoire consists of a plurality of Jλ gene segments derived from the genome sequence of a (or said) first human individual, supplemented with one or more Jλ gene segments derived from the genome sequence of a (or said) second, different human individual; optionally wherein the individuals are not related; optionally wherein the Vλ segments are derived from the genome sequence of the first human individual; optionally wherein the Jλ gene segments from the genome sequence of the second individual are selected from the Jλ gene segments listed in Table 7.

285. The vertebrate or cell of claim 284, wherein

(i) the transgenic light chain locus comprises more than 26 functional human Jκ gene segments; optionally wherein the locus comprises at least 27 or 28functional human Jκ gene segments (eg, wherein the locus comprises the full functional Jκ repertoire of said first individual supplemented with one or more Jκ gene segments derived from the genome sequence of the second human individual and optionally with one or more Jκ gene segments derived from the genome sequence of a third human individual); and/or

(i) the transgenic light chain locus comprises more than 26 functional human Jλ gene segments; optionally wherein the locus comprises at least 27 or 28 functional human Jλ gene segments (eg, wherein the locus comprises the full functional Jλ repertoire of said first individual supplemented with one or more Jλ gene segments derived from the genome sequence of the second human individual and optionally with one or more Jλ gene segments derived from the genome sequence of a third human individual).

286. The vertebrate or cell of claim 284 or 285, wherein

(i) the kappa light transgenic locus comprises a first Jκ gene segment derived from the genome sequence of the first individual and a second Jκ gene segment derived from the genome sequence of the second individual, wherein the second Jκ gene segment is a polymorphic variant of the first Jκ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Jκ gene segment (eg, a variant derived from the genome sequence of a third human individual); and

(ii) the lambda light transgenic locus comprises a first Jλ gene segment derived from the genome sequence of the first individual and a second Jλ gene segment derived from the genome

sequence of the second individual, wherein the second Jκ gene segment is a polymorphic variant of the first Jλ gene segment; optionally wherein the locus comprises a further polymorphic variant of the first Jλ gene segment (eg, a variant derived from the genome sequence of a third human individual).

287. A library of antibody-producing transgenic cells whose genomes collectively encode a repertoire of antibodies, wherein

(a) a first transgenic cell expresses a first antibody having a chain encoded by a first immunoglobulin gene, the gene comprising a first variable domain nucleotide sequence produced following recombination of a first human unrearranged immunoglobulin gene segment;

(b) a second transgenic cell expresses a second antibody having a chain encoded by a second immunoglobulin gene, the second gene comprising a second variable domain nucleotide sequence produced following recombination of a second human unrearranged immunoglobulin gene segment, the first and second antibodies being non-identical;

(c) the first and second gene segments are different and derived from the genome sequences of first and second human individuals respectively, wherein the individuals are different; and optionally not related;

(d) wherein the cells are non-human vertebrate (eg, mouse or rat) cells.

288. The library of claim 23, wherein in step (c) the individuals or not related.

289. The vertebrate, cell or library of any one of claims 265 to 288, wherein the first and second human individuals are members of first and second ethnic populations respectively, wherein the populations are different; optionally wherein the ethnic populations are selected from those identified in the 1000 genomes database.

290. The vertebrate, cell or library of claim 289, wherein the human immunoglobulin gene segment derived from the genome sequence of the second individual is low-frequency (optionally rare) within the second ethnic population.

291. The library of claim 289 or 290, wherein the first variable region nucleotide sequence is produced by recombination of the first human immunoglobulin gene segment with a first J gene segment and optionally a first D gene segment, wherein the first human immunoglobulin gene segment is a V gene segment and the V, D and J segments are derived from the first human population, optionally from the genome of one individual of the first human population.

292. The library of claim 291, wherein the second variable region nucleotide sequence is produced by recombination of the second human immunoglobulin gene segment with a second J gene segment and optionally a second D gene segment, wherein the second human immunoglobulin gene segment is a V gene segment derived from the second population and the

D and/or J segments are derived from the first human population, optionally the D and J gene segments being from the genome of one individual of the first human population.

293. The library of claim 292, wherein all of the D and J segments that have been recombined with the first and second V gene segments are D and J segments derived from the first human population, optionally the D and J gene segments being from the genome of one individual of the first human population.

294. The vertebrate, cell or library of any one of claims 289 to 293, wherein the first and second ethnic populations are selected from the group consisting of an ethnic population with European ancestry, an ethnic population with East Asian, an ethnic population with West African ancestry, an ethnic population with Americas ancestry and an ethnic population with South Asian ancestry.

295. The vertebrate, cell or library of claim 294, wherein the first and second ethnic populations are selected from the group consisting of an ethnic population with Northern European ancestry; or an ethnic population with Western European ancestry; or an ethnic population with Toscani ancestry; or an ethnic population with British ancestry; or an ethnic population with Icelandic ancestry; or an ethnic population with Finnish ancestry; or an ethnic population with Iberian ancestry; or an ethnic population with Japanese ancestry; or an ethnic population with Chinese ancestry; or an ethnic population Vietnamese ancestry; or an ethnic population with Yoruba ancestry; or an ethnic population with Luhya ancestry; or an ethnic population with Gambian ancestry; or an ethnic population with Malawian ancestry; or an ethnic population with Native American ancestry; or an ethnic population with Afro-Caribbean ancestry; or an ethnic population with Mexican ancestry; or an ethnic population with Puerto Rican ancestry; or an ethnic population with Columbian ancestry; or an ethnic population with Peruvian ancestry; or an ethnic population with Ahom ancestry; or an ethnic population with Kayadtha ancestry; or an ethnic population with Reddy ancestry; or an ethnic population with Maratha; or an ethnic population with Punjabi ancestry.

296. The library of claim 287, wherein the second human immunoglobulin gene segment is a polymorphic variant of the first human immunoglobulin gene segment; optionally wherein the second gene segment is selected from the group consisting of a gene segment in any of Tables 1 to 7 and 9 to 14, eg, the second gene segment is a polymorphic variant of VH1-69.

297. The library of any one of claims 287 to 296, wherein the first and second human immunoglobulin gene segments are both (i) V_H gene segments; (ii) D segments; (iii) J segments (optionally both J_H segments, both J_K segments or both J_λ segments); (iv) constant regions (optionally both a gamma constant region, optionally both a C gamma-1 constant region); (v) CH1 regions; (vi) CH2 regions; or (vii) CH3 regions.

298. The library of any one of claims 287 to 297, wherein the library is naive and optionally has a library size of from 10^2 to 10^9 cells.

299. The library of any one of claims 287 to 298, wherein the library has been selected against a predetermined antigen optionally has a library size of from 10^2 to 10^9 cells.

300. The library of any one of claims 287 to 299, wherein said first and second cells are progeny of first and second ancestor non-human vertebrate cells respectively, wherein the first ancestor cell comprises a genome comprising said first human immunoglobulin gene segment; and the second ancestor cell comprises a genome comprising said second human immunoglobulin gene segment.

301. A library of antibody-producing transgenic cells whose genomes collectively encode a repertoire of antibodies, wherein the library comprises the first and second ancestor cells recited in claim 300.

302. A library of hybridoma cells produced by fusion of the library of any one of claims 265 to 301 with fusion partner cells; and optionally the hybridoma library has a size of from 10^2 to 10^9 cells

303. An isolated antibody having

(a) a heavy chain encoded by a nucleotide sequence produced following recombination in a transgenic non-human vertebrate cell of an unrearranged human immunoglobulin V gene segment with a human D and human J segment, optionally with affinity maturation in said cell, wherein one of the gene segments is derived from the genome of an individual from a first human ethnic population; and the other two gene segments are derived from the genome of an individual from a second, different, human ethnic population, and wherein the antibody comprises heavy chain constant regions of said non-human vertebrate (eg, rodent, mouse or rat heavy chain constant regions); and/or

(b) a light chain encoded by a nucleotide sequence produced following recombination in a transgenic non-human vertebrate cell of an unrearranged human immunoglobulin V gene segment with a human J segment, optionally with affinity maturation in said cell, wherein one of the gene segments is derived from the genome of an individual from a first human ethnic population (optionally the same as the first population in (a)); and the other gene segment is derived from the genome of an individual from a second, different, human ethnic population (optionally the same as the second population in (a)), and wherein the antibody comprises light chain constant regions of said non-human vertebrate (eg, rodent, mouse or rat heavy light constant regions);

(c) Optionally wherein each variable domain of the antibody is a human variable domain.

(d) Optionally wherein the heavy chain constant regions are gamma-type constant regions.

304. An isolated nucleotide sequence encoding the antibody of claim 303, optionally wherein the sequence is provided in an antibody expression vector, optionally in a host cell.

305. A method of producing a human antibody, the method comprising replacing the non-human vertebrate constant regions of the antibody of claim 303 with human antibody constant regions.

306. A pharmaceutical composition comprising an antibody according to claim 303, or an antibody produced according to the method of claim 305 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.

307. An antibody-producing cell that expresses the second antibody recited in any one of claims 287 to 302 or the isolated antibody of claim 303.

308. A non-human vertebrate or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises a transgenic immunoglobulin locus (eg, a heavy chain locus or a light chain locus), said locus comprising immunoglobulin gene segments according to the first and second human immunoglobulin gene segments (optionally V segments) recited in any one of claims 287 to 302 operably connected upstream of an immunoglobulin constant region; optionally wherein the genome is homozygous for said transgenic immunoglobulin locus;

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

309. A transgenic non-human vertebrate (eg, a mouse or rat) or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises a transgenic immunoglobulin locus comprising a plurality of human immunoglobulin gene segments operably connected upstream of a non-human vertebrate constant region for the production of a repertoire of chimaeric antibodies, or chimaeric light or heavy chains, having a non-human vertebrate constant region and a human variable region;

wherein the transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments, a first (optionally a V segment) of said gene segments and a second (optionally a V segment) of said gene segments being different and derived from the genomes of first and second human individuals respectively, wherein the individuals are different; and optionally not related;

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or

optionally wherein the immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

310. A transgenic non-human vertebrate (eg, a mouse or rat) or vertebrate cell (optionally an ES cell or antibody-producing cell) whose genome comprises first and second transgenic immunoglobulin loci, each locus comprising a plurality of human immunoglobulin gene segments operably connected upstream of a non-human vertebrate constant region for the production of a repertoire of chimaeric antibodies, or chimaeric light or heavy chains, having a non-human vertebrate constant region and a human variable region;

wherein (i) the first transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments, (ii) the second transgenic locus comprises one or more human immunoglobulin V gene segments, one or more human J gene segments and optionally one or more human D gene segments; and (iii) wherein a first (optionally a V) gene segment of said first locus and a second (optionally a V) gene segment of said second gene locus are different and derived from the genomes of first and second human individuals respectively, wherein the individuals are different; and optionally not related

optionally wherein the first and second loci are on different chromosomes (optionally chromosomes with the same chromosome number) in said genome;

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional V gene segments; and/or

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional D gene segments; and/or

optionally wherein each immunoglobulin locus comprises more than the natural human complement of functional J gene segments.

311. The non-human vertebrate or cell of claim 309 or 310, wherein the immunoglobulin gene segments are as recited in any one of claims 23 to 37.

312. The non-human vertebrate or cell of claim 308, 309, 310 or 311, wherein the genome comprises a third immunoglobulin gene segment (optionally a V segment), the third gene segment being derived from a human individual that is different from the individual from which the first (and optionally also the second) gene segment is derived; optionally wherein the first, second and third gene segments are polymorphic variants of a human immunoglobulin gene segment.

313. The non-human vertebrate or cell of any one of claims 308 to 312, wherein the genome of the vertebrate or cell is homozygous for the first, second and optional third gene segment,

wherein a copy of the first, second and optional third gene segments are provided together on the same chromosome operably connected upstream of a common non-human vertebrate constant region.

314. The non-human vertebrate or cell of any one of claims 308 to 313, wherein each first, second and optional third gene segment is a V gene segment.

315. The library of any one of claims 287 to 302, wherein the library is provided by a collection of non-human vertebrates (optionally a collection of rodents, mice or rats); optionally, wherein a first member of said collection produces said first antibody but not said second antibody, and a second member of the collection produces said second antibody (but optionally not said first antibody).

316. A repertoire of antibodies expressed from a library of cells according to any one of claims 287 to 302 and 315.

317. A method of constructing a cell (eg, an ES cell) according to any one of claims 265 to 286, the method comprising

(a) identifying functional V and J (and optionally D) gene segments of the genome sequence of a (or said) first human individual;

(b) identifying one or more functional V and/or D and/or J gene segments of the genome sequence of a (or said) second human individual, wherein these additional gene segments are not found in the genome sequence of the first individual;

(c) and constructing a transgenic immunoglobulin locus in the cell, wherein the gene segments of (a) and (b) are provided in the locus operably connected upstream of a constant region.

318. The method of claim 317, wherein the cell comprises a heavy chain locus constructed according to steps (a) to (c) and/or a light chain locus (kappa and/or lambda loci) constructed according to steps (a) to (c).

319. The method of claim 317 or 318, wherein the cell is homozygous for the or each transgenic locus; optionally wherein antibody expression from loci endogenous to said cell has been inactivated.

320. The method of any one of claims 317 to 319, wherein the gene segment(s) in step (b) are identified from an immunoglobulin gene database selected from the 1000 genomes, Ensembl, Genbank and IMGT databases.

321. The method of any one of claims 317 to 320, wherein the first and second human individuals are members of first and second ethnic populations respectively, wherein the populations are different, optionally wherein the human immunoglobulin gene segment derived from the genome sequence of the second individual is low-frequency (optionally rare) within the second ethnic population.

322. A method of making a transgenic non-human vertebrate (eg, a mouse or rat), the method comprising
(a) constructing an ES cell (eg, a mouse C57BL/6N, C57BL/6J, 129S5 or 129Sv strain ES cell) by carrying out the method of any one of claims 317 to 321;
(b) injecting the ES cell into a donor non-human vertebrate blastocyst (eg, a mouse C57BL/6N, C57BL/6J, 129S5 or 129Sv strain blastocyst);
(c) implanting the blastocyst into a foster non-human vertebrate mother (eg, a C57BL/6N, C57BL/6J, 129S5 or 129Sv strain mouse); and
(d) obtaining a child from said mother, wherein the child genome comprises a transgenic immunoglobulin locus.

323. A transgenic non-human vertebrate (eg, a mouse or rat) made by the method of claim 322 or a progeny thereof.

324. A population of non-human vertebrates according to claim 323.

325. A method of isolating an antibody that binds a predetermined antigen (eg, a bacterial or viral pathogen antigen), the method comprising
(a) providing a vertebrate (optionally a mouse or rat) according to any one of claims 265 to 286, 289, 290, 294, 295, 308 to 314 and 323;
(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; and
(f) optionally producing a derivative or variant of the antibody.

326. The method of claim 325, further comprising after step (e) 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.

327. A non-human vertebrate (eg, a mouse or rat) or cell (eg, a mouse cell or rat cell) comprising a genome that comprises a transgenic heavy chain immunoglobulin locus, wherein the locus comprises at least 42 (optionally at least 43, 44, 45, 46, 47, 48, 49, 50) functional human VH gene segments, one or more functional human D gene segments, one or more functional human JH gene segments operably connected upstream of a non-human vertebrate constant region (eg, a mouse constant region, eg, a Cmu and/or a C gamma).

328. The vertebrate or cell of claim 327, wherein the locus comprises at least 23 functional human D gene segments and/or at least 6 functional human JH gene segments, optionally at

least a full human repertoire of functional VH, D and JH segments (optionally derived from the genome sequence of the same human individual).

329. A non-human vertebrate (eg, a mouse or rat) or cell (eg, a mouse cell or rat cell) comprising a genome that comprises a transgenic heavy chain immunoglobulin locus, wherein the locus comprises a full human repertoire of functional VH, D and JH segments derived from the genome sequence of the same human individual, supplemented with one or more additional functional human VH, D and/or JH gene segment that is not found in the genome sequence of said human individual.

330. The vertebrate or cell of claim 329, wherein each additional VH, D and JH segment is selected from the group consisting of (i) a gene segment derived from a second, different individual, optionally wherein the first and second individuals are members of different ethnic populations; (ii) a mutant of a human gene segment (optionally a human germline gene segment having up to 5, 10 or 15 mutations); and (iii) a hybrid human gene segment comprising a first and second nucleotide sequences derived from genome sequences of different human individuals or different polymorphic variants of a human immunoglobulin gene segment (eg, variants of a germline human VH, D or JH gene segment).

331. The vertebrate or cell of any one of claims 328 to 330, wherein the genome is homozygous for said heavy chain transgene and optionally endogenous immunoglobulin heavy chain expression has been inactivated.

332. A non-human vertebrate (eg, a mouse or rat) or cell (eg, a mouse cell or rat cell) comprising a genome that comprises a transgenic kappa light chain immunoglobulin locus, wherein the locus comprises at least 39 functional human Vk gene segments and one or more functional human Jk gene segments operably connected upstream of a non-human vertebrate constant region (eg, a mouse constant region).

333. The vertebrate or cell of claim 332, wherein the locus comprises at least a full human repertoire of functional Vk and Jk segments (optionally derived from the genome sequence of the same human individual).

334. A non-human vertebrate (eg, a mouse or rat) or cell (eg, a mouse cell or rat cell) comprising a genome that comprises a transgenic kappa light chain immunoglobulin locus, wherein the locus comprises a full human repertoire of functional Vk and Jk segments derived from the genome sequence of the same human individual, supplemented with one or more additional functional human Vk and/or Jk gene segment that is not found in the genome sequence of said human individual.

335. The vertebrate or cell of claim 334, wherein each additional Vk and Jk segment is selected from the group consisting of (i) a gene segment derived from a second, different individual, optionally wherein the first and second individuals are members of different ethnic populations; (ii) a mutant of a human gene segment (optionally a human germline gene segment having up to 5, 10 or 15 mutations); and (iii) a hybrid human gene segment comprising a first and second nucleotide sequences derived from: genome sequences of different human individuals or

different polymorphic variants of a human immunoglobulin gene segment (eg, variants of a germline human Vk or Jk gene segment).

336. The vertebrate or cell of any one of claims 332 to 335, wherein the genome is homozygous for said kappa light chain transgene and optionally endogenous immunoglobulin kappa light chain expression has been inactivated.

337. A non-human vertebrate (eg, a mouse or rat) or cell (eg, a mouse cell or rat cell) comprising a genome that comprises a transgenic lambda light chain immunoglobulin locus, wherein the locus comprises at least 32 functional human V lambda gene segments and one or more functional human J lambda gene segments operably connected upstream of a non-human vertebrate constant region (eg, a mouse constant region).

338. The vertebrate or cell of claim 337, wherein the locus comprises at least a full human repertoire of functional V lambda and J lambda segments (optionally derived from the genome sequence of the same human individual).

339. A non-human vertebrate (eg, a mouse or rat) or cell (eg, a mouse cell or rat cell) comprising a genome that comprises a transgenic lambda light chain immunoglobulin locus, wherein the locus comprises a full human repertoire of functional V lambda and J lambda segments derived from the genome sequence of the same human individual, supplemented with one or more additional functional human V lambda and/or J lambda gene segment that is not found in the genome sequence of said human individual.

340. The vertebrate or cell of claim 339, wherein each additional V lambda and J lambda segment is selected from the group consisting of (i) a gene segment derived from a second, different individual, optionally wherein the first and second individuals are members of different ethnic populations; (ii) a mutant of a human gene segment (optionally a human germline gene segment having up to 5, 10 or 15 mutations); and (iii) a hybrid human gene segment comprising a first and second nucleotide sequences derived from genome sequences of different human individuals or different polymorphic variants of a human immunoglobulin gene segment (eg, variants of a germline human V lambda or J lambda gene segment).

341. The vertebrate or cell of any one of claims 337 to 340, wherein the genome is homozygous for said lambda light chain transgene and optionally endogenous immunoglobulin lambda light chain expression has been inactivated.

342. A transgenic immunoglobulin locus comprising a synthetic immunoglobulin gene haplotype, the haplotype comprising first and second human gene segments (each being a V, D or J), a switch region and a constant region, wherein

a) the second gene segment is a polymorphic variant of the first gene segment; or
b) the first and second gene segments are derived respectively from genome sequence of individuals from different, first and second, ethnic populations (eg, according to the 1000 Genomes database) and the second gene segment is not found in the first population (eg,

according to the 1000 Genomes database), optionally wherein the second gene segment occurs at low-frequency (or is rare) in the second population; and

Wherein the constant region, and optionally the switch region, are non-human vertebrate (eg, mouse or rat) constant and switch regions (eg, mouse or rat Cmu; mouse or rat Smu; mouse or rat C gamma; or mouse or rat S gamma).

343. The transgenic immunoglobulin locus of claim 342 wherein the transgene comprises a third human immunoglobulin gene segment (a V, D or J) which is (a) a polymorphic variant of the first and second gene segments; or (b) derived from a genome sequence of the second population (and optionally is a low-frequency or rare gene in that population) or is from a third human ethnic population, wherein the third gene segment is not found in the first population (eg, according to the 1000 Genomes database).

344. A non-human vertebrate or cell comprising the transgene of claim 342 or 343; optionally wherein the vertebrate or cell is homozygous for said transgene.

345. An isolated antibody for administration to a Chinese patient, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the constant region is a human constant region selected from a constant region (eg, an IGHG constant region) inTable 13 found in a Chinese population and with a cumulative frequency of at least 5%;, and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

346. The antibody of claim 345 wherein the constant region is a IGHG1a, IGHG2a, IGHG3a, IGHG3b or IGHG4a constant region.

347. The antibody of claim 345 or 346, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

348. The antibody of claim 345, 346 or 347, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

349. The antibody of claim 345, 346, 347 or 348 wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in

a Chinese population and with a cumulative frequency of at least 5%.

350. An isolated VH domain identical to a variable domain as recited in any one of claims 347 to 349, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

351. A pharmaceutical composition comprising the antibody or variable domain of any one of claims 345 to 350 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

352. An isolated antibody for administration to a Chinese patient, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%; and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

353. The antibody of claim 352, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

354. The antibody of claim 352 or 353, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%.

355. An isolated VH domain identical to a variable domain as recited in any one of claims 352 to 354, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

356. A pharmaceutical composition comprising the antibody or variable domain of any one of claims 352 to 355 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

357. An antibody heavy chain or VH domain (eg, provided as part of an antibody) for therapy and/or prophylaxis of a disease or medical condition in a Chinese patient, wherein the heavy chain is a heavy chain produced by the following steps (or is a copy of such a heavy chain):-

(a) Selection of an antigen-specific antibody heavy chain or VH domain from a non-human vertebrate (eg, a mouse or a rat), wherein the heavy chain or VH domain is derived

from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a Chinese population and with a cumulative frequency of at least 5%;

(b) Optional humanisation of the heavy chain by combining the variable domain of the heavy chain with a human constant region; or optional humanisation of the selected VH domain by combining with a human constant region.

358. The antibody heavy chain or VH domain of claim 357, wherein the human constant region is as recited in claim 345 or 346.

359. An antibody heavy chain or VH domain as recited in claim 357 or 358 for use in a medicament for therapy and/or prophylaxis of a disease or medical condition in a Chinese patient.

360. A method of treating and/or preventing a disease or medical condition in a Chinese patient, the method comprising administering to the patient a therapeutically or prophylactically-effective amount of the antibody heavy chain or VH domain as recited in claim 357 or 358.

361. An isolated antibody for administration to a patient of European, East Asian, West African, South Asian or Americas ancestry, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the constant region is a human constant region selected from a constant region (eg, an IGHG constant region) in Table 13 found in a population of European, East Asian, West African, South Asian or Americas ancestry respectively and with a cumulative frequency of at least 5%;, and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

362. The antibody of claim 361 wherein the constant region is a IGHG1a, IGHG2a, IGHG3a, IGHG3b or IGHG4a constant region and the patient is of European ancestry.

363. The antibody of claim 361 or 362, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment, the VH gene segment being selected from a VH in Table 13 found in said population and with a cumulative frequency of at least 5%.

364. The antibody of claim 361, 362 or 363, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in

said population and with a cumulative frequency of at least 5%.

365. The antibody of claim 361, 362, 363 or 364 wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in said population and with a cumulative frequency of at least 5%.

366. An isolated VH domain identical to a variable domain as recited in any one of claims 363 to 365, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

367. A pharmaceutical composition comprising the antibody or variable domain of any one of claims 361 to 366 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

368. An isolated antibody for administration to a patient of European, East Asian, West African or Americas ancestry, the antibody comprising a human heavy chain, the heavy chain comprising a variable domain that is specific for an antigen and a constant region, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in a population of European, East Asian, West African, South Asian or Americas ancestry respectively and with a cumulative frequency of at least 5%; and wherein

(i) the variable domain is derived from the recombination of said human gene segments in a non-human vertebrate (eg, in a mouse or a rat); or (ii) the variable domain comprises non-human vertebrate (eg, mouse or rat) AID-pattern mutations and non-human vertebrate (eg, mouse or rat) terminal deoxynucleotidyl transferase (TdT)-pattern mutations.

369. The antibody of claim 368, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the D gene segment being selected from a D in Table 13 found in said population and with a cumulative frequency of at least 5%.

370. The antibody of claim 368 or 369, wherein the variable domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the JH gene segment being selected from a JH in Table 13 found in said population and with a cumulative frequency of at least 5%.

371. An isolated VH domain identical to a variable domain as recited in any one of claims 368 to 370, optionally fused at its C-terminus to a polypeptide (eg, an antibody Fc).

372. A pharmaceutical composition comprising the antibody or variable domain of any one of claims 368 to 371 together with a pharmaceutically-acceptable excipient, diluent or a medicament (eg, a further antigen-specific variable domain, antibody chain or antibody).

373. An antibody heavy chain or VH domain (eg, provided as part of an antibody) for therapy and/or prophylaxis of a disease or medical condition in a patient of European, East Asian, West African, South Asian or Americas ancestry, wherein the heavy chain is a heavy chain produced by the following steps (or is a copy of such a heavy chain):-

(a) Selection of an antigen-specific antibody heavy chain or VH domain from a non-human vertebrate (eg, a mouse or a rat), wherein the heavy chain or VH domain is derived from the recombination of a human VH gene segment with a human D gene segment and a human JH gene segment , the VH gene segment being selected from a VH in Table 13 found in said population and with a cumulative frequency of at least 5%;

(b) Optional humanisation of the heavy chain by combining the variable domain of the heavy chain with a human constant region; or optional humanisation of the selected VH domain by combining with a human constant region.

374. The antibody heavy chain or VH domain of claim 373, wherein the human constant region is as recited in claim 361 or 362.

375. An antibody heavy chain or VH domain as recited in claim 373 or 374 for use in a medicament for therapy and/or prophylaxis of a disease or medical condition in a patient of said ancestry.

376. A method of treating and/or preventing a disease or medical condition in a patient of European, East Asian, West African, South Asian or Americas ancestry, the method comprising administering to the patient a therapeutically or prophylactically-effective amount of the antibody heavy chain or VH domain as recited in claim 373 or 374.