(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 March 2024 (21.03.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/056044 A1

(51) International Patent Classification:
C12N 15/85 (2006.01) *A01K 67/027* (2006.01)

(21) International Application Number:
PCT/CN2023/118958

(22) International Filing Date:
15 September 2023 (15.09.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/CN2022/119188
16 September 2022 (16.09.2022) CN
PCT/CN2022/136246
02 December 2022 (02.12.2022) CN

(71) Applicant: **BIOCYTOGEN PHARMACEUTICALS (BEIJING) CO., LTD.** [CN/CN]; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN).

(72) Inventors: **HU, Yiqing**; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN). **ZHANG, Qi**; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN). **ZHANG, Lijun**; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN). **ZHANG, Yabo**; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN). **ZHAO, Huizhen**; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN). **YAO, Jiawei**; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN). **SHEN, Yuelei**; No.12, Baoshen South Street, Daxing Bio-Medicine Industry Park, Daxing District, Beijing 102600 (CN).

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

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

WO 2024/056044 A1

(54) Title: GENETICALLY MODIFIED NON-HUMAN ANIMALS AND METHODS FOR PRODUCING HEAVY-CHAIN ANTIBODIES

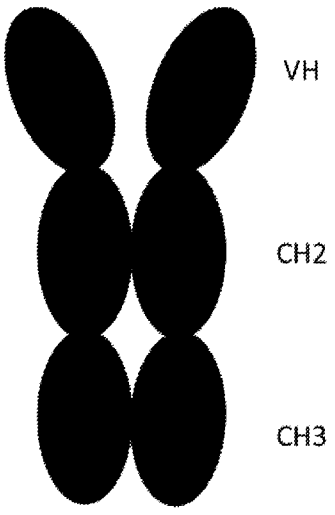


FIG. 1

(57) Abstract: The present disclosure relates to genetically modified animals and methods for producing heavy-chain antibodies. In one aspect, the genetically modified non-human animal comprises a modified immunoglobulin heavy chain locus, wherein the modified immunoglobulin heavy chain locus comprises an IgG constant region gene, wherein the IgG constant region gene encodes an IgG heavy chain constant region lacking a CH1 domain, wherein the genetically modified non-human animal expresses a heavy-chain antibody. Also disclosed herein are anti-TFR1 antibodies, antigen-binding fragments, and the uses thereof.

[Continued on next page]

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

GENETICALLY MODIFIED NON-HUMAN ANIMALS AND METHODS FOR PRODUCING HEAVY-CHAIN ANTIBODIES

CLAIM OF PRIORITY

This application claims priority to PCT/CN2022/119188, filed on September 16, 2022 and PCT/CN2022/136246, filed on December 02, 2022. The entire contents of the foregoing are incorporated herein by reference.

TECHNICAL FIELD

This disclosure relates to genetically modified animals and methods for producing heavy-chain antibodies. The disclosure also relates to anti-TFR1 antibodies, antigen-binding fragments, and the uses thereof.

BACKGROUND

Therapeutic antibodies are one of the fastest growing classes of therapeutic compounds, rapidly outpacing the growth of small-molecule drugs. For example, monoclonal antibodies have revolutionized cancer therapy. However, delivery to tumor cells *in vivo* is hampered by the large size of conventional antibodies. The minimal target recognition module of a conventional antibody is composed of two non-covalently associated variable domains (VH and VL). The inherent hydrophobic interaction of VH and VL domains limits the stability and solubility of engineered antibodies, often causing aggregation and/or mispairing of V-domains.

The discovery of heavy-chain antibodies has given rise to unprecedented opportunities in impacting cancer therapy. These unique form of camelid-derived antibodies lack the entire light chain and the CH1 domain and are only composed of a single variable domain termed VHH. Recombinant VHHs are small (15-20 kDa) and strictly monomeric; they bind their target with nM affinity as well as with being stable in a broad pH and temperature ranges. Molecular manipulation is also easier with VHH; this facilitates the production of multivalent formats of monoclonal antibodies compared with conventional recombinant antibodies and their fragments, which is problematic due to aggregation and reduced affinity. Moreover, VHH often binds to epitopes that are less immunogenic for conventional antibodies.

Usually, the therapeutic antibodies are human or humanized antibodies. The human or humanized antibodies can be generated by humanization of a rodent antibody (e.g., a mouse antibody) or by using phage libraries. However, these animals or phage libraries usually cannot produce heavy chain antibodies. Instead, heavy chain antibodies are often obtained from camelid heavy chain antibodies. These camelid heavy chain antibodies need to be humanized. The humanization process may adversely affect the binding affinity and introduce immunogenic epitopes to the antibodies. Iterative and time-consuming experiments are often required to improve the properties of these antibodies. And in some cases, these antibodies can also be immunogenic in patients, leading to attenuation of their efficacy over time. Therefore, there is a need for an efficient and reliable platform to produce human or humanized heavy chain antibodies and nanobodies.

SUMMARY

The present disclosure relates to genetically-modified animals and cells with humanized immunoglobulin heavy chain variable region locus and truncated immunoglobulin heavy chain constant region locus. For example, the CH1 coding region within the IGHG1 gene can be knocked out such that the expressed IgG does not include the CH1 domain. In some embodiments, the immunoglobulin light chain (e.g., kappa and lambda) loci are also knocked out. Upon immunization, the animals can produce heavy-chain antibodies with high affinity/diversity. In some embodiments, the heavy-chain antibodies can be further processed to generate nanobodies.

In one aspect, the disclosure is related to a genetically modified non-human animal comprising a modified immunoglobulin heavy chain locus, in some embodiments, the modified immunoglobulin heavy chain locus comprises an IgG constant region gene, in some embodiments, the IgG constant region gene encodes an IgG heavy chain constant region lacking a CH1 domain, in some embodiments, the genetically modified non-human animal expresses a heavy-chain antibody. In some embodiments, the animal comprises exactly one IgG constant region gene. In some embodiments, the IgG heavy chain constant region gene is IGHG1. In some embodiments, the IgG heavy chain constant region comprises or consists of a CH2 domain and a CH3 domain, and optionally a hinge region.

In one aspect, the disclosure is related to a genetically modified non-human animal whose genome comprises a germline genetic modification comprising a deletion of IGHG3, IGHG2b, and IGHG2c genes and a deletion of the CH1 exon of IGHG1 gene at an endogenous immunoglobulin heavy chain gene locus. In some embodiments, the germline genetic modification further comprises a deletion of endogenous IGHE gene at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the genetic modification further comprises a deletion of endogenous Sγ2b, Sγ2c, and Sε switch regions at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the modified immunoglobulin heavy chain gene locus comprises a modified IGHG1 gene lacking a sequence encoding a CH1 domain, in some embodiments, the modified IGHG1 gene comprises a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 1. In some embodiments, the genetic modification further comprises a deletion of endogenous Sγ3 switch region at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and endogenous IGHM, IGHδ, IGHA genes. In some embodiments, the genetic modification further comprises a deletion of endogenous IGHM and IGHδ genes at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the animal genome comprises endogenous Sμ, Sγ1, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and the endogenous IGHA gene. In some embodiments, the Sμ and Sγ1 switch regions are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 8. In some embodiments, the genetic modification further comprises a deletion of the CH1 coding sequence of IGHM gene at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, a modified IGHM gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and endogenous IGHδ, IGHA genes.

In some embodiments, the Sμ switch region and the modified IGHM gene are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 10. In some embodiments, the genetic modification further comprises a deletion of the CH1 exon of IGHM gene and a deletion of IGHδ gene at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, a modified IGHM gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and an endogenous IGHA gene. In some embodiments, the genetic modification further comprises a deletion of the CH1 exon of IGHM gene and a deletion of the CH1 coding sequence of IGHδ gene at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, a modified IGHM gene lacking a sequence encoding a CH1 domain, a modified IGHδ gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and an endogenous IGHA gene. In some embodiments, the modified IGHM gene are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 10, and the modified IGHδ gene comprises a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 41. In some embodiments, the modified IGHM gene comprises a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 13. In some embodiments, the genetic modification further comprises a deletion of endogenous Sγ1 switch region at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the animal's genome comprises endogenous Sμ, Sγ3, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and endogenous IGHM, IGHδ, IGHA genes. In some embodiments, the genetic modification further comprises a deletion of endogenous Sγ3 switch region at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the genetic modification further comprises a deletion of endogenous IGHM and IGHδ genes at the endogenous immunoglobulin heavy chain gene locus. In some embodiments, the animal's genome comprises endogenous Sμ, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and the endogenous IGHA gene. In some embodiments, the Sμ switch region and the modified IGHG1 gene are linked with a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 9. In some embodiments, the modified genome comprises a functional IGHM gene.

In some embodiments, an animal can still have the endogenous sequence when the sequence in the genome is replaced by the same sequence or the sequence from the same animal.

In one aspect, the disclosure is related to a genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene. In some embodiments, the elements are operably linked. In some embodiments, the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

In one aspect, the disclosure is related to a genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy

chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene, IGHδ gene, Sγ3 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene. In some embodiments, the elements are operably linked. In some embodiments, the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene, IGHδ gene, Sγ3 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

In one aspect, the disclosure is related to a genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene. In some embodiments, the elements are operably linked. In some embodiments, the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

In one aspect, the disclosure is related to a genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene. In some embodiments, the elements are operably linked. In some embodiments, the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

In one aspect, the disclosure is related to a genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene. In some embodiments, the elements are operably linked. In some embodiments, the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

In one aspect, the disclosure is related to a genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene. In some embodiments, the elements are operably linked. In some embodiments, the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

In one aspect, the disclosure is related to a genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy

chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene. In some embodiments, the elements are operably linked. In some embodiments, the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

In some embodiments, the animal expresses a heavy-chain antibody comprising an IgG heavy chain constant region lacking the CH1 domain. In some embodiments, the heavy-chain antibody binds to it target antigen with a KD of less than 10^{-7} M, less than 10^{-8} M, or less than 10^{-9} M. In some embodiments, the heavy-chain antibody comprises or consists of a variable region, a CH2 domain and a CH3 domain. In some embodiments, the heavy-chain antibody further comprises a transmembrane domain and/or a cytoplasmic domain. In some embodiments, the genetically modified non-human animal does not express IgG antibodies comprising light chains. In some embodiments, the animal expresses IgM, IgD, and/or IgA (e.g., functional IgM, IgD, and/or IgA).

In some embodiments, the animal comprises at an endogenous immunoglobulin heavy chain gene locus, one or more human IGHV genes, one or more human IGHD genes, and one or more human IGHJ genes, in some embodiments, the human IGHV genes, the human IGHD genes, and the human IGHJ genes are operably linked and can undergo VDJ rearrangement. In some embodiments, the animal comprises at least 150 human IGHV genes selected from Table 1, at least 20 human IGHD genes selected from Table 2, and at least 5 human IGHJ genes selected from Table 3. In some embodiments, the animal comprises all human IGHV genes, all human IGHD genes, and all human IGHJ genes at the endogenous immunoglobulin heavy chain gene locus of human chromosome 14 of a human subject. In some embodiments, the animal comprises all human IGHV genes, all human IGHD genes, and all human IGHJ genes at the endogenous immunoglobulin heavy chain gene locus of human chromosome 14 of a human cell. In some embodiments, the animal is a mouse and the genetic modification in the animal's endogenous immunoglobulin heavy chain gene locus comprises a deletion of one or more mouse IGHV genes in Table 4, one or more mouse IGHD genes in Table 5, and/or one or more mouse IGHJ genes in Table 6. In some embodiments, the animal is a mouse and the genetic modification in the animal's endogenous heavy chain immunoglobulin gene locus comprises a deletion of a contiguous sequence starting from mouse IGHV1-85 gene to mouse IGHJ4 gene. In some embodiments, the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus, in some embodiments, the unmodified human sequence is at least 800 kb. In some embodiments, the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHV1-2. In some embodiments, the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHV6-1. In some embodiments, the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHD1-1 to human IGHJ6. In some embodiments, the animal comprises an unmodified

human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHJ6.

In some embodiments, the genetically modified non-human animal described herein has a genome comprising, at the endogenous immunoglobulin heavy chain locus: a replacement of one or more endogenous IGHV, endogenous IGHD, and endogenous IGHJ genes with one or more human IGHV, human IGHD, and human IGHJ genes, in some embodiments, human IGHV, human IGHD, and human IGHJ genes are operably linked to one or more of endogenous IGHM, IGHδ, IGHG1 lacking a sequence encoding the CH1 domain, and IGHA genes. In some embodiments, one or more endogenous IGHV, endogenous IGHD, and endogenous IGHJ genes are replaced by at least 150 human IGHV genes in Table 1, at least 20 human IGHD genes in Table 2, and at least 5 human IGHJ genes in Table 3. In some embodiments, the animal is a mouse, and at least 180 mouse IGHV genes in Table 4, all mouse IGHD genes in Table 5, and all mouse IGHJ genes in Table 6 are replaced. In some embodiments, the animal is homozygous with respect to the immunoglobulin heavy chain gene locus. In some embodiments, the animal is heterozygous with respect to the immunoglobulin heavy chain gene locus. In some embodiments, the animal comprises an endogenous light chain immunoglobulin gene locus. In some embodiments, the animal comprises a disruption in the endogenous immunoglobulin light chain gene locus. In some embodiments, the animal lacks an endogenous immunoglobulin heavy chain variable region locus that is capable of rearranging and forming a nucleic acid sequence that encodes an endogenous heavy chain variable domain. In some embodiments, the animal can produce a humanized antibody. In some embodiments, the animal is a mammal. In some embodiments, the animal is a rodent. In some embodiments, the animal is a mouse. In some embodiments, the animal has substantially normal B cell development and maturation.

In one aspect, the disclosure is related to a cell obtained from the genetically modified non-human animal as described herein. In some embodiments, the cell is a B cell that expresses a chimeric immunoglobulin heavy chain comprising an immunoglobulin heavy chain variable domain that is derived from a rearrangement of one or more human IGHV genes, one or more human IGHD genes, and one or more human IGHJ genes, in some embodiments, the immunoglobulin heavy chain variable domain is operably linked to a non-human heavy chain constant region. In some embodiments, the cell is an embryonic stem (ES) cell.

In one aspect, the disclosure is related to a method of making an antibody that specifically binds to an antigen, the method comprising a) exposing the genetically modified non-human animal as described herein to the antigen; b) producing a hybridoma from a cell collected from the animal; and c) collecting a heavy-chain antibody produced by the hybridoma. In some embodiments, the method further comprises sequencing the genome of the hybridoma.

In one aspect, the disclosure is related to a method of making an antibody that specifically binds to an antigen, the method comprising a) exposing the genetically modified non-human animal as described herein to the antigen; b) sequencing nucleic acids encoding human immunoglobulin heavy chain variable regions in a cell that expresses a heavy-chain antibody that specifically binds to the antigen; and c) operably linking in a cell the nucleic acid encoding the human immunoglobulin heavy chain variable region with a nucleic acid encoding a human immunoglobulin heavy chain constant region.

In one aspect, the disclosure is related to a method of making an antibody that specifically binds to an antigen, the method comprising a) obtaining a nucleic acid sequence encoding human immunoglobulin heavy chain variable regions in a cell that expresses a heavy-chain antibody that specifically binds to the antigen, in some embodiments, the cell is obtained by exposing the genetically modified non-human animal as described herein to the antigen; b) operably linking the nucleic acid encoding the human immunoglobulin heavy chain variable region with a nucleic acid encoding a human immunoglobulin heavy chain constant region; and c) expressing the nucleic acid in a cell, thereby obtaining the antibody.

In one aspect, the disclosure is related to a method of obtaining a nucleic acid that encodes an antibody binding domain that specifically binds to an antigen, the method comprising a) exposing the genetically modified non-human animal as described herein to the antigen; and b) sequencing nucleic acids encoding human immunoglobulin heavy chain variable regions in a cell that expresses a heavy-chain antibody that specifically binds to the antigen.

In one aspect, the disclosure is related to a method of making an antibody that specifically binds to antigen, the method comprising a) exposing the genetically modified non-human animal as described herein to the antigen; b) constructing a phage plasmid library using RNA prepared from immune cells (e.g., splenocytes) of the animal; c) screening the phage plasmid library; and d) sequencing nucleic acids encoding human immunoglobulin heavy chain variable regions from a phage plasmid that encodes a heavy-chain antibody that specifically binds to the antigen. In some embodiments, screening comprises isolating phages expressing immunoglobulin heavy chain variable regions based on binding affinity to the antigen.

In one aspect, the disclosure is related to a method of obtaining a sample, the method comprising a) exposing the genetically modified non-human animal as described herein to the antigen; and b) collecting the sample from the animal. In some embodiments, the sample is an immune cell, a lymphoid tissue, a spleen tissue, a spleen cell, or a B cell.

In one aspect, the disclosure is related to an antibody or antigen-binding fragment thereof that binds to transferrin receptor 1 (TFR1), comprising: a heavy chain single variable domain (VHH) comprising complementarity determining regions (CDRs) 1, 2, and 3, in some embodiments, the VHH CDR1 region comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a selected VHH CDR1 amino acid sequence, the VHH CDR2 region comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a selected VHH CDR2 amino acid sequence, and the VHH CDR3 region comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a selected VHH CDR3 amino acid sequence; in some embodiments, the selected VHH CDRs 1, 2, and 3 amino acid sequences are one of the following:

(1) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 42, 43, and 44, respectively;
(2) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 45, 46, and 47, respectively;
(3) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 48, 49, and 50, respectively;

(4) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 51, 52, and 53, respectively;
(5) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 54, 55, and 56, respectively;
(6) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 57, 58, and 59, respectively;
(7) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 60, 61, and 62, respectively; and
(8) the selected VHH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 63, 64, and 65, respectively.

In some embodiments, the VHH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 42, 43, and 44, respectively. In some embodiments, the VHH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 45, 46, and 47, respectively. In some embodiments, the VHH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 48, 49, and 50, respectively. In some embodiments, the VHH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 51, 52, and 53, respectively.

In one aspect, the disclosure is related to an antibody or antigen-binding fragment thereof that binds to TFR1 comprising a heavy chain single variable region (VHH) comprising an amino acid sequence that is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a selected VHH sequence, in some embodiments, the selected VHH sequence is selected from the group consisting of SEQ ID NOs: 66, 67, 68, and 69. In some embodiments, the VHH comprises the sequence of SEQ ID NO: 66. In some embodiments, the VHH comprises the sequence of SEQ ID NO: 67. In some embodiments, the VHH comprises the sequence of SEQ ID NO: 68. In some embodiments, the VHH comprises the sequence of SEQ ID NO: 69. In some embodiments, the antibody or antigen-binding fragment specifically binds to a human TFR1, a monkey TFR1, a mouse TFR1, or a chimeric TFR1. In some embodiments, the antibody or antigen-binding fragment is a human or humanized antibody or antigen-binding fragment thereof. In some embodiments, the antibody or antigen-binding fragment is a multi-specific antibody (e.g., a bispecific antibody).

In one aspect, the disclosure is relate to an antibody or antigen-binding fragment thereof comprising the VHH CDRs 1, 2, 3, of the antibody or antigen-binding fragment thereof as described herein.

In some embodiments, the antibody or antigen-binding fragment comprises a human IgG Fc (e.g., a human IgG1 Fc). In some embodiments, the human IgG Fc comprises a non-asparagine residue (e.g., alanine) at position 297 according to EU numbering. In some embodiments, the antibody or antigen-binding fragment comprises two or more heavy chain single variable domains.

In one aspect, the disclosure is related to a nucleic acid comprising a polynucleotide encoding the antibody or antigen-binding fragment thereof as described herein. In some embodiments, the nucleic acid is cDNA.

In one aspect, the disclosure is related to a vector comprising one or more of the nucleic acids as described herein.

In one aspect, the disclosure is related to a cell comprising the vector as described herein. In some embodiments, the cell is a CHO cell. In one aspect, the disclosure is related to a cell comprising one or more of the nucleic acids described herein.

In one aspect, the disclosure is related to a method of producing an antibody or an antigen-binding fragment thereof, the method comprising (a) culturing the cell as described herein under conditions sufficient for the cell to produce the antibody or the antigen-binding fragment thereof; and (b) collecting the antibody or the antigen-binding fragment thereof produced by the cell.

In one aspect, the disclosure is related to an antibody-drug conjugate comprising the antibody or antigen-binding fragment thereof as described herein covalently bound to a therapeutic agent. In some embodiments, the therapeutic agent is a cytotoxic or cytostatic agent.

In one aspect, the disclosure is related to a method of treating a subject having a brain disease (e.g., a brain cancer), the method comprising administering a therapeutically effective amount of a composition comprising the antibody or antigen-binding fragment thereof, or the antibody-drug conjugate as described herein, to the subject. In some embodiments, the antibody or antigen-binding fragment thereof, or the antibody-drug conjugate can pass across the blood-brain barrier (BBB) of the subject.

In one aspect, the disclosure is related to a method of treating a subject having a cancer, the method comprising administering a therapeutically effective amount of a composition comprising the antibody or antigen-binding fragment thereof, or the antibody-drug conjugate as described herein, to the subject. In some embodiments, the cancer is brain cancer, lung cancer, gastric cancer, colorectal cancer, liver cancer, ovarian cancer, prostate cancer, leukemia, or breast cancer. In one aspect, the disclosure is related to a method of identifying a subject as having a brain disease (e.g., a brain cancer), the method comprising detecting a sample collected from the subject as having the brain disease by the antibody or antigen-binding fragment thereof as described herein, thereby identifying the subject as having the brain disease. In some embodiments, the sample is a brain parenchyma sample from the subject. In some embodiments, the subject described herein is a human subject.

In one aspect, the disclosure is related to a method of delivering an agent to cross blood brain barrier, the method comprising administering the agent covalently linked to the antibody or antigen-binding fragment thereof as described herein to the subject. In some embodiments, the agent is an antibody or an antibody drug conjugate. In some embodiments, the agent is anti-amyloid antibody.

In one aspect, the disclosure is related to a pharmaceutical composition comprising the antibody or antigen-binding fragment thereof as described herien, and a pharmaceutically acceptable carrier. In one aspect, the disclosure is related to a pharmaceutical composition comprising the antibody drug conjugate as described herein, and a pharmaceutically acceptable carrier.

In one aspect, the disclosure is related to an antibody or antigen-binding fragment thereof that cross-competes with the antibody or antigen-binding fragment thereof as described herein.

In one aspect, the disclosure provides a method of making an antibody that specifically binds to an antigen. The method involves exposing the animal as described herein to the antigen; obtaining the sequence of (e.g. by sequencing) nucleic acids encoding human heavy chain

immunoglobulin variable regions in a cell that expresses a chimeric heavy chain antibody that specifically binds to the antigen; and operably linking in a cell the nucleic acid encoding the human heavy chain immunoglobulin variable region with a nucleic acid encoding a human heavy chain immunoglobulin constant region.

The disclosure also relates to an offspring of the non-human mammal. In some embodiments, the non-human mammal is a rodent. In some embodiments, the non-human mammal is a mouse.

The disclosure also provides to a cell including the targeting vector as described herein. The disclosure also relates to a cell (e.g., a stem cell, an embryonic stem cell, an immune cell, a B cell, a T cell, or a hybridoma) or a cell line, or a primary cell culture thereof derived from the non-human mammal or an offspring thereof. The disclosure further relates to the tissue, organ or a culture thereof derived from the non-human mammal or an offspring thereof.

The disclosure further relates to the use of the non-human mammal or an offspring thereof, the animal model generated through the method as described herein in the development of a product related to an immunization process, the manufacture of a human antibody, or the model system for research in pharmacology, immunology, microbiology and medicine.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Methods and materials are described herein for use in the present invention; other, suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

Other features and advantages of the invention will be apparent from the following detailed description and figures, and from the claims.

DESCRIPTION OF DRAWINGS

FIG. 1 is a schematic diagram of heavy-chain antibody.

FIG. 2 shows schematic structures of IgM, IgD, IgG, IgE, and IgA immunoglobulin isotypes. Mouse and human immunoglobulin heavy chain locus genes are aligned and labelled with corresponding names.

FIG. 3 is a schematic diagram showing the constant region genes of mouse immunoglobulin heavy chain locus. The Cγ1 gene structure including exons CH1, H, CH2, CH3, M1, and M2 are shown.

FIGS. 4A-4B shows mutant alleles after genetic modifications at mouse immunoglobulin heavy chain constant region locus.

FIG. 5 shows gene structure of Cγ1ΔCH1. The CH1 coding region is either missing (left) or replaced by a Neo cassette (right).

FIG. 6 shows a workflow of genetic modifications using targeting vector V1 (Example 3). A 100883 bp sequence from Sγ3 to Cε at mouse immunoglobulin heavy chain constant region locus was replaced in a single step with a 16076 bp sequence including the mouse Sγ1 and the Cγ1ΔCH1 knock-in sequence to generate *Mutant allele 1*.

FIG. 7 shows a workflow of genetic modifications using targeting vector V2 (Example 3). A 92859 bp sequence from Cγ3 to Cε at mouse immunoglobulin heavy chain constant region locus was replaced in a single step with a sequence including the Cγ1ΔCH1 knock-in sequence to generate *Mutant allele 1'*.

FIG. 8 shows a workflow of genetic modifications using targeting vector V3 (Example 3). A 16434 bp sequence including Cμ and Cδ was knocked out from *Mutant allele 1* to generate *Mutant allele 2*.

FIG. 9 shows a workflow of genetic modifications using targeting vector V4 (Example 3). A sequence including Cμ, Cδ, and Sγ1 was knocked out from *Mutant allele 1* such that Sμ and Cγ1ΔCH1' were directly connected, generating *Mutant allele 2'*.

FIG. 10 shows a workflow of genetic modifications using targeting vector V5 (Example 3). The CH1 coding sequence in Cμ was knocked out from *Mutant allele 1* to generate *Mutant allele 3*.

FIG. 11 shows a workflow of genetic modifications using targeting vector V6 (Example 3). The CH1 coding sequence in Cμ and the entire Cδ were knocked out from *Mutant allele 1* to generate *Mutant allele 4*.

FIG. 12 shows a workflow of genetic modifications using targeting vector V7 (Example 3). The CH1 coding sequence in Cμ and the CH1 coding sequence in Cδ were knocked out from *Mutant allele 1* to generate *Mutant allele 5*.

FIGS. 13A-13B show PCR assay results using primer pairs L-GT-F1/L-GT-R1 and R-GT-F2/R-GT-R2, respectively, to verify the genotype of *Mutant allele 1*. WT is a wild-type control. H_2O is a blank control. M is a marker.

FIG. 14A shows Southern Blot results of *Mutant allele 1* positive clones digested with BclI and hybridized with LR probe. M is a marker. WT is wild-type.

FIG. 14B shows Southern Blot results of *Mutant allele 1* positive clones digested with ScaI and hybridized with 3' probe. M is a marker. WT is wild-type.

FIG. 14C shows Southern Blot results of *Mutant allele 1* positive clones digested with XmnI and hybridized with A probe. M is a marker. WT is wild-type.

FIG. 14D shows Southern Blot results of *Mutant allele 1* positive clones digested with BglII and hybridized with 5' probe. M is a marker. WT is wild-type.

FIG. 15 shows PCR assay results using primers DE-F1 and DE-R1 to verify the knock-out of the sequence from Cμ to Cδ in *Mutant allele 2*.

FIG. 16 shows PCR assay results using primers GT-Mut-F, GT-Mut-R and GT-WT-R to verify the knock-out of the sequence from Cμ to Sγ1 in *Mutant allele 2'*.

FIG. 17 shows PCR assay results using primers GT-3F and GT-3R to verify the knock-out of the CH1 coding sequence of Cμ in *Mutant Allele 3*.

FIG. 18A shows PCR assay results using primers Mut-F and Mut-R to verify the sequence of CμΔCH1 in *Mutant Allele 4*. WT is a wild-type control. H_2O is a blank control.

FIG. 18B shows PCR assay results using primers F4 and R4 to verify the absence of Cδ in *Mutant Allele 4*.

FIGS. 19A-19B show PCR assay results using primer pairs Mut-F/Mut-R and F3/R3 to verify the sequence of CμΔCH1 and CδΔCH1, respectively, in *Mutant allele 5*. WT is a wild-type control. H_2O is a blank control.

FIG. 20 is an exemplary flow chart of a method of introducing human immunoglobulin genes into the mouse genome.

FIG. 21 is an overview of replacing mouse immunoglobulin heavy chain variable region locus sequences with human immunoglobulin heavy chain variable region locus sequences.

FIG. 22 shows the length distribution of CDR3 in the heavy chain variable region of antibodies produced by immunizing heterozygous Mut3 mice with antigen A.

FIG. 23 shows the germline gene usage of the variable region genes in mice heterozygous for heavy chain *Mutant allele 3* genotype.

FIG. 24 shows the KD value distribution of antibodies produced in heterozygous Mut3 mice (H/-) against antigen A.

FIG. 25 shows the KD value distribution of antibodies produced in homozygous Mut2 mice against antigen A.

FIG. 26 shows the germline gene usage of the variable region genes in homozygous Mut2 mice.

FIG. 27 is a schematic diagram showing human immunoglobulin heavy chain (IGH) locus on chromosome 14 (14q32.33).

FIG. 28 is a schematic diagram showing mouse (*Mus musculus*) IGH locus on chromosome 12 (12F2) (strain C57BL/6).

FIG. 29 lists IMGT repertoire for human heavy chain immunoglobulin locus (IGH).

FIG. 30 lists IMGT repertoire for mouse IGH.

FIG. 31 lists sequences discussed in the disclosure.

FIG. 32 shows Western blot results of serum IgG levels in mice with *Mutant allele 2'*, *Mutant allele 3*, and *Mutant allele 4* genotypes, respectively. Biot. Ladder is a protein marker. WT represents a wild-type mouse.

FIG. 33 shows the length distribution of CDR3 in the heavy chain variable region of antibodies produced by immunizing heterozygous Mut3 mice with human 4-1BB.

FIG. 34 shows the germline gene usage of the variable region genes in mice homozygous for heavy chain *Mutant allele 3* genotype.

FIG. 35 shows the length distribution of CDR3 in the heavy chain variable region of antibodies produced by immunizing heterozygous Mut3 mice with human CD3ED and cynomolgus CD3ED.

FIG. 36 shows the germline gene usage of the variable region genes in mice homozygous for heavy chain *Mutant allele 3* genotype.

FIG. 37 lists CDR sequences of heavy chain variable region of anti-TFR1 antibodies according to Kabat numbering.

FIG. 38 lists CDR sequences of heavy chain variable region of anti-TFR1 antibodies according to IMGT numbering.

FIG. 39 list amino acid sequences of the heavy chain variable region of the anti-TFR1 antibodies.

FIG. 40A shows the antibody concentration in total brain protein of hTFR1 mice within 72 hours of intravenous (i.v.) administration of hIgG1 (G1), JR141-N (G2), 23B8-N (G3), 24A1-N (G4), 24G5-N (G5), or 24C9-N (G6) by .

FIG. 40B shows the ratio of antibody concentration in brain total protein to serum antibody concentration of hTFR1 mice within 72 hours of intravenous (i.v.) administration of hIgG1 (G1), JR141-N (G2), 23B8-N (G3), 24A1-N (G4), 24G5-N (G5), or 24C9-N (G6).

FIG. 40C shows the antibody concentration in brain parenchyma of hTFR1 mice within 72 hours of intravenous (i.v.) administration of hIgG1 (G1), JR141-N (G2), 23B8-N (G3), 24A1-N (G4), 24G5-N (G5), or 24C9-N (G6).

FIG. 40D shows the ratio of antibody concentration in brain parenchyma to serum antibody concentration of hTFR1 mice within 72 hours of intravenous (i.v.) administration of hIgG1 (G1), JR141-N (G2), 23B8-N (G3), 24A1-N (G4), 24G5-N (G5), or 24C9-N (G6).

FIG. 41A shows the antibody concentration test results in brain parenchyma after 24 hours of intravenous (i.v.) administration of hIgG1 (G1), JR141-N (G2), 23B8-N (G3), 24A1-N (G4), or 24G5-N (G5).

FIG. 41B shows the antibody concentration test results in brain total protein (Whole Brain) after 24 hours of intravenous (i.v.) administration of hIgG1 (G1), JR141-N (G2), 23B8-N (G3), 24A1-N (G4), or 24G5-N (G5).

FIG. 42 shows the antibody concentration results after 6 hours or 24 hours of intravenous (i.v.) administration of JR141-N (G2-G4) or 24G5-N (G5-G7)). hIgG1 was used as a negative control;

FIG. 43 shows the germline gene usage of the variable region genes in mice homozygous for heavy chain Mutant allele 3 genotype;

FIG. 44 shows the germline gene usage of the variable region genes in heterozygous Mut2' mice (H/-);

FIG. 45 shows the KD value distribution of antibodies produced in heterozygous Mut2' mice (H/-) against Human Serum Albumin;

FIG. 46 shows the germline gene usage of the variable region genes in mice homozygous for heavy chain Mutant allele 4 genotype;

FIG. 47 shows the germline gene usage of the variable region genes in mice homozygous for heavy chain Mutant allele 5 genotype.

DETAILED DESCRIPTION

The present disclosure relates to genetically modified animals and methods for producing heavy-chain antibodies.

A heavy-chain antibody (or heavy chain-only antibody) is an antibody which has only heavy chains (generally two heavy chains) and lacks the two light chains usually found in antibodies. Naturally occurring heavy-chain antibodies have been discovered in cartilaginous fishes (e.g., shark) and camelids (e.g., llama). For example, in cartilaginous fishes, the immunoglobulin new antigen receptor (IgNAR) is a heavy-chain antibody. IgNAR shows significant structural differences to other antibodies. It has five constant domains (CH) per chain instead of the usual three, several disulfide bonds in unusual positions, and the complementarity-determining region 3 (CDR3) forms an extended loop covering the site which binds to a light chain in other antibodies. These differences, in combination with the phylogenetic age of the cartilaginous fishes, have led to the hypothesis that IgNAR could be more closely related to a primordial antigen-binding protein than the mammalian immunoglobulins.

The only mammals with heavy-chain (IgG-like) antibodies are camelids such as dromedaries, camels, llamas and alpacas. Like all mammals, camelids (e.g., llamas) can produce conventional antibodies made of two heavy chains and two light chains bound together with disulfide bonds in a Y shape (e.g., IgG1). However, they also produce two unique subclasses of IgG: IgG2 and IgG3, also known as heavy chain IgG. These antibodies are made of only two heavy chains, which lack the CH1 region but still bear an antigen-binding domain (e.g., VHH) at their N-terminus. Conventional Ig require the association of variable regions from both heavy and light chains to allow a high diversity of antigen-antibody interactions. Although isolated heavy and light chains still show this capacity, they exhibit very low affinity when compared to paired heavy and light chains. The unique feature of heavy chain IgG is the capacity of their monomeric antigen binding regions to bind antigens with specificity, affinity and especially diversity that are comparable to conventional antibodies without the need of pairing with another region. This feature is mainly due to a couple of major variations within the amino acid sequence of the variable region of the two heavy chains, which induce deep conformational changes when compared to conventional Ig. Major substitutions in the variable regions prevent the light chains from binding to the heavy chains, but also prevent unbound heavy chains from being recycled by the Immunoglobulin Binding Protein.

The single variable domain of these heavy-chain antibodies (designated VHH, sdAb, or nanobody) is the smallest antigen-binding domain generated by adaptive immune systems. The Complementarity Determining Region 3 (CDR3) of the variable region of these antibodies has often been found to be twice as long as the conventional ones. This results in an increased interaction surface with the antigen as well as an increased diversity of antigen-antibody interactions, which compensates the absence of the light chains. With a long complementarity-determining region 3 (CDR3), VHHs can extend into crevices on proteins that are not accessible to conventional antibodies, including functionally interesting sites such as the active site of an enzyme or the receptor-binding canyon on a virus surface. Moreover, an additional cysteine residue allows the structure to be more stable, thus increasing the strength of the interaction.

VHHs offer numerous other advantages compared to conventional antibodies carrying variable domains (VH and VL) of conventional antibodies, including higher stability, solubility, expression yields, and refolding capacity, as well as better *in vivo* tissue penetration. Moreover, in contrast to the VH domains of conventional antibodies, VHH do not display an intrinsic tendency to bind to light chains. This facilitates the induction of heavy chain antibodies in the presence of a functional light chain loci. Further, since VHH do not bind to VL domains, it is much easier to reformat VHHs into bispecific antibody constructs than constructs containing conventional VH-VL pairs or single domains based on VH domains.

A notable difference between the camelid VHH and the human VH domain is the length and orientation of the CDR3 loop. The CDR3 corresponds to the unique region of the antibody molecule that is encoded by a DNA element newly generated during B-cell development. Genetic recombination results in the fusion of a D-element with flanking V- and J-elements. During recombination further genetic diversity is generated by addition and/or deletion of nucleotides at the junctions. Thereby, the CDR3 loop provides the major contribution to antibody diversity and specificity. A limited number of variable region genes (IGHV, IGHD, and IGHJ) in some early transgenic heavy chain antibody animals, results in some antigens not being

recognized by these animals, despite potent antigen response by wildtype animals (Janssens, Rick, et al. "Generation of heavy-chain-only antibodies in mice." Proceedings of the National Academy of Sciences 103.41 (2006): 15130-15135). The present disclosure provides genetically modified animals that have complete human heavy chain antibody repertoires. Thus, the variable domains generated by these animals can have the maximal possible diversity for human heavy chain variable domains, thus maximizing the chance to obtain a fully humanized heavy chain antibody.

Furthermore, because the entire sequence at the human immunoglobulin locus is introduced into the animal genome (with no modifications or limited modifications), these genes can undergo the V(D)J recombination in a way that is very similar to what happens in human, reducing the risk of the generating new immunogenic epitopes that can be recognized in a human immune system, thereby decreasing immunogenicity. The immunogenicity can lead to production of anti-drug-antibodies and may comprise efficacy. Here, the endogenous IGHV, IGHD, and IGHJ genes have been effectively deleted. It is less likely that the antibodies generated by the antibody repertoires are immunogenic in humans. In addition, the antibody production can be very efficient and has a production rate that is similar to the normal production rates due to the efficient V(D)J recombination. Thus, the antibodies are more suitable as therapeutics in humans. Therefore, the genetically modified animals provide an advantageous platform to produce humanized heavy chain antibodies.

In addition, IgG1 is the most abundant antibody subtype in serum, with long serum half-life, strong FcγR affinity, antibody-dependent cellular cytotoxicity (ADCC), complement dependent cytotoxicity (CDC) activity, etc. IgG1 has unique advantages in the field of antibody drug development. Thus, in one aspect, the present disclosure is particularly related to preparation of humanized mice that can produce heavy-chain antibodies of IgG1 subtype. In the meantime, coding sequences for all other IgG subtypes can be deleted. This creates an efficient and reliable platform to create heavy chain antibodies in the animals.

As used herein, the term "antibody" refers to any antigen-binding molecule that contains at least one (e.g., one, two, three, four, five, or six) complementary determining region (CDR) (e.g., any of the three CDRs from an immunoglobulin light chain or any of the three CDRs from an immunoglobulin heavy chain) and is capable of specifically binding to an epitope. Non-limiting examples of antibodies include: monoclonal antibodies, polyclonal antibodies, multi-specific antibodies (e.g., bi-specific antibodies), single-chain antibodies, heavy chain antibodies, chimeric antibodies, human antibodies, and humanized antibodies. In some embodiments, an antibody can contain an Fc region of a human antibody. The term antibody also includes derivatives, e.g., bi-specific antibodies, single-chain antibodies, diabodies, linear antibodies, and multi-specific antibodies formed from antibody fragments.

As used herein, the term "antigen-binding fragment" refers to a portion of a full-length antibody, wherein the portion is capable of specifically binding to an antigen. In some embodiments, the antigen-binding fragment contains at least one variable domain (e.g., a variable domain of a heavy chain or a variable domain of light chain). Non-limiting examples of antibody fragments include, e.g., Fab, Fab', F(ab')2, and Fv fragments.

As used herein, the term "human antibody" refers to an antibody that is encoded by a nucleic acid (e.g., rearranged human immunoglobulin heavy or light chain locus) present in a

human. In some embodiments, a human antibody is collected from a human or produced in a human cell culture (e.g., human hybridoma cells). In some embodiments, a human antibody is produced in a non-human cell (e.g., a mouse or hamster cell line). In some embodiments, a human antibody is produced in a bacterial or yeast cell. In some embodiments, a human antibody is produced in a transgenic non-human animal (e.g., a mouse) containing an unrearranged or rearranged human immunoglobulin locus (e.g., heavy or light chain human immunoglobulin locus).

As used herein, the term "chimeric antibody" refers to an antibody that contains a sequence present in at least two different antibodies (e.g., antibodies from two different mammalian species such as a human and a mouse antibody). A non-limiting example of a chimeric antibody is an antibody containing the variable domain sequences (e.g., all or part of a light chain and/or heavy chain variable domain sequence) of a human antibody and the constant domains of a non-human antibody. Additional examples of chimeric antibodies are described herein and are known in the art.

As used herein, the term "humanized antibody" refers to a non-human antibody which contains sequence derived from a non-human (e.g., mouse) immunoglobulin and contains sequences derived from a human immunoglobulin.

As used herein, the term "single-chain antibody" refers to a single polypeptide that contains at least two immunoglobulin variable domains (e.g., a variable domain of a mammalian immunoglobulin heavy chain or light chain) that is capable of specifically binding to an antigen.

As used herein, the term "heavy-chain antibody" refers to an antibody molecule which is composed only of heavy chains (generally two) and does not have any light chains.

As used herein, the term "VHH" refers to the variable domain derived from a heavy-chain antibody. The VHH can specifically recognize an antigen without the need to be paired with a VL. In some embodiments, the VHH (also know as sdAb or nanobody) described herein is derived from any of the humanized heavy-chain antibody described herein. In some embodiments, the VHH, sdAb, or nanobody described herein is derived from the heavy chain antibody produced by any of the genetically modified non-human animal described herein.

As used herein, the terms "subject" and "patient" are used interchangeably throughout the specification and describe an animal, human or non-human. Veterinary and non-veterinary applications are contemplated by the present disclosure. Human patients can be adult humans or juvenile humans (e.g., humans below the age of 18 years old). In addition to humans, patients include but are not limited to mice, rats, hamsters, guinea-pigs, rabbits, ferrets, cats, dogs, and primates. Included are, for example, non-human primates (e.g., monkey, chimpanzee, gorilla, and the like), rodents (e.g., rats, mice, gerbils, hamsters, ferrets, rabbits), lagomorphs, swine (e.g., pig, miniature pig), equine, canine, feline, bovine, and other domestic, farm, and zoo animals.

As used herein, when referring to an antibody, the phrases "specifically binding" and "specifically binds" mean that the antibody interacts with its target molecule preferably to other molecules, because the interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) on the target molecule; in other words, the reagent is recognizing and binding to molecules that include a specific structure rather than to all molecules in general. An antibody that specifically binds to the target molecule may be referred to as a target-specific antibody.

As used herein, the terms "polypeptide," "peptide," and "protein" are used interchangeably to refer to polymers of amino acids of any length of at least two amino acids.

As used herein, the terms "polynucleotide," "nucleic acid molecule," and "nucleic acid sequence" are used interchangeably herein to refer to polymers of nucleotides of any length of at least two nucleotides, and include, without limitation, DNA, RNA, DNA/RNA hybrids, and modifications thereof.

As used herein, the term "an unmodified human sequence" refers to a sequence that is derived from a human subject, a human cell, a cultured human cell or a human cell line, wherein the sequence is identical to the genetic sequence of a human subject, a human cell, a cultured human cell or a human cell line.

Immunoglobulin heavy chain constant region locus

Heavy chain immunoglobulin locus (also known as IGH or immunoglobulin heavy locus) is a region on the chromosome (e.g., mouse chromosome 12) that contains genes for the heavy chains of antibodies (or immunoglobulins). It includes sequences encoding the heavy chain variable region and heavy chain constant region.

As shown in **FIG. 2**, the mouse immunoglobulin heavy chain constant region genes include (as shown in the following order): immunoglobulin heavy constant mu (IGHM, or Cμ), immunoglobulin heavy constant delta (IGHδ, or Cδ), immunoglobulin heavy constant gamma 3 (IGHG3, or Cγ3), immunoglobulin heavy constant gamma 1 (IGHG1, or Cγ1), immunoglobulin heavy constant gamma 2b (IGHG2b, or Cγ2b), immunoglobulin heavy constant gamma 2a (IGHG2c, or Cγ2c), immunoglobulin heavy constant epsilon (IGHE, or Cε), and immunoglobulin heavy constant alpha (IGHA, or Cα) genes. In some embodiments, immunoglobulin heavy constant gamma 2a (IGHG2a) is at the position of IGHG2c. In contrast, human immunoglobulin constant region genes include (as shown in the following order): immunoglobulin heavy constant mu (IGHM, or Cμ), immunoglobulin heavy constant delta (IGHδ, Cδ), immunoglobulin heavy constant gamma 3 (IGHG3, or Cγ3), immunoglobulin heavy constant gamma 1 (IGHG1, or Cγ1), immunoglobulin heavy constant epsilon P1 (pseudogene) (IGHEP1, or ψCε), immunoglobulin heavy constant alpha 1 (IGHA1, or Cα1), immunoglobulin heavy constant gamma P (non-functional) (IGHGP, or CγP; not shown), immunoglobulin heavy constant gamma 2 (IGHG2, or Cγ2), immunoglobulin heavy constant gamma 4 (IGHG4, Cγ4), immunoglobulin heavy constant epsilon (IGHE, or Cε), and immunoglobulin heavy constant alpha 2 (IGHA2, or Cα2) genes.

Immunoglobulin class switching (or isotype switching, or isotypic commutation, or class switch recombination (CSR)) is a biological mechanism that changes a B cell's production of antibody from one class to another; for example, from isotype IgM to isotype IgG. During this process, the constant region portion of the antibody-heavy chain is changed, but the variable region of the heavy chain stays the same. Since the variable region does not change, class switching does not affect antigen specificity. Instead, the antibody retains affinity for the same antigens, but can interact with different effector molecules. This allows different daughter cells from the same activated B cell to produce antibodies of different isotypes or subtypes (e.g. IgG1, IgG2 etc.). Class switching occurs by a mechanism called class switch recombination (CSR) binding. Class switch recombination is a biological mechanism that allows the class of antibody

produced by an activated B cell to change during a process known as isotype or class switching. During CSR, portions of the antibody-heavy chain locus are removed from the chromosome, and the gene segments surrounding the deleted portion are rejoined to retain a functional antibody gene that produces antibody of a different isotype. Double-stranded breaks are generated in DNA at conserved nucleotide motifs, called switch (S) regions, which are upstream from gene segments that encode the constant regions of antibody heavy chains; these occur adjacent to all heavy chain constant region genes with the exception of Cδ. DNA is nicked and broken at two selected switch regions by the activity of a series of enzymes, including Activation-Induced (Cytidine) Deaminase (AID), uracil DNA glycosylase and apyrimidinic/apurinic (AP)-endonucleases. The intervening DNA between the switch regions is subsequently deleted from the chromosome, e.g., removing unwanted Cμ or Cδ heavy chain constant region sequences and allowing substitution of Cγ, Cα, or Cε constant region gene segment. The free ends of the DNA are rejoined by a process called non-homologous end joining (NHEJ) to link the variable domain exon to the desired downstream constant domain exon of the antibody heavy chain. In the absence of non-homologous end joining, free ends of DNA can be rejoined by an alternative pathway biased toward microhomology joins. With the exception of the Cμ and Cδ genes, only one antibody class is expressed by a B cell at any point in time. **FIG. 3** shows the location of each switch region (e.g., Sμ, Sγ3, Sγ1, Sγ2b, Sγ2c, Sε, and Sα) in the mouse immunoglobulin heavy chain constant region locus. Specifically, the Cγ1 gene comprises exons encoding the CH1 domain, the hinge region, the CH2 domain, the CH3 domain, and two transmembrane domains. The exons are labelled with CH1, H, CH2, CH3, M1, and M2, respectively.

The five main classes of immunoglobulins are IgM, IgD, IgG, IgE, and IgA, each of which can occur as transmembrane antigen receptors or secreted antibodies. In humans, IgG is found as four subclasses (IgG1, IgG2, IgG3, and IgG4), named by decreasing order of their abundance in serum, and IgA antibodies are found as two subclasses (IgA1 and IgA2). The different heavy chains that define these classes are known as isotypes and are designated by the lowercase Greek letters μ (IgM), δ (IgD), γ (IgG), ε (IgE), and α (IgA), respectively.

Although the position and number of disulfide bonds are different between IgG1-IgG4, structures of the four IgG subtypes are very similar. In fact, as the most abundant IgG subtype in plasma, IgG1 is widely used for making recombinant therapeutic antibodies. In contrast, IgG3 is rarely used for antibody drug development, because it has a weak binding affinity to FcRn and a short half-life (about 9 days). Therefore, IgG3-based antibody drugs have to be administered more frequently for pharmacokinetics reasons. In addition, the antibody levels of different subtypes can change during body development, which is described e.g., in Elena Blanco et al., "Age-associated distribution of normal B-cell and plasma cell subsets in peripheral blood," Journal of Allergy and Clinical Immunology, Volume 141, Issue 6 (2018), which is incorporated herein by reference in its entirety. IgG1 is the most abundant antibody subtype in serum, with long serum half-life, strong FcγR affinity, antibody-dependent cellular cytotoxicity (ADCC), complement dependent cytotoxicity (CDC) activity, etc. IgG1 has unique advantages in the field of antibody drug development. Thus, in one aspect, the disclosure is related to preparation of humanized mice that can produce heavy-chain antibodies of IgG1 subtype. In one aspect, the present disclosure is particularly related to preparation of humanized mice that can produce

heavy-chain antibodies of IgG1 subtype. In some embodiments, the heavy-chain antibodies can be further processed to generate nanobodies.

In one aspect, the disclosure relates to a genetically modified non-human animal comprising a modified immunoglobulin heavy chain locus, wherein the modified immunoglobulin heavy chain locus comprises an IgG constant region gene, wherein the IgG constant region gene encodes an IgG heavy chain constant region lacking a CH1 domain, wherein the genetically modified non-human animal expresses a heavy-chain antibody.

In some embodiments, the IgG constant region gene is one of the following: IGHG3, IGHG1, IGHG2a, IGHG2b, and IGHG2c. In some embodiments, the modified immunoglobulin heavy chain locus has one, two, three, four, or five IgG constant region genes, wherein none of them encodes a CH1 domain. In some embodiments, only one, two, or three IGHG genes do not encode a CH1 domain, while at least one, two, or three remaining IGHG genes can encode a CH1 domain. In some embodiments, the modified immunoglobulin heavy chain locus has only one (e.g., exactly one) IgG constant region gene, e.g., IGHG1. In some embodiments, the modified immunoglobulin heavy chain locus does not include IGHG3, IGHG2a, IGHG2b, and/or IGHG2c genes.

In some embodiments, the modified immunoglobulin heavy chain locus has only one IGHG gene (e.g., IGHG3, IGHG1, IGHG2a, IGHG2b, or IGHG2c), and the sequence encoding the CH1 domain in the IGHG is deleted. In some embodiments, the IGHG gene is operably linked to Sγ3, Sγ1, Sγ2a, Sγ2b, or Sγ2c (e.g., Sγ3 or Sγ1).

In some embodiments, the IGHM is an intact functional endogenous IGHM gene. In some embodiments, the sequence encoding the CH1 domain in the IGHM is deleted. In some embodiments, IGHM is deleted. In some embodiments, the IGHδ is an intact functional endogenous IGHδ gene. In some embodiments, the sequence encoding the CH1 domain in the IGHδ is deleted. In some embodiments, IGHδ is deleted. In some embodiments, both IGHM and IGHδ are deleted.

In some embodiments, the IGHE is an intact functional endogenous IGHE gene. In some embodiments, the sequence encoding the CH1 domain in the IGHE is deleted. In some embodiments, IGHE is deleted.

In some embodiments, the IGHA is an intact functional endogenous IGHA gene. In some embodiments, the sequence encoding the CH1 domain in the IGHA is deleted. In some embodiments, IGHA is deleted.

In various embodiments, the humanized heavy chain antibody comprises a unique immunoglobulin constant region (Fc), that lacks at least the CH1 domain. In some embodiment, it also lacks the hinge region of a human Fc. In some embodiment, the heavy-chain antibody comprises the CH2 and CH3 regions of an immunoglobulin G (IgG) heavy chain constant region. In some embodiments, the constant region of the heavy chain antibody includes the hinge, CH2 and CH3 regions of the IgG heavy chain Fc.

In some embodiments, a suitable number of rearranged heavy chain variable regions that can effectively survive selection when presented during B cell development is needed. In one aspect, the disclosure provides a transgenic animal comprising a germline genetic modification that comprises a deletion of a nucleotide sequence encoding a CH1 domain of an IgG, wherein the animal expresses a functional IgM and the animal expresses in its serum an IgG heavy chain

antibody (e.g., with IgG1 heavy chain CH2 and CH3 domains). In some embodiments, the IgM comprises two heavy chains, and they associate with two lambda or kappa lights chains to recognize an antigen. In some embodiments, the functional IgM includes a CH1 domain. In some embodiments, the functional IgM does not include a CH1 domain. While not intending to be bound by any theory, it is believed that the deleting a sequence encoding a CH1 domain from endogenous IGHM gene does not substantially change the IgM function.

In some embodiments, the modified immunoglobulin heavy chain constant region locus includes a modified IGHδ gene lacking a sequence encoding a CH1 domain. In some embodiments, the modified IGHδ expresses a functional IgD. While not intending to be bound by any theory, it is believed that the deleting a sequence encoding a CH1 domain from endogenous IGHδ gene does not substantially change the IgD function.

Immunoglobulin heavy chain variable region locus

To produce a humanized heavy chain antibody (HCAb), the heavy chain immunoglobulin variable region locus in the animal can be humanized. This heavy chain immunoglobulin variable region represents the germline organization of the heavy chain locus. The locus includes V (variable), D (diversity), J (joining), and C (constant) segments. The genes in the V region form a V gene cluster (also known as IGHV gene cluster). The genes in the D region form a D gene cluster (also known as IGHD gene cluster). The genes in the J region form a J gene cluster (also known as IGHJ gene cluster).

During B cell development, a recombination event at the DNA level joins a single D segment (also known as an IGHD gene) with a J segment (also known as an IGHJ gene); the fused D-J exon of this partially rearranged D-J region is then joined to a V segment (also known as an IGHV gene). The rearranged V-D-J region containing a fused V-D-J exon is then transcribed and fused at the RNA level to the IGHM constant region; this transcript encodes a mu heavy chain. Later in development, B cells generate V-D-J-Cμ-Cδ pre-messenger RNA, which is alternatively spliced to encode either a mu or a delta heavy chain. Mature B cells in the lymph nodes undergo switch recombination, so that the fused V-D-J gene segment is brought in proximity to one of the IGHG, IGHA, or IGHE gene segments and each cell expresses either the gamma, alpha, or epsilon heavy chain. Potential recombination of many different IGHV genes with several IGHJ genes provides a wide range of antigen recognition. Additional diversity is attained by junctional diversity, resulting from the random addition of nucleotides by terminal deoxynucleotidyl transferase, and by somatic hypermutation, which occurs during B cell maturation in the spleen and lymph nodes. Several V, D, J, and C segments are known to be incapable of encoding a protein and are considered pseudogenous gene segments (often simply referred to as pseudogenes).

The human heavy chain immunoglobulin locus is located on human chromosome 14 (**FIG. 27** and **FIG. 29**). Table 1 lists IGHV genes and its relative orders in this locus.

Table 1. List of IGHV genes on human chromosome 14

Gene names	Order	Gene names	Order	Gene names	Order	Gene names	Order
IGHV(III)-82	1	IGHV3-57	42	IGHV3-36	83	IGHV3-23	124
IGHV7-81	2	IGHV7-56	43	IGHV3-35	84	IGHV(III)-22-2	125
IGHV4-80	3	IGHV4-55	44	IGHV7-34-1	85	IGHV(II)-22-1	126

IGHV3-79	4	IGHV3-54	45	IGHV4-34	86	IGHV3-22	127
IGHV(II)-78-1	5	IGHV(II)-53-1	46	IGHV3-33-2	87	IGHV3-21	128
IGHV5-78	6	IGHV3-53	47	IGHV(II)-33-1	88	IGHV(II)-20-1	129
IGHV7-77	7	IGHV3-52	48	IGHV3-33	89	IGHV3-20	130
IGHV(III)-76-1	8	IGHV(II)-51-2	49	GOLGA4P1 (Golgin)	90	IGHV3-19	131
IGHV3-76	9	IGHV(III)-51-1	50	IGHV3-32	91	IGHV1-18	132
IGHV3-75	10	IGHV5-51	51	IGHV(II)-31-1	92	SLC20A1P2	133
IGHV(II)-74-1	11	IGHV3-50	52	IGHV4-31	93	IGHV1-17	134
IGHV3-74	12	IGHV(II)-49-1	53	IGHV3-30-52	94	IGHV(III)-16-1	135
IGHV3-73	13	IGHV3-49	54	IGHV(II)-30-51	95	IGHV3-16	136
IGHV3-72	14	IGHV3-48	55	IGHV3-30-5	96	IGHV(II)-15-1	137
IGHV3-71	15	IGHV(III)-47-1	56	IGHV3-30-42	97	IGHV3-15	138
IGHV2-70	16	IGHV3-47	57	IGHV(II)-30-41	98	IGHV1-14	139
IGHV1-69D	17	IGHV(II)-46-1	58	IGHV4-30-4	99	IGHV(III)-13-1	140
IGHV1-69-2	18	IGHV1-46	59	IGHV3-30-33	100	IGHV3-13	141
IGHV3-69-1	19	IGHV1-45	60	IGHV(II)-30-32	101	IGHV1-12	142
IGHV2-70D	20	IGHV(II)-44-2	61	IGHV3-30-3	102	IGHV(III)-11-1	143
IGHV1-69	21	IGHV(IV)-44-1	62	IGHV3-30-22	103	IGHV3-11	144
IGHV1-68	22	IGHV(III)-44	63	IGHV(II)-30-21	104	IGHV2-10	145
IGHV(III)-67-4	23	IGHV(II)-43-1	64	IGHV4-30-2	105	IGHV3-9	146
IGHV(III)-67-3	24	IGHV3-43	65	IGHV4-30-1	106	IGHV1-8	147
IGHV(III)-67-2	25	IGHV3-42	66	IGHV3-30-2	107	IGHV5-10-1	148
IGHV(II)-67-1	26	IGHV3-41	67	IGHV(II)-30-1	108	IGHV3-64D	149
SLC20A1P1 (GLVR1)	27	IGHV(II)-40-1	68	IGHV3-30	109	IGHV3-7	150
IGHV1-67	28	IGHV7-40	69	GOLGA4P2 (Golgin)	110	IGHV3-6	151
IGHV3-66	29	IGHV4-39	70	IGHV3-29	111	IGHV(III)-5-2	152
IGHV(II)-65-1	30	IGHV1-38-4	71	IGHV(II)-28-1	112	IGHV(III)-5-1	153
IGHV3-65	31	IGHV(III)-38-1D	72	IGHV4-28	113	IGHV2-5	154
IGHV3-64	32	IGHV3-38-3	73	IGHV7-27	114	IGHV7-4-1	155
GOLGA4P3 (Golgin)	33	IGHV(III)-44D	74	IGHV(II)-26-2	115	IGHV4-4	156
IGHV3-63	34	IGHV(II)-43-1D	75	IGHV(III)-26-1	116	IGHV1-3	157
IGHV(II)-62-1	35	IGHV3-43D	76	IGHV2-26	117	IGHV(III)-2-1	158
IGHV3-62	36	IGHV3-42D	77	IGHV(III)-25-1	118	IGHV1-2	159
IGHV4-61	37	IGHV7-40D	78	IGHV3-25	119		*
IGHV(II)-60-1	38	IGHV4-38-2	79	IGHV1-24	120		*

IGHV3-60	39	IGHV(III)-38-1	80	IGHV3-23D	121	IGHV(II)-1-1	162
IGHV4-59	40	IGHV3-38	81	IGHV(III)-22-2D	122	IGHV6-1	163
IGHV1-58	41	IGHV3-37	82	IGHV(II)-22-1D	123		*

RPS8P1, ADAM6, and KIAA0125 are also located in this locus. The relative order of RPS8P1 is 160, the relative order of ADAM6 is161, and the relative order of KIAA0125 is 164. Table 2 lists all IGHD genes and its relative orders on human chromosome 14. Table 3 lists all IGHJ genes and its relative orders on human chromosome 14. The genes for immunoglobulin constant domains are located after the IGHV, IGHD, and IGHJ genes. These genes include (as shown in the following order): immunoglobulin heavy constant mu (IGHM), immunoglobulin heavy constant delta (IGH δ), immunoglobulin heavy constant gamma 3 (IGHG3), immunoglobulin heavy constant gamma 1 (IGHG1), immunoglobulin heavy constant epsilon P1 (pseudogene) (IGHEP1), immunoglobulin heavy constant alpha 1 (IGHA1), immunoglobulin heavy constant gamma P (non-functional) (IGHGP), immunoglobulin heavy constant gamma 2 (IGHG2), immunoglobulin heavy constant gamma 4 (IGHG4), immunoglobulin heavy constant epsilon (IGHE), and immunoglobulin heavy constant alpha 2 (IGHA2). These genes and the order of these genes are also shown in **FIG. 27** and **FIG. 29**.

Table 2. List of IGHD genes on human chromosome 14

Gene names	Order	Gene names	Order	Gene names	Order	Gene names	Order
IGHD1-1	165	IGHD2-8	172	IGHD2-15	179	IGHD3-22	186
IGHD2-2	166	IGHD3-9	173	IGHD3-16	180	IGHD4-23	187
IGHD3-3	167	IGHD3-10	174	IGHD4-17	181	IGHD5-24	188
IGHD4-4	168	IGHD4-11	175	IGHD5-18	182	IGHD6-25	189
IGHD5-5	169	IGHD5-12	176	IGHD6-19	183	IGHD1-26	190
IGHD6-6	170	IGHD6-13	177	IGHD1-20	184		*
IGHD1-7	171	IGHD1-14	178	IGHD2-21	185	IGHD7-27	192

Table 3. List of IGHJ genes on human chromosome 14

Gene names	Order	Gene names	Order
IGHJ1P	191	IGHJ4	197
IGHJ1	193	IGHJ5	198
IGHJ2	194	IGHJ3P	199
IGHJ2P	195	IGHJ6	200
IGHJ3	196		

The mouse heavy chain immunoglobulin locus is located on mouse chromosome 12 (**FIG. 28** and **FIG. 30**). Table 4 lists IGHV genes and its relative orders in this locus.

Table 4. List of IGHV genes on mouse chromosome 12

Gene names	Order	Gene names	Order	Gene names	Order	Gene names	Order
IGHV1-86	1	IGHV1-51	50	IGHV15-2	99	IGHV4-1	148

IGHV1-85	2	IGHV1-50	51	IGHV1-8	100	IGHV14-1	149
IGHV1-84	3	IGHV8-5	52	IGHV10-4	101	IGHV7-2	150
IGHV1-83	4	IGHV1-49	53	IGHV1-7	102	IGHV7-1	151
IGHV1-82	5	IGHV1-48	54	IGHV1-6	103	IGHV5-19	152
IGHV1-81	6	IGHV8-4	55	IGHV10-3	104	IGHV2-9	153
IGHV1-80	7	IGHV8-3	56	IGHV1-5	105	IGHV2-8	154
IGHV1-79	8	IGHV1-47	57	IGHV10-2	106	IGHV5-18	155
IGHV1-78	9	IGHV1-46	58	IGHV1-4	107	IGHV5-17	156
IGHV1-77	10	IGHV1-45	59	IGHV1-3	108	IGHV5-16	157
IGHV8-16	11	IGHV1-44	60	IGHV10-1	109	IGHV5-15	158
IGHV1-76	12	IGHV1-43	61	IGHV1-2	110	IGHV2-7	159
IGHV8-15	13	IGHV1-42	62	IGHV8-2	111	IGHV2-6-8	160
IGHV1-75	14	IGHV1-41	63	IGHV6-7	112	IGHV2-9-1	161
IGHV8-14	15	IGHV1-40	64	IGHV6-6	113	IGHV5-12-4	162
IGHV1-74	16	IGHV1-39	65	IGHV6-5	114	IGHV5-9-1	163
IGHV1-73	17	IGHV1-38	66	IGHV6-4	115	IGHV2-6	164
IGHV8-13	18	IGHV1-37	67	IGHV6-3	116	IGHV5-12	165
IGHV1-72	19	IGHV1-36	68	IGHV12-3	117	IGHV5-11	166
IGHV1-71	20	IGHV1-35	69	IGHV13-2	118	IGHV2-5	167
IGHV1-70	21	IGHV1-34	70	IGHV1-1	119	IGHV5-10	168
IGHV8-12	22	IGHV1-33	71	IGHV8-1	120	IGHV5-9	169
IGHV1-69	23	IGHV1-32	72	IGHV3-8	121	IGHV5-8	170
IGHV1-68	24	IGHV1-31	73	IGHV5-21	122	IGHV2-4	171
IGHV1-67	25	IGHV1-30	74	IGHV3-7	123	IGHV5-7	172
IGHV1-66	26	IGHV1-29	75	IGHV9-4	124	IGHV5-6	173
IGHV8-11	27	IGHV1-28	76	IGHV3-6	125	IGHV5-5	174
IGHV1-65	28	IGHV1-27	77	IGHV13-1	126	IGHV2-3	175
IGHV8-10	29	IGHV1-26	78	IGHV3-5	127	IGHV6-1	176
IGHV1-64	30	IGHV1-25	79	IGHV3-4	128	IGHV5-4	177
IGHV1-63	31	IGHV1-24	80	IGHV7-4	129	IGHV5-3	178
IGHV8-9	32	IGHV1-23	81	IGHV3-3	130	IGHV2-2	179
IGHV1-62-3	33	IGHV1-22	82	IGHV14-4	131	IGHV5-2	180
IGHV1-62-2	34	IGHV1-21	83	IGHV15-1	132	IGHV2-1	181
IGHV1-62-1	35	IGHV1-21-1	84	IGHV7-3	133	IGHV5-1	182
IGHV1-62	36	IGHV1-20	85	IGHV9-3	134		
IGHV1-61	37	IGHV1-19	86	IGHV12-2	135		
IGHV1-60	38	IGHV1-19-1	87	IGHV9-2	136		
IGHV1-59	39	IGHV1-18	88	IGHV12-1	137		
IGHV1-58	40	IGHV1-17	89	IGHV9-1	138		
IGHV8-8	41	IGHV1-17-1	90	IGHV6-2	139		
IGHV1-57	42	IGHV1-16	91	IGHV16-1	140		

IGHV8-7	43	IGHV1-15	92	IGHV14-3	141		
IGHV1-56	44	IGHV1-14	93	IGHV11-2	142		
IGHV1-55	45	IGHV1-13	94	IGHV3-2	143		
IGHV1-54	46	IGHV1-12	95	IGHV4-2	144		
IGHV8-6	47	IGHV1-11	96	IGHV14-2	145		
IGHV1-53	48	IGHV1-10	97	IGHV11-1	146		
IGHV1-52	49	IGHV1-9	98	IGHV3-1	147		

Table 5 lists all IGHD genes and its relative orders on mouse chromosome 12. Table 6 lists all IGHJ genes and its relative orders on mouse chromosome 12. The genes for immunoglobulin constant domains are after the IGHV, IGHD, and IGHJ genes. These genes include (as shown in the following order): immunoglobulin heavy constant mu (IGHM), immunoglobulin heavy constant delta (IGHδ), immunoglobulin heavy constant gamma 3 (IGHG3), immunoglobulin heavy constant gamma 1 (IGHG1), immunoglobulin heavy constant gamma 2b (IGHG2b), immunoglobulin heavy constant gamma 2c (IGHG2c), immunoglobulin heavy constant epsilon (IGHE), and immunoglobulin heavy constant alpha (IGHA) genes. In some embodiments, immunoglobulin heavy constant gamma 2a (IGHG2a) is at the position of IGHG2c. These genes and the order of these genes are also shown in **FIG. 28** and **FIG. 30**.

Table 5. List of IGHD genes on mouse chromosome 12

Gene names	**Order**	**Gene names**	**Order**	**Gene names**	**Order**
IGHD5-1	183	IGHD2-5	191	IGHD5-5	198
IGHD3-1	184	IGHD5-3	192	IGHD2-8	199
IGHD1-1	185	IGHD5-7	193	IGHD5-6	200
IGHD6-1	186	IGHD2-6	194	IGHD3-2	201
IGHD2-3	187	IGHD5-4	195	IGHD4-1	202
IGHD6-2	188	IGHD5-8	196		
IGHD2-4	189	IGHD2-7	197		

Table 6. List of IGHJ genes on mouse chromosome 12

Gene names	**Order**	**Gene names**	**Order**
IGHJ1	203	IGHJ3	205
IGHJ2	204	IGHJ4	206

The present disclosure provides genetically-modified, non-human animal comprising one or more human IGHV genes, one or more human IGHD genes, and/or one or more human IGHJ genes. In some embodiments, the human IGHV genes, the human IGHD genes, and the human IGHJ genes are operably linked together and can undergo VDJ rearrangement. In some embodiments, the human IGHV genes, the human IGHD genes, and the human IGHJ genes are at the endogenous heavy chain immunoglobulin gene locus.

In some embodiments, the animal compromises about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160 or 161 human IGHV genes (e.g., genes as shown in Table 1).

In some embodiments, the animal comprises about or at least 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160 or 161 human IGHV genes selected from Table 1, about or at least 20, 21, 22, 23, 24, 25, 26, or 27 human IGHD genes selected from Table 2, and about or at least 5, 6, 7, 8, or 9 human IGHJ genes selected from Table 3. In some embodiments, the animal comprises all human IGHV genes in Table 1 except IGHV2-10, IGHV3-9, and IGHV1-8, all human IGHD genes in Table 2, and all human IGHJ genes in Table 3. In some embodiments, the animal comprises all human IGHV genes in Table 1 except IGHV5-10-1 and IGHV3-64D, all human IGHD genes in Table 2, and all human IGHJ genes in Table 3. In some embodiments, the animal comprises all human IGHV genes, all human IGHD genes, and all human IGHJ genes at the endogenous heavy chain immunoglobulin gene locus of human chromosome 14 of a human subject. In some embodiments, the animal comprises all human IGHV genes, all human IGHD genes, and all human IGHJ genes at the endogenous heavy chain immunoglobulin gene locus of human chromosome 14 of a human cell (e.g., a somatic cell, a cultured cell, a non-immune cell, a cell without any V(D)J rearrangement).

In some embodiments, the animal compromises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 genes selected from IGHV(III)-82, IGHV7-81, IGHV4-80, IGHV3-79, IGHV(II)-78-1, IGHV5-78, IGHV7-77, IGHV(III)-76-1, IGHV3-76, and IGHV3-75.

In some embodiments, the animal compromises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 genes selected from IGHV(III)-5-2, IGHV(III)-5-1, IGHV2-5, IGHV7-4-1, IGHV4-4, IGHV1-3, IGHV(III)-2-1, IGHV1-2, IGHV(II)-1-1, and IGHV6-1.

In some embodiments, the animal compromises an unmodified human sequence comprising a sequence starting from a gene selected from IGHV(III)-82, IGHV7-81, IGHV4-80, IGHV3-79, IGHV(II)-78-1, IGHV5-78, IGHV7-77 , IGHV(III)-76-1, IGHV3-76, and IGHV3-75, and ending at a gene selected from IGHV(III)-5-2, IGHV(III)-5-1, IGHV2-5, IGHV7-4-1, IGHV4-4, IGHV1-3, IGHV(III)-2-1, IGHV1-2, IGHV(II)-1-1, and IGHV6-1. In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHV1-2. In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHV(II)-1-1. In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHV-6-1.

In some embodiments, the animal compromises about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, or 27 human IGHD genes (e.g., genes as shown in Table 2). In some embodiments, the animal compromises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 genes selected from IGHD1-1, IGHD2-2, IGHD3-3, IGHD4-4, IGHD5-5, IGHD4-23, IGHD5-24, IGHD6-25, IGHD1-26, and IGHD7-27.

In some embodiments, the animal compromises about or at least 1, 2, 3, 4, 5, 6, 7, 8, or 9 human IGHJ genes (e.g., genes as shown in Table 3). In some embodiments, the animal compromises 1, 2, 3, 4, 5, 6, 7, 8, or 9 human IGHJ genes selected from IGHJ1P, IGHJ1, IGHJ2, IGHJ2P, IGHJ3, IGHJ4, IGHJ5, IGHJ3P, and IGHJ6.

In some embodiments, the animal compromises an unmodified human sequence comprising a sequence starting from a gene selected from IGHD1-1, IGHD2-2, IGHD3-3, IGHD4-4, IGHD5-5, IGHD4-23, IGHD5-24, IGHD6-25, IGHD1-26, and IGHD7-27, and ending

at a gene selected from IGHJ1P, IGHJ1, IGHJ2, IGHJ2P, IGHJ3, IGHJ4, IGHJ5, IGHJ3P, and IGHJ6. In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHD1-1 to human IGHJ6.

In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHD1-1 to human IGHD7-27.

In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHJ1P to human IGHJ6. In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHJ1 to human IGHJ6.

In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHJ6.

In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV1-2 to human IGHJ6. In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(II)-1-1 to human IGHJ6. In some embodiments, the unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV6-1 to human IGHJ6.

In some embodiments, the animal can have one, two, three, four, five, six, seven, eight, nine, or ten unmodified human sequences. In some embodiments, the unmodified human sequence has a length of about or at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 kb.

In some aspects, the disclosure relates to a genetically-modified animal comprising at an endogenous heavy chain immunoglobulin gene locus, a first sequence comprising one or more human IGHV genes; a second sequence comprising an endogenous sequence; and a third sequence comprising one or more human IGHD genes, and one or more human IGHJ genes, wherein the first sequence, the second sequence, and the third sequence are operably linked.

In some embodiments, the first sequence comprises about or at least 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160 or 161 human IGHV genes selected from Table 1. In some embodiments, the first sequence comprises about or at least 20, 21, 22, 23, 24, 25, 26, or 27 human IGHD genes selected from Table 2.

In some embodiments, the first sequence is an unmodified sequence derived from a human heavy chain immunoglobulin gene locus. In some embodiments, the first sequence is about or at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 kb.

In some embodiments, the second sequence comprises an endogenous sequence that is about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 kb.

In some embodiments, the third sequence comprises about or at least 20, 21, 22, 23, 24, 25, 26, or 27 human IGHD genes selected from Table 2. In some embodiments, the third sequence comprises about or at least 5, 6, 7, 8, or 9 human IGHJ genes selected from Table 3. In some embodiments, the third sequence comprises all human IGHD genes in Table 2, and all human IGHJ genes in Table 3.

In some embodiments, the animal comprises one or more endogenous genes selected from the group consisting of immunoglobulin heavy constant mu (IGHM), immunoglobulin

heavy constant delta (IGHδ), immunoglobulin heavy constant gamma 3 (IGHG3), immunoglobulin heavy constant gamma 1 (IGHG1), immunoglobulin heavy constant gamma 2b (IGHG2b), immunoglobulin heavy constant gamma 2c (IGHG2c), immunoglobulin heavy constant epsilon (IGHE), and immunoglobulin heavy constant alpha (IGHA) genes. In some embodiments, immunoglobulin heavy constant gamma 2a (IGHG2a) is at the position of IGHG2c. In some embodiments, these endogenous genes are operably linked together. In some embodiments, these endogenous genes have the same order as in a wildtype animal. In some embodiments, isotype switching (immunoglobulin class switching) can occur in the animal.

In some embodiments, the IGHV genes, the IGHD genes, and/or the IGHJ genes are operably linked together. The VDJ recombination can occur among these genes and produce functional antibodies. In some embodiments, these genes are arranged in an order that is similar to the order in human heavy chain immunoglobulin locus. This arrangement offers various advantages, e.g., the arrangement of these genes allow the production of heavy chain variable domains with a diversity that is very similar to the diversity of the heavy chain variable domains in human. As some random sequences may be inserted to the sequence during VDJ recombination, in some embodiments, the complete human antibody repertoires with no or minimum modifications can reduce the likelihood that non-human sequence is inserted during the VDJ recombination.

In some embodiments, the IGHV genes, the IGHD genes, and/or the IGHJ genes are operably linked together to one or more genes (e.g., all genes) selected from IGHM, IGHδ, IGHG3, IGHG1, IGHG2a, IGHG2b, IGHG2c, IGHE, and IGHA genes.

In some embodiments, the animal comprises a disruption in the animal's endogenous heavy chain immunoglobulin gene locus. In some embodiments, the disruption in the animal's endogenous heavy chain immunoglobulin gene locus comprises a deletion of one or more endogenous IGHV genes, one or more endogenous IGHD genes, and one or more endogenous IGHJ genes.

In some embodiments, the animal is a mouse. The disruption in the animal's endogenous heavy chain immunoglobulin gene locus comprises a deletion of at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, or 182 mouse IGHV genes (e.g., genes as shown in Table 4). In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGHV genes selected from IGHV1-86, IGHV1-85, IGHV1-84, IGHV1-83, IGHV1-82, IGHV1-81, IGHV1-80, IGHV1-79, IGHV1-78, and IGHV1-77. In some embodiments, the mouse still compromises about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGHV genes selected from IGHV1-86, IGHV1-85, IGHV1-84, IGHV1-83, IGHV1-82, IGHV1-81, IGHV1-80, IGHV1-79, IGHV1-78, and IGHV1-77 (e.g., IGHV1-86).

In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGHV genes selected from IGHV5-6, IGHV5-5, IGHV2-3, IGHV6-1, IGHV5-4, IGHV5-3, IGHV2-2, IGHV5-2, IGHV2-1, and IGHV5-1. In some embodiments, the mouse still compromises a deletion of about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGHV genes selected from IGHV5-6, IGHV5-5, IGHV2-3, IGHV6-1, IGHV5-4, IGHV5-3, IGHV2-2, IGHV5-2, IGHV2-1, and IGHV5-1.

In some embodiments, the disruption in the animal's endogenous heavy chain immunoglobulin gene locus comprises a deletion of at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 mouse IGHD genes (e.g., genes as shown in Table 5). In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGHD genes selected from IGHD5-1, IGHD3-1, IGHD1-1, IGHD6-1, IGHD2-3, IGHD2-7, IGHD2-8, IGHD5-6, IGHD3-2, and IGHD4-1. In some embodiments, the mouse still compromises about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGHD genes selected from IGHD5-1, IGHD3-1, IGHD1-1, IGHD6-1, IGHD2-3, IGHD2-7, IGHD2-8, IGHD5-6, IGHD3-2, and IGHD4-1.

In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, or 4 mouse IGHJ genes selected from IGHJ1, IGHJ2, IGHJ3, and IGHJ4. In some embodiments, the mouse still compromises about or at least 1, 2, 3, or 4 mouse IGHJ genes selected from IGHJ1, IGHJ2, IGHJ3, and IGHJ4.

In some embodiments, the disruption in the animal's endogenous heavy chain immunoglobulin gene locus comprises a deletion of about or at least 500 kb, 600 kb, 700 kb, 800 kb, 900 kb, 1000 kb, 1500 kb, 2000 kb, 2500 kb, or 3000 kb of an endogenous sequence.

In some embodiments, the deleted sequence starts from IGHV1-86 to IGHJ4, from IGHV1-85 to IGHJ4, from IGHV1-84 to IGHJ4, from IGHV1-83 to IGHJ4, or from IGHV1-82 to IGHJ4 (e.g., from IGHV1-85 to IGHJ4).

In some embodiments, the animal comprises about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 sequences that are at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a sequence in the human heavy chain immunoglobulin gene locus. In some embodiments, the sequence has a length of about or at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000 or 3500 kb. In some embodiments, the sequence starts from human IGHV(III)-82 to IGHV1-2. In some embodiments, the sequence starts from human IGHV7-81 to IGHV1-2. In some embodiments, the sequence starts from human IGHV(II)-1-1 to IGHVJ6. In some embodiments, the sequence starts from human IGHV6-1 to IGHVJ6.

The human IGHV genes, the human IGHD genes, and the human IGHJ genes are operably linked together and can undergo VDJ rearrangement. In some embodiments, the modified mouse has complete human IGHV, IGHD, and IGHJ gene repertoires (e.g., including all non-pseudo human IGHV, IGHD, and IGHJ genes). Thus, the modified mouse can produce a complete human antibody repertory. In some embodiments, after VDJ recombination, one IGHV gene (e.g., IGHV3-21 or IGHV3-74) contributes to the sequence that encodes an antibody heavy chain variable region. One IGHD gene contributes to the sequence that encodes an antibody heavy chain variable region. And one IGHJ gene contributes to the sequence that encodes an antibody heavy chain variable region. In some embodiments, the IGHV gene is IGHV3-21 or IGHV3-74.

In some embodiments, one IGHV gene (e.g., IGHV3-30, IGHV3-33, IGHV4-39, or IGHV4-34) contributes to the sequence that encodes an antibody heavy chain variable region. One IGHD gene (e.g., IGHD6-19) contributes to the sequence that encodes an antibody heavy chain variable region. And one IGHJ gene (e.g., IGHJ4 or IGHJ6) contributes to the sequence that encodes an antibody heavy chain variable region.

Furthermore, in some cases, the entire mouse IGHV genes, IGHD genes, and IGHJ genes (e.g., including all none-pseudo genes) are knocked out, and the heavy chain variable region will not have any sequence that is encoded by a sequence derived from the mouse, thereby minimizing immunogenicity in human.

In some embodiments, the locus can have about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 30 or 40 non-human exogenous IGHV genes (e.g., from camelid). In some embodiments, the locus can have about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 30 or 40 non-human exogenous IGHD genes (e.g., from camelid). In some embodiments, the locus can have about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 30 or 40 non-human exogenous IGHJ genes (e.g., from camelid). These non-human exogenous genes can help improve the diversity of VHH domains.

Various modifications on the heavy chain immunoglobulin locus are described e.g., in WO2020169022A1, which is incorporated herein by reference in its entirety.

Immunoglobulin light chain locus

In some embodiments, the animal has an intact kappa and/or lambda chain immunoglobulin locus. In some embodiments, the animal has a disrupted kappa and/or lambda chain immunoglobulin locus.

Kappa chain immunoglobulin locus (also known as IGK or immunoglobulin kappa locus) is a region on the chromosome (e.g., chromosome 6) that contains genes for the light chains of human antibodies (or immunoglobulins). Similarly, the immunoglobulin light chain genes can also undergo a series rearrangement that lead to the production of a mature immunoglobulin light-chain nucleic acid (e.g., a kappa chain).

The joining of a V segment (also known as an IGKV gene) and a J segment (also known as an IGKJ gene) creates a continuous exon that encodes the whole of the light-chain variable domain. In the unrearranged DNA, the V gene segments (or IGKV gene cluster) are located relatively far away from the C region. The J gene segments (or IGKJ gene cluster) are located close to the C region. Joining of a V segment to a J gene segment also brings the V gene close to a C-region sequence. The J gene segment of the rearranged V region is separated from a C-region sequence only by an intron. To make a complete immunoglobulin light-chain messenger RNA, the V-region exon is joined to the C-region sequence by RNA splicing after transcription.

In some embodiments, the animal comprises a disruption in the animal's endogenous light chain immunoglobulin gene locus. In some embodiments, the disruption in the animal's endogenous light chain immunoglobulin gene locus comprises a deletion of one or more endogenous IGKV genes, and one or more endogenous IGKJ genes.

In some embodiments, the animal is a mouse. The disruption in the animal's endogenous kappa chain immunoglobulin gene locus comprises a deletion of at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, or 163 mouse IGKV genes. In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGKV genes selected from IGKV2-137, IGKV1-136, IGKV1-135, IGKV14-134-1, IGKV17-134, IGKV1-133, IGKV1-132, IGKV1-131, IGKV14-130, and IGKV9-129. In some embodiments, the mouse still compromises about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGKV genes

selected from IGKV2-137, IGKV1-136, IGKV1-135, IGKV14-134-1, IGKV17-134, IGKV1-133, IGKV1-132, IGKV1-131, IGKV14-130, and IGKV9-129.

In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGKV genes selected from IGKV3-10, IGKV3-9, IGKV3-8, IGKV3-7, IGKV3-6, IGKV3-5, IGKV3-4, IGKV3-3, IGKV3-2, and IGKV3-1. In some embodiments, the mouse still compromises about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 mouse IGKV genes selected from IGKV3-10, IGKV3-9, IGKV3-8, IGKV3-7, IGKV3-6, IGKV3-5, IGKV3-4, IGKV3-3, IGKV3-2, and IGKV3-1.

In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, 4, or 5 mouse IGKJ genes selected from IGKJ1, IGKJ2, IGKJ3, IGKJ4, and IGKJ5. In some embodiments, the mouse still compromises about or at least 1, 2, 3, 4, or 5 mouse IGKJ genes selected from IGKJ1, IGKJ2, IGKJ3, IGKJ4, and IGKJ5 (e.g., IGKJ5).

In some embodiments, the disruption in the animal's endogenous kappa light chain immunoglobulin gene locus comprises a deletion of about or at least 500 kb, 600 kb, 700 kb, 800 kb, 900 kb, 1000 kb, 1500 kb, 2000 kb, 2500 kb, 3000 kb or 3500 kb of an endogenous sequence.

In some embodiments, the deleted sequence starts from IGKV2-137 to IGKJ4, from IGKV1-136 to IGKJ4, from IGKV1-135 to IGKJ4, from IGKV2-137 to IGKJ5, from IGKV1-136 to IGKJ5, or from IGKV1-135 to IGKJ5 (e.g., from IGKV2-137 to IGKJ5).

In some embodiments, the animal comprises a disruption in the animal's endogenous lambda light chain immunoglobulin gene locus. In some embodiments, the disruption in the animal's endogenous light chain immunoglobulin gene locus comprises a deletion of one or more endogenous IGLV genes, one or more endogenous IGLJ genes, and/or one or more immunoglobulin lambda constant (IGLC) genes (e.g., IGLC1, IGLC2, IGLC3, and IGLC4).

The disruption in the animal's endogenous lambda light chain immunoglobulin gene locus comprises a deletion of at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 mouse IGLV, IGLJ, and IGLC genes. In some embodiments, the deletion compromises about or at least 1, 2, 3, or 4 mouse IGKC genes selected from IGLC1, IGLC2, IGLC3, and IGLC4. In some embodiments, the disruption compromises a deletion of about or at least 1, 2, or 3 mouse IGLV genes selected from IGLV1, IGLV2, and IGLV3. In some embodiments, the disruption compromises a deletion of about or at least 1, 2, 3, 4, or 5 mouse IGLJ genes selected from IGLJ1, IGLJ2, IGLJ3, IGLJ3P, and IGLJ4.

In some embodiments, the disruption in the animal's endogenous lambda light chain immunoglobulin gene locus comprises a deletion of about or at least 10 kb, 20 kb, 30 kb, 40 kb, 50 kb, 60 kb, 70 kb, 80 kb, 90 kb, 100 kb, 110 kb, 120 kb, 130 kb, 140 kb, 150 kb, 160 kb, 170 kb, 180 kb, 190 kb, 200 kb, 210 kb, 220 kb, 230 kb, 240 kb, 250 kb, 260 kb, 270 kb, 280 kb, 290 kb, 300 kb, 350 kb, 400 kb, 450 kb, 500 kb, or 1000 kb of nucleotides. In some embodiments, there is no disruption in the animal's endogenous lambda light chain immunoglobulin gene.

In some embodiments, the deleted sequence starts from IGLV2 to IGLC1, from IGLV3 to IGLC1, or from IGLJ2 to IGLC1.

Various modifications on the light chain immunoglobulin locus are described e.g., in WO2020169022A1, which is incorporated herein by reference in its entirety.

Genetically modified animals

The present application provides genetically modified non-human animals that produce heavy chain-only antibodies, i.e., antibodies that lack light chains. In some embodiments, the genetically modified non-human animal does not produce IgG (e.g., IgG1) molecules comprising light chains. In some embodiments, the genetically modified non-human animal does not produce any conventional IgG molecules, i.e., IgG antibodies that have two heavy chains and two light chains.

In some embodiments, the genetically modified non-human animal has a fully functional endogenous light chain locus. In some embodiments, the immunoglobulin heavy chain locus of the animal comprises a fully functional IGHM, IGHδ, and/or IGHA gene. In some embodiments, the genome of the animal does not comprise exogenous sequences (e.g., human immunoglobulin heavy chain constant region genes) in the endogenous immunoglobulin heavy chain constant region locus.

In some embodiments, the modified immunoglobulin heavy chain locus lacks the CH1 exon of endogenous IGHG gene (e.g., IGHG1 gene). In some embodiments, the immunoglobulin heavy chain locus of the animal comprises fully functional H, CH2, CH3, M1, and/or M2 exons of IGHG1 gene.

In some embodiments, the immunoglobulin heavy chain locus of the animal only has modifications to the CH1 exon of endogenous IGHG1 gene. In some embodiments, the modification does not comprise a mutation, such as deletion or loss-of-function mutation, of a gene encoding a hinge region of the IgG1. In some embodiments, the modification does not comprise a mutation, such as deletion or loss-of-function mutation to the CH2 or the CH3 exons of the IGHG1 gene. In some embodiments, the immunoglobulin heavy chain locus does not have mutations or modifications to the CH1 exon of IGHM gene. In some embodiments, the genetically modified non-human animal has a functional gene segment encoding the CH1 domain of IgM at the endogenous IGHM locus.

In some embodiments, the genetically modified non-human animal has fully functional genes encoding other heavy chain constant isotypes, such as IgM, IgD, and/or IgA. In some embodiments, the genetically modified non-human animal has fully functional IGHM, IGHδ and/or IGHA genes. In some embodiments, the modified immunoglobulin heavy chain locus has wild-type IGHM, IGHδ, and/or IGHA genes. In some embodiments, the modified immunoglobulin heavy chain locus does not comprise any mutation to the endogenous IGHM, IGHδ and/or IGHA genes. In some embodiments, the endogenous immunoglobulin heavy chain locus has an intact endogenous IGHM, IGHδ, or IGHA gene. In some embodiments, the genetically modified non-human animal expresses wild-type IgM, IgD, and/or IgA proteins.

In some embodiments, the immunoglobulin heavy chain locus includes a mutation (e.g., deletion) or modifications to the CH1 exon of IGHM and/or IGHδ genes. In some embodiments, the endogenous immunoglobulin heavy chain locus has a modified IGHM gene and/or a modified IGHδ gene. In some embodiments, the genetically modified non-human animal does not express wild-type IgM and/or IgD. In some embodiments, the genetically modified non-human animal expresses IgM lacking a CH1 domain and/or IgD lacking a CH1 domain.

In some embodiments, introduction of modifications to the endogenous immunoglobulin heavy chain locus in the genetically modified non-human animal maintain health of the animal, including substantially normal B cell development and maturation. In some embodiments,

introduction of modifications to the endogenous immunoglobulin heavy chain locus in the genetically modified non-human animal reduces or avoids immunogenicity of exogenous sequences. In some embodiments, introduction of minimal changes to the endogenous immunoglobulin heavy chain locus in the genetically modified non-human animal preserves the normal functions of the endogenous immunoglobulin heavy chain locus, including V-D-J recombination, classic switch recombination, and somatic hyper-mutation.

In some embodiments, the genetically modified non-human animal has one or more fully functional light chain loci, e.g., lambda light chain locus, and/or kappa light chain locus. In some embodiments, the genetically modified non-human animal has an unaltered endogenous light chain locus. In some embodiments, no mutations are introduced to the endogenous light chain loci of the genetically modified non-human animal. In some embodiments, the lambda and/or kappa light chain variable region locus of the genetically modified non-human animal is functional, not silenced. In some embodiments, the genetically modified non-human animal expresses a wild-type lambda light chain, and/or wild-type kappa light chain. In some embodiments, the genetically modified non-human animal expresses functional IgM molecules comprising light chains. In some embodiments, the genetically modified non-human animal expresses functional IgA, IgD, and/or IgM molecules comprising light chains. In some embodiments, the genetically modified non-human animal does not have an exogenous light chain gene or gene cluster. For example, the lambda and/or kappa light chain variable region locus of the genetically modified non-human animal can be knocked out.

In some embodiments, the modified immunoglobulin heavy chain locus does not comprises rearranged genes (e.g., rearranged IGHV, IGHD, and/or IGHJ genes). In some embodiments, the modified immunoglobulin heavy chain locus comprises the full set of unrearranged human IGHV, IGHD, and IGHJ genes.

In some embodiments, the modified immunoglobulin heavy chain locus comprises a functional splice site immediately after the CH1 exon of endogenous IGHG1 gene. In some embodiments, the modified immunoglobulin heavy chain locus comprises a wild-type splice site immediately after the CH1 exon of endogenous IGHG1 gene

In some embodiments, the animal described herein expresses a membrane-bound IgG1 lacking a CH1 domain. In some embodiments, the animal described herein expresses a soluble IgG1 lacking a CH1 domain.

In one aspect, the present disclosure provides genetically-modified, non-human animal comprising a humanized heavy chain immunoglobulin locus. In some embodiments, the animal comprises one or more human IGHV genes, one or more human IGHD genes, and/or one or more human IGHJ genes. In some embodiments, these genes are at the endogenous immunoglobulin gene locus.

In some embodiments, the animal comprises an endogenous kappa or lambda chain immunoglobulin locus. In some embodiments, the animal does not comprise an endogenous kappa or lambda chain immunoglobulin locus. In some embodiments, the animal comprises a disruption in the animal's endogenous kappa or lambda light chain immunoglobulin gene locus. In some embodiments, the animal does not have a disruption in the animal's endogenous kappa or lambda light chain immunoglobulin gene locus.

The genetically modified non-human animal can be various animals, e.g., a mouse, rat, rabbit, pig, bovine (e.g., cow, bull, buffalo), deer, sheep, goat, chicken, cat, dog, ferret, primate (e.g., marmoset, rhesus monkey). For the non-human animals where suitable genetically modifiable embryonic stem (ES) cells are not readily available, other methods are employed to make a non-human animal comprising the genetic modification. Such methods include, e.g., modifying a non-ES cell genome (e.g., a fibroblast or an induced pluripotent cell) and employing nuclear transfer to transfer the modified genome to a suitable cell, e.g., an oocyte, and gestating the modified cell (e.g., the modified oocyte) in a non-human animal under suitable conditions to form an embryo. These methods are known in the art, and are described, e.g., in A. Nagy, et al., "Manipulating the Mouse Embryo: A Laboratory Manual (Third Edition)," Cold Spring Harbor Laboratory Press, 2003, which is incorporated by reference herein in its entirety. Thus, in various embodiments, human V, D, and/or J segments can be operably linked to non-human animal (e.g., rodent, mouse, rat, hamster) constant region gene sequences. During B cell development, these rearranged human V, D, and/or J segments are linked to the non-human animal immunoglobulin constant region.

In one aspect, the animal is a mammal, e.g., of the superfamily *Dipodoidea* or *Muroidea*. In some embodiments, the genetically modified animal is a rodent. The rodent can be selected from a mouse, a rat, and a hamster. In some embodiments, the genetically modified animal is from a family selected from *Calomyscidae* (e.g., mouse-like hamsters), *Cricetidae* (e.g., hamster, New World rats and mice, voles), *Muridae* (true mice and rats, gerbils, spiny mice, crested rats), *Nesomyidae* (climbing mice, rock mice, with-tailed rats, Malagasy rats and mice), *Platacanthomyidae* (e.g., spiny dormice), and *Spalacidae* (e.g., mole rates, bamboo rats, and zokors). In some embodiments, the genetically modified rodent is selected from a true mouse or rat (family *Muridae*), a gerbil, a spiny mouse, and a crested rat. In some embodiments, the non-human animal is a mouse.

In some embodiments, the animal is a mouse of a C57BL strain selected from C57BL/A, C57BL/An, C57BL/GrFa, C57BL/KaLwN, C57BL/6, C57BL/6J, C57BL/6ByJ, C57BL/6NJ, C57BL/10, C57BL/10ScSn, C57BL/10Cr, and C57BL/Ola. In some embodiments, the mouse is a 129 strain selected from the group consisting of a strain that is 129P1, 129P2, 129P3, 129X1, 129S1 (e.g., 129S1/SV, 129S1/SvIm), 129S2, 129S4, 129S5, 129S9/SvEvH, 129S6 (129/SvEvTac), 129S7, 129S8, 129T1, 129T2. These mice are described, e.g., in Festing et al., Revised nomenclature for strain 129 mice, Mammalian Genome 10: 836 (1999); Auerbach et al., Establishment and Chimera Analysis of 129/SvEv- and C57BL/6-Derived Mouse Embryonic Stem Cell Lines (2000), both of which are incorporated herein by reference in the entirety. In some embodiments, the genetically modified mouse is a mix of the 129 strain and the C57BL/6 strain. In some embodiments, the mouse is a mix of the 129 strains, or a mix of the BL/6 strains. In some embodiments, the mouse is a BALB strain, e.g., BALB/c strain. In some embodiments, the mouse is a mix of a BALB strain and another strain. In some embodiments, the mouse is from a hybrid line (e.g., 50% BALB/c-50% 12954/Sv; or 50% C57BL/6-50% 129). In some embodiments, the non-human animal is a rodent. In some embodiments, the non-human animal is a mouse having a BALB/c, A, A/He, A/J, A/WySN, AKR, AKR/A, AKR/J, AKR/N, TA1, TA2, RF, SWR, C3H, C57BR, SJL, C57L, DBA/2, KM, NIH, ICR, CFW, FACA, C57BL/A, C57BL/An, C57BL/GrFa, C57BL/KaLwN, C57BL/6, C57BL/6J, C57BL/6ByJ, C57BL/6NJ,

C57BL/10, C57BL/10ScSn, C57BL (C57BL/10Cr and C57BL/Ola), C58, CBA/Br, CBA/Ca, CBA/J, CBA/st, or CBA/H background.

Different animals have different germline organization and genes at their endogenous immunoglobulin heavy chain (IgH) loci. The IgH loci of many species have been sequenced. The gene positions and exon/intron organization of the IgH loci in mouse, rat and rabbit can be found, for example, at IMGT repertoire and NCBI databases, which are incorporated herein by reference in the entirety.

In some embodiments, the animal is a rat. The rat can be selected from a Wistar rat, an LEA strain, a Sprague Dawley strain, a Fischer strain, F344, F6, and Dark Agouti. In some embodiments, the rat strain is a mix of two or more strains selected from the group consisting of Wistar, LEA, Sprague Dawley, Fischer, F344, F6, and Dark Agouti.

The animal can have one or more other genetic modifications, and/or other modifications, that are suitable for the particular purpose for which the humanized animal is made.

Genetically modified non-human animals that comprise a modification of an endogenous non-human immunoglobulin gene locus. In some embodiments, the modification can comprise a human nucleic acid sequence encoding at least a portion of a human protein (e.g., at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% identical to the human heavy chain variable domain or light chain variable domain sequence). Although genetically modified cells are also provided that can comprise the modifications described herein (e.g., ES cells, somatic cells), in many embodiments, the genetically modified non-human animals comprise the modification of the endogenous locus in the germline of the animal.

Genetically modified animals can express a humanized antibody and/or a chimeric antibody from endogenous mouse loci, wherein one or more endogenous mouse immunoglobulin genes have been replaced with human immunoglobulin genes and/or a nucleotide sequence that is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% identical to the human immunoglobulin gene sequences (e.g., IGHV, IGHD, IGHJ, IGKV and/or IGKJ genes). In various embodiments, an endogenous non-human immunoglobulin gene locus is modified in whole or in part to comprise human nucleic acid sequence.

Genetic, molecular and behavioral analyses for the non-human mammals described above can performed. The present disclosure also relates to the progeny produced by the non-human mammal provided by the present disclosure mated with the same or other genotypes. Non-human mammals can be any non-human animal known in the art and which can be used in the methods as described herein. Preferred non-human mammals are mammals, (e.g., rodents). In some embodiments, the non-human mammal is a mouse.

The present disclosure also provides a cell line or primary cell culture derived from the non-human mammal or a progeny thereof. A model based on cell culture can be prepared, for example, by the following methods. Cell cultures can be obtained by way of isolation from a non-human mammal, alternatively cell can be obtained from the cell culture established using the same constructs and the standard cell transfection techniques. The integration of genetic constructs containing DNA sequences encoding human or humanized immunoglobulins can be detected by a variety of methods.

There are many analytical methods that can be used to detect exogenous DNA or modifications on the genomic DNA, including methods at the level of nucleic acid (including the

mRNA quantification approaches using reverse transcriptase polymerase chain reaction (RT-PCR) or Southern blotting, and in situ hybridization) and methods at the protein level (including histochemistry, immunoblot analysis and in vitro binding studies). In addition, the expression level of the gene of interest can be quantified by ELISA techniques well known to those skilled in the art. Many standard analysis methods can be used to complete quantitative measurements. For example, transcription levels can be measured using RT-PCR and hybridization methods including RNase protection, Southern blot analysis, RNA dot analysis (RNAdot) analysis. Immunohistochemical staining, flow cytometry, Western blot analysis can also be used to assess the presence of human or humanized proteins.

Antibodies and Antigen Binding Fragments

The present disclosure provides antibodies and antigen-binding fragments thereof (e.g., heavy chain antibodies, humanized heavy chain antibodies, or multi-specific antibodies) that are produced by the methods described herein.

In general, conventional antibodies are made up of two classes of polypeptide chains, light chains and heavy chains. A non-limiting antibody of the present disclosure can be an intact, four immunoglobulin chain antibody comprising two heavy chains and two light chains. The heavy chain of the antibody can be of any isotype including IgM, IgG, IgE, IgA, or IgD or subclasses including IgG1, IgG2, IgG2a, IgG2b, IgG3, IgG4, IgE1, IgE2, etc. The light chain can be a kappa light chain or a lambda light chain. An antibody can comprise two identical copies of a light chain and two identical copies of a heavy chain. The heavy chains, which each contain one variable domain (or variable region, V_H) and multiple constant domains (or constant regions), bind to one another via disulfide bonding within their constant domains to form the "stem" of the antibody. The light chains, which each contain one variable domain (or variable region, V_L) and one constant domain (or constant region), each bind to one heavy chain via disulfide binding. The variable region of each light chain is aligned with the variable region of the heavy chain to which it is bound. The variable regions of both the light chains and heavy chains contain three hypervariable regions sandwiched between more conserved framework regions (FR).

These hypervariable regions, known as the complementary determining regions (CDRs), form loops that comprise the principle antigen binding surface of the antibody. The four framework regions largely adopt a beta-sheet conformation and the CDRs form loops connecting, and in some cases forming part of, the beta-sheet structure. The CDRs in each chain are held in close proximity by the framework regions and, with the CDRs from the other chain, contribute to the formation of the antigen-binding region.

Methods for identifying the CDR regions of an antibody by analyzing the amino acid sequence of the antibody are well known, and a number of definitions of the CDRs are commonly used. The Kabat definition is based on sequence variability, and the Chothia definition is based on the location of the structural loop regions. These methods and definitions are described in, *e.g.*, Martin, "Protein sequence and structure analysis of antibody variable domains," Antibody engineering, Springer Berlin Heidelberg, 2001. 422-439; Abhinandan, et al. "Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains," Molecular immunology 45.14 (2008): 3832-3839; Wu, T.T. and Kabat, E.A. (1970) J. Exp. Med. 132: 211-250; Martin et al., Methods Enzymol. 203:121-53 (1991); Morea et al.,

Biophys Chem. 68(1-3):9-16 (Oct. 1997); Morea et al., J Mol Biol. 275(2):269-94 (Jan .1998); Chothia et al., Nature 342(6252):877-83 (Dec. 1989); Ponomarenko and Bourne, BMC Structural Biology 7:64 (2007); each of which is incorporated herein by reference in its entirety.

The CDRs are important for recognizing an epitope of an antigen. As used herein, an "epitope" is the smallest portion of a target molecule capable of being specifically bound by the antigen binding domain of an antibody. The minimal size of an epitope may be about three, four, five, six, or seven amino acids, but these amino acids need not be in a consecutive linear sequence of the antigen's primary structure, as the epitope may depend on an antigen's three-dimensional configuration based on the antigen's secondary and tertiary structure.

In some embodiments, the antibody is an intact immunoglobulin molecule (e.g., IgG1, IgG2a, IgG2b, IgG2c, IgG3, IgG4, IgM, IgD, IgE, IgA). The IgG subclasses (IgG1, IgG2, IgG3, and IgG4) are highly conserved, differ in their constant region, particularly in their hinges and upper CH2 domains. The sequences and differences of the IgG subclasses are known in the art, and are described, e.g., in Vidarsson, et al, "IgG subclasses and allotypes: from structure to effector functions." Frontiers in immunology 5 (2014); Irani, et al. "Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases." Molecular immunology 67.2 (2015): 171-182; Shakib, Farouk, ed. The human IgG subclasses: molecular analysis of structure, function and regulation. Elsevier, 2016; each of which is incorporated herein by reference in its entirety. The heavy chain constant regions in the heavy chain antibodies can be derived from any immunoglobulin molecules described herein (e.g., IgG1, IgG2a, IgG2b, IgG2c, IgG3, IgG4, IgM, IgD, IgE, IgA).

The antibody can also be an immunoglobulin molecule that is derived from any species (e.g., human, rodent, mouse, rat, camelid). Antibodies disclosed herein also include, but are not limited to, polyclonal, monoclonal, monospecific, multi-specific antibodies, and chimeric antibodies that include an immunoglobulin binding domain fused to another polypeptide. The term "antigen binding domain" or "antigen binding fragment" is a portion of an antibody that retains specific binding activity of the intact antibody, i.e., any portion of an antibody that is capable of specific binding to an epitope on the intact antibody's target molecule. It includes, e.g., Fab, Fab', $F(ab')_2$, and variants of these fragments. Thus, in some embodiments, an antibody or an antigen binding fragment thereof can be, e.g., a scFv, a Fv, a Fd, a dAb, a bispecific antibody, a bispecific scFv, a diabody, a linear antibody, a single-chain antibody molecule, a multi-specific antibody formed from antibody fragments, and any polypeptide that includes a binding domain which is, or is homologous to, an antibody binding domain. Non-limiting examples of antigen binding domains include, e.g., the heavy chain and/or light chain CDRs of an intact antibody, the heavy and/or light chain variable regions of an intact antibody, full length heavy or light chains of an intact antibody, or an individual CDR from either the heavy chain or the light chain of an intact antibody.

In some embodiments, the antigen binding fragment can form a part of a chimeric antigen receptor (CAR). In some embodiments, the chimeric antigen receptor are fusions of VHH as described herein, fused to CD3-zeta transmembrane- and endodomain.

The antibodies and antigen-binding fragments thereof (e.g., humanized antibodies or chimeric antibodies) that are produced by the methods described herein have various advantages.

In some embodiments, no further optimization is required to obtain desired properties (e.g., binding affinities, thermal stabilities, and/or limited aggregation).

In some implementations, the antibody (or antigen-binding fragments thereof) specifically binds to a target with a dissociation rate (koff) of less than 0.1 s^{-1}, less than 0.01 s^{-1}, less than 0.001 s^{-1}, less than 0.0001 s^{-1}, or less than 0.00001 s^{-1}. In some embodiments, the dissociation rate (koff) is greater than 0.01 s^{-1}, greater than 0.001 s^{-1}, greater than 0.0001 s^{-1}, greater than 0.00001 s^{-1}, or greater than 0.000001 s^{-1}.

In some embodiments, kinetic association rates (kon) are greater than 1 x 10^2/Ms, greater than 1 x 10^3/Ms, greater than 1 x 10^4/Ms, greater than 1 x 10^5/Ms, or greater than 1 x 10^6/Ms. In some embodiments, kinetic association rates (kon) are less than 1 x 10^5/Ms, less than 1 x 10^6/Ms, or less than 1 x 10^7/Ms.

Affinities can be deduced from the quotient of the kinetic rate constants (KD=koff/kon). In some embodiments, KD is less than 1 x 10^{-6} M, less than 1 x 10^{-7} M, less than 1 x 10^{-8} M, less than 1 x 10^{-9} M, or less than 1 x 10^{-10} M. In some embodiments, the KD is less than 50nM, 40 nM, 30 nM, 20 nM, 15 nM, 10 nM, 9 nM, 8 nM, 7 nM, 6 nM, 5 nM, 4 nM, 3 nM, 2 nM, or 1 nM. In some embodiments, KD is greater than 1 x 10^{-7} M, greater than 1 x 10^{-8} M, greater than 1 x 10^{-9} M, greater than 1 x 10^{-10} M, greater than 1 x 10^{-11} M, or greater than 1 x 10^{-12} M. In some embodiments, the antibody binds to a target with KD less than or equal to about 0.9 nM, 0.8 nM, 0.7 nM, 0.6 nM, 0.5 nM, 0.4 nM, 0.3 nM, 0.2 nM, or 0.1 nM.

In some embodiments, thermal stabilities are determined. The antibodies or antigen binding fragments as described herein can have a Tm greater than 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, or 95 °C.

In various embodiments, substitutions are performed to a parental heavy chain antibody sequence to make a variant heavy chain antibody. In general, a heavy chain antibody variant of a parental heavy chain antibody has an antigen binding affinity that is at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or at least 100% (e.g., at least 150%, at least 200%, at least 500%, at least 1000%, or up to at least 10,000%) of the binding affinity of the parental heavy chain antibody to a particular antigen. In some embodiments, a variant heavy chain antibody will comprise a single substitution as compared to a parental heavy chain antibody. However, in other embodiments, several amino acids, e.g., up to about 5 or 10 or more, are substituted as compared to the parental heavy chain antibody sequence that are derived from other human heavy chain sequences that share identity at a given position. In various embodiments, the resultant variant heavy chain antibody is tested to confirm that the desired binding affinity and/or specificity has not been significantly decreased by the replacement residues. In some embodiments, an improved variant heavy chain antibody is produced by the substitution of amino acids from a different human heavy chain sequence. In various embodiments, the VHH is at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a parental VHH.

The VHH described herein can be used to make multi-specific (e.g., bispecific antibodies). In one aspect, the present disclosure provides a multi-specific antibody comprising: a first antigen binding portion and a second antigen binding portion. In some embodiments, the

first antigen binding portion comprises a heavy chain variable domain (VH) and a light chain variable domain (VL), wherein the VH and VL together form an antigen-binding site that specifically binds a first epitope. In some embodiments, the first antigen binding portion comprises a VHH that specifically binds a first epitope. In some embodiments, the second antigen binding portion comprises a VHH that specifically binds a second epitope. In some embodiments, the first epitope and the second epitope are from the same antigen. In some embodiments, the first epitope and the second epitope are from different antigens.

In some embodiments, the first antigen binding portion is a full-length antibody consisting of two heavy chains and two light chains. In some embodiments, the first antigen binding portion is an antibody fragment comprising a heavy chain comprising the VH and a light chain comprising the VL. In some embodiments, the second antigen binding portion comprises a single polypeptide chain. In some embodiments, the C terminus of the second antigen binding portion is fused to the N-terminus of at least one heavy chain of the first antigen binding portion. In some embodiments, the C terminus of the second antigen binding portion is fused to the N-terminus of at least one light chain of the first antigen binding portion. In some embodiments, the N terminus of the second antigen binding portion is fused to the C-terminus of at least one heavy chain of the first antigen binding portion. In some embodiments, the N terminus of the second antigen binding portion is fused to the C-terminus of at least one light chain of the first antigen binding portion. In some embodiments, the second antigen binding portion is a Fab-like domain comprising a first polypeptide chain comprising a first VHH fused to a CH1 domain, and a second polypeptide chain comprising a second VHH fused to a CL domain.

In some embodiments, the antibody or antigen-binding fragment thereof is a tri-specific antibody. In some embodiments, the tri-specific antibody is a tri-specific VHH-Fc. In some embodiments, the tri-specific antibody comprises the same VHHs. In some embodiments, the tri-specific antibody comprises different VHHs. In some embodiments, the VHHs bind to the same epitope. In some embodiments, the VHHs bind to different epitopes.

In some embodiments, the antibody or antigen-binding fragment thereof has four or more than four VHHs. In some embodiments, in order to increase developability, at least four VHHs are combined without the addition of IgG Fc domain to construct tetra-specific VHHs. These molecules would have the added advantage of increased affinity and avidity towards the antigen compared to bi- and tri-specific VHH-Fcs, despite lacking the Fc effector functions.

In some embodiments, these the antibody or antigen-binding fragment thereof (e.g., comprising 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10) has a functional Fc.

In some embodiments, the heavy-chain antibody produced by the genetically modified non-human animal described herein has a VHH domain that includes CDR1, CDR2, and CDR3. In some embodiments, the CDR3 length is between 6-23, e.g., 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or 23. In some embodiments, the CDR3 length is at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, or at least 23.

Methods of making genetically modified animals

The genetically modified animals can be made by modifying the immunoglobulin gene locus. **FIGS. 6-12** shows workflows of genetic modifications using targeting vectors V1-V7,

respectively. **FIG. 20** shows the methods of making the humanized animals. In some embodiments, the methods first involve modifying the human immunoglobulin locus on the human chromosome. The modified human chromosomes are then introduced into the mouse recipient cell. The human immunoglobulin variable region is then introduced into the corresponding region of the mouse genome by direct replacement. Then, the recipient cells are screened. In some embodiments, the cells do not contain the human chromosomes. The cells are then injected to blastocysts to prepare chimeric mice. Subsequent breeding can be performed to obtain mice containing intact humanized immunoglobulin locus.

Several other techniques may be used in making genetically modified animals, including, e.g., nonhomologous end-joining (NHEJ), homologous recombination (HR), zinc finger nucleases (ZFNs), transcription activator-like effector-based nucleases (TALEN), and the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas system. In some embodiments, homologous recombination is used. In some embodiments, CRISPR-Cas9 genome editing is used to generate genetically modified animals. Many of these genome editing techniques are known in the art, and is described, e.g., in Yin et al., "Delivery technologies for genome editing," Nature Reviews Drug Discovery 16.6 (2017): 387-399, which is incorporated by reference in its entirety. Many other methods are also provided and can be used in genome editing, e.g., micro-injecting a genetically modified nucleus into an enucleated oocyte, and fusing an enucleated oocyte with another genetically modified cell.

The genetic modification process can involve replacing endogenous sequence with human sequence by homologous recombination. In some embodiments, the cleavage at the upstream and the downstream of the target site (e.g., by zinc finger nucleases, TALEN or CRISPR) can result in DNA double strands break, and the homologous recombination is used to replace endogenous sequence with human sequence.

In some embodiments, the methods for deleting CH1 sequence in an IGHG gene involve one or the combination of the following methods. These modifications can be performed in various cells. In some embodiments, the cell is a stem cell, an embryonic stem cell, or a fertilized egg cell.

In some embodiments, a sequence starting from the Sγ3 switch region to Cε is knocked out, then a sequence comprising the Sγ1 switch region and a Cγ1 sequence without CH1 (abbreviated Cγ1ΔCH1) is inserted. As a result, the Cγ1ΔCH1 is operably linked with Sγ1.

In some embodiments, a sequence comprising the Sγ1 switch region and a Cγ1 sequence without CH1 (abbreviated Cγ1ΔCH1) is used to directly replace all sequences starting from Sγ3 switch region to Cε (**FIG. 6**). As a result, the Cγ1ΔCH1 is operably linked with Sγ1.

In some embodiments, Sγ3 switch region and the Cγ3 sequence are knocked out first, and then a Cγ1 sequence without CH1 (abbreviated as Cγ1ΔCH1) is used to replace all the sequences from Cγ1 to Cε. As a result, the Cγ1ΔCH1 is operably linked with Sγ1.

In some embodiments, a Cγ1 sequence without CH1 (abbreviated as Cγ1ΔCH1) is used to directly replace all sequences starting from Cγ3 to Cε (**FIG. 7**). As a result, the Cγ1ΔCH1 is operably linked with Sγ3.

In some embodiments, based on the allele shown in **FIG. 6**, a sequence including Cμ and Cδ is knocked out (**FIG. 8**). As a result, both IGHM and IGHδ genes are knocked out, and the Cγ1ΔCH1 is operably linked with Sγ1.

In some embodiments, based on the allele shown in **FIG. 6**, a sequence including Cμ, Cδ, and Sγ1 is knocked out (**FIG. 9**). As a result, both IGHM and IGHδ genes are knocked out, and the Cγ1ΔCH1' is operably linked with Sμ.

In some embodiments, based on the allele shown in **FIG. 6**, the CH1 coding region of Cμ is knocked out (**FIG. 10**). As a result, the modified locus includes sequences encoding IgM lacking a CH1 domain, IgD, and IgG1 lacking a CH1 domain. In addition, the Cγ1ΔCH1 is operably linked with Sγ1.

In some embodiments, based on the allele shown in **FIG. 6**, a Cμ sequence without the CH1 coding region (abbreviated as CμΔCH1) is used to directed replace a sequence including Cμ and Cδ (**FIG. 11**). As a result, the modified locus includes sequences encoding IgM lacking a CH1 domain, and IgG1 lacking a CH1 domain. In addition, the Cγ1ΔCH1 is operably linked with Sγ1.

In some embodiments, based on the allele shown in **FIG. 6**, a sequence including CμΔCH1 and Cδ without the CH1 coding region (abbreviated as CδΔCH1) is used to directed replace a sequence including Cμ and Cδ (**FIG. 12**). As a result, the modified locus includes sequences encoding IgM lacking a CH1 domain, IgD lacking a CH1 domain, and IgG1 lacking a CH1 domain. In addition, the Cγ1ΔCH1 is operably linked with Sγ1.

In some embodiments, provided herein is a genetically modified non-human animal comprising a modified immunoglobulin heavy chain constant region locus. In some embodiments, the modified immunoglobulin heavy chain constant region locus comprises a sequence that is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100 identical to the sequence of Cγ1ΔCH1, Cγ1ΔCH1', CμΔCH1, or CδΔCH1. In some embodiments, the sequence of Cγ1ΔCH1' includes a deletion of at least 1, at least 2, at least 3, or at least 4 nucleotides at the 5' end of the sequence of Cγ1ΔCH1 (SEQ ID NO: 1).

The present disclosure also relates to a genetically modified non-human animal comprising a nucleic acid sequence, wherein the nucleic acid sequence can be selected from the group consisting of:

a) a nucleic acid sequence as shown in SEQ ID NO: 1, 8, 9, 10, 13, or 41;

b) a nucleic acid sequence that is able to hybridize to the nucleotide sequence as shown in SEQ ID NO: 1, 8, 9, 10, 13, or 41 under a low stringency condition or a strict stringency condition;

c) a nucleic acid sequence that has a homology of at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, or at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the nucleotide sequence as shown in SEQ ID NO: 1, 8, 9, 10, 13, or 41; and

d) a nucleic acid sequence that encodes an amino acid sequence, wherein the amino acid sequence has a homology of at least 90% with or at least 90% identical to the amino acid sequence of endogenous IgG, IgM, IgD, or IgA.

The method further involves transplanting the genetically modified cell to the fallopian tube or uterus of the recipient female non-human mammal, allowing the cell to develop in the uterus of the female non-human mammal. In some embodiments, experiments are performed to identify the germline transmission in the offspring the genetically modified genes.

In some embodiments, the methods for making a genetically modified, humanized animal, can also include the step of replacing at an endogenous locus (or site), a nucleic acid (e.g., V, D,

J regions, or V, J regions) with a corresponding region of human sequence. The sequence can include a region (e.g., a part or the entire region) of IGHV, IGHD, IGHJ, IGKV, and/or IGKJ genes. In some embodiments, the replacement is mediated by homologous recombination. In some embodiments, the replacement is mediated by Cre recombinase.

The 5' end homology arm and/or the 3' end homology arm can have a desired length to facilitate homologous recombination. In some embodiments, the homology arm is about or at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, or 50 kb (e.g., about 3kb). In some embodiments, the homology arm is less than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, or 50 kb.

In some embodiments, the vector may also optionally include a reporter protein, e.g., a luciferase (e.g., Gluc) or a fluorescent protein (e.g., EGFP, BFP, etc.).

These modifications can be performed in various cells. In some embodiments, the cell is a stem cell, an embryonic stem cell, or a fertilized egg cell.

The present disclosure further provides a method for establishing a humanized animal model, involving the following steps:

(a) providing the cell (e.g. a fertilized egg cell) based on the methods described herein;

(b) culturing the cell in a liquid culture medium;

(c) transplanting the cultured cell to the fallopian tube or uterus of the recipient female non-human mammal, allowing the cell to develop in the uterus of the female non-human mammal;

(d) identifying the germline transmission in the offspring genetically modified humanized non-human mammal of the pregnant female in step (c).

In some embodiments, the non-human mammal in the foregoing method is a mouse (e.g., a C57 mouse, a BALB/c mouse, or a C57BL/6 mouse).

In some embodiments, the non-human mammal in step (c) is a female with pseudo pregnancy (or false pregnancy).

In some embodiments, the fertilized eggs for the methods described above are C57BL/6 fertilized eggs. Other fertilized eggs that can also be used in the methods as described herein include, but are not limited to, FVB/N fertilized eggs, BALB/c fertilized eggs, DBA/1 fertilized eggs and DBA/2 fertilized eggs.

Fertilized eggs can come from any non-human animal, e.g., any non-human animal as described herein. In some embodiments, the fertilized egg cells are derived from rodents. The genetic construct can be introduced into a fertilized egg by microinjection of DNA. For example, by way of culturing a fertilized egg after microinjection, a cultured fertilized egg can be transferred to a false pregnant non-human animal, which then gives birth of a non-human mammal, so as to generate the non-human mammal mentioned in the methods described above.

Cells, tissues, and animals (e.g., mouse) are also provided that comprise the nucleotide sequences as described herein, as well as cells, tissues, and animals (e.g., mouse) that express humanized or chimeric antibodies from an endogenous non-human locus.

The present disclosure also provides various targeting vectors (e.g., vectors that are useful for making the genetically modified animals). In some embodiments, the vector can comprise: a) a DNA fragment homologous to the 5' end of a region to be altered (5' homology arm); b) a sequence comprising desired genetic elements (e.g., LoxP recognition site, drug resistance genes, and/or reporter genes etc.); and c) a second DNA fragment homologous to the 3'

end of the region to be altered (3' homology arm). The disclosure also relates to a cell comprising the targeting vectors as described herein.

In some embodiments, the genes in the cell are heterozygous. In some embodiments, the genes in the cell are homozygous.

In some embodiments, the non-human mammalian cell is a mouse cell. In some embodiments, the cell is a fertilized egg cell.

The disclosure also provides a nucleic acid sequence that 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%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to any nucleotide sequence as described herein, and an amino acid sequence that 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%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to any amino acid sequence as described herein.

In some embodiments, the disclosure relates to nucleotide sequences encoding any peptides that are described herein, or any amino acid sequences that are encoded by any nucleotide sequences as described herein. In some embodiments, the nucleic acid sequence is less than 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 150, 200, 250, 300, 350, 400, 500, or 600 nucleotides. In some embodiments, the amino acid sequence is less than 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, or 400 amino acid residues.

In some embodiments, the amino acid sequence (i) comprises an amino acid sequence; or (ii) consists of an amino acid sequence, wherein the amino acid sequence is any one of the sequences as described herein.

In some embodiments, the nucleic acid sequence (i) comprises a nucleic acid sequence; or (ii) consists of a nucleic acid sequence, wherein the nucleic acid sequence is any one of the sequences as described herein.

To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. For purposes of illustration, the comparison of sequences and determination of percent identity between two sequences can be accomplished using a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5 when appropriate.

The percentage of residues conserved with similar physicochemical properties (percent homology), e.g. leucine and isoleucine, can also be used to measure sequence similarity.

Families of amino acid residues having similar physicochemical properties have been defined in the art. These families include e.g., amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). The homology percentage, in many cases, is higher than the identity percentage. Thus, the present disclosure also provides an amino acid sequence that has 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%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% homology percentage to any amino acid sequence as described herein, or a nucleic acid encoding these amino acid sequences.

Methods of using genetic modified animals

The genetic modified animals can be used to generate heavy-chain antibodies that can bind specifically to a target. In some embodiments, the target (e.g., a protein or a fragment of the protein) can be used as an immunogen to generate antibodies in these animals using standard techniques for polyclonal and monoclonal antibody preparation. In some embodiments, the genetic modified animal is exposed to a selected antigen for a time and under conditions which permit the animal to produce antibody specific for the antigen.

Polyclonal antibodies can be raised in animals by multiple injections (e.g., subcutaneous or intraperitoneal injections) of an antigenic peptide or protein. In some embodiments, the antigenic peptide or protein is injected with at least one adjuvant. In some embodiments, the antigenic peptide or protein can be conjugated to an agent that is immunogenic in the species to be immunized. Animals can be injected with the antigenic peptide or protein more than one time (e.g., twice, three times, or four times).

The full-length polypeptide or protein can be used or, alternatively, antigenic peptide fragments thereof can be used as immunogens. The antigenic peptide of a protein comprises at least 8 (e.g., at least 10, 15, 20, or 30) amino acid residues of the amino acid sequence and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with the protein.

An immunogen typically is used to prepare antibodies by immunizing a suitable subject (e.g., the genetically modified animal as described herein). An appropriate immunogenic preparation can contain, for example, a recombinantly-expressed or a chemically-synthesized polypeptide (e.g., a fragment of the protein). The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or a similar immunostimulatory agent.

The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme-linked immunosorbent assay (ELISA) using the immobilized polypeptide or peptide. If desired, the antibody molecules can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A of protein G chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma

technique originally described by Kohler et al. (*Nature* 256:495-497, 1975), the human B cell hybridoma technique (Kozbor et al., *Immunol. Today* 4:72, 1983), the EBV-hybridoma technique (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96, 1985), or trioma techniques. The technology for producing hybridomas is well known (see, generally, Current Protocols in Immunology, 1994, Coligan et al. (Eds.), John Wiley & Sons, Inc., New York, NY). Hybridoma cells producing a monoclonal antibody are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide or epitope of interest, e.g., using a standard ELISA assay.

In one aspect, the disclosure provides a mouse that comprises a modification of an endogenous immunoglobulin heavy chain locus, wherein the mouse produces a B cell that comprises a rearranged immunoglobulin sequence operably linked to a heavy chain constant region gene sequence. In some embodiment, the rearranged immunoglobulin sequence operably linked to the heavy chain constant region gene sequence comprises a human heavy chain V, D, and/or J sequence. In some embodiments, the heavy chain constant region gene sequence comprises a human or a mouse heavy chain sequence selected from the group consisting of a CH1, a hinge, a CH2, a CH3, and a combination thereof.

In one aspect, the disclosure relates to a method for making a somatically mutated heavy chain antibody in an animal. The method involves immunizing an animal with an antigen, maintaining the animal under conditions sufficient to initiate an immune response to the antigen; and, isolating from the animal a somatically mutated heavy chain antibody comprising a variable domain that is derived from the human or endogenous heavy chain immunoglobulin variable region gene segment, and wherein the somatically mutated heavy chain antibody specifically binds the antigen. In some embodiments, the animal comprises an unrearranged human or endogenous heavy chain immunoglobulin variable region gene segment, and wherein the animal lacks a nucleotide sequence of at least one allele that encodes a functional IgG CH1 domain, and wherein the animal expresses an IgM that comprises a CH1 domain;

The B cells or spleen cells can comprise a rearranged non-mouse immunoglobulin variable gene sequence, e.g., operably linked to a mouse immunoglobulin constant region gene. The sequences for encoding human heavy chain variable region are determined. The sequences can be determined by e.g., sequencing the hybridoma of interest or B cells. In some embodiments, single B cell screening is used. It can screen the natural antibody repertoire without the need for hybridoma fusion and combinatorial display. For example, B cells can be mixed with a panel of DNA-barcoded antigens, such that both the antigen barcode(s) and B-cell receptor (BCR) sequences of individual B cells are recovered via single-cell sequencing protocols.

The antibodies can be further modified to obtain a humanized antibody or a human antibody, e.g., by operably linking the sequence encoding human heavy chain variable region to a sequence encoding a human heavy chain constant region.

In some embodiments, if the mouse expresses a protein that is very similar to the antigen of interest, it can be difficult to elicit an immune response in the mouse. This is because during immune cell development, B-cells and T-cells that recognize MHC molecules bound to peptides of self-origin are deleted from the repertoire of immune cells. In those cases, the humanized mouse can be further modified. The corresponding gene in the mouse can be knocked out, and

the mouse is then exposed to the antigen of interest. Because the mouse does not go through negative selection for the gene product, the mouse can generate an antibody that can specifically bind to the target easily.

The disclosure also provides methods of making antibodies, nucleic acids, cells, tissues (e.g., spleen tissue). In some embodiments, the methods involve exposing the animal as described herein to the antigen. Antibodies (e.g., hybrid antibodies), nucleic acids encoding the antibodies, cells, and/or tissues (e.g., spleen tissue) can be obtained from the animal. In some embodiments, the nucleic acids encoding the variable regions are determined, e.g., by sequencing. In some embodiments, the nucleic acid encoding the human heavy chain immunoglobulin variable region can be operably linked with a nucleic acid encoding a human heavy chain immunoglobulin constant region. In some embodiments, the cells containing the nucleic acids as described herein are cultured and the antibodies are collected.

In some embodiments, no mouse immunoglobulin V, D, J genes (e.g., no mouse IGHV, IGHD, IGHJ, IGKV, or IGKJ genes) contributes to the heavy chain variable region sequence. In some embodiments, the heavy chain variable region sequence produced by the animal are fully human, and are completely contributed by human immunoglobulin V, D, J genes (e.g., human IGHV, IGHD, IGHJ, IGKV, and IGKJ genes). In some embodiments, the rearranged VDJ sequences can further undergo somatic hypermutations.

Variants of the antibodies or antigen-binding fragments described herein can be prepared by introducing appropriate nucleotide changes into the DNA encoding a human, humanized, or chimeric antibody, or antigen-binding fragment thereof described herein, or by peptide synthesis. Such variants include, for example, deletions, insertions, or substitutions of residues within the amino acids sequences that make-up the antigen-binding site of the antibody or an antigen-binding domain. In a population of such variants, some antibodies or antigen-binding fragments will have increased affinity for the target protein. Any combination of deletions, insertions, and/or combinations can be made to arrive at an antibody or antigen-binding fragment thereof that has increased binding affinity for the target. The amino acid changes introduced into the antibody or antigen-binding fragment can also alter or introduce new post-translational modifications into the antibody or antigen-binding fragment, such as changing (e.g., increasing or decreasing) the number of glycosylation sites, changing the type of glycosylation site (e.g., changing the amino acid sequence such that a different sugar is attached by enzymes present in a cell), or introducing new glycosylation sites.

Antibodies disclosed herein can be derived from any species of animal, including mammals. Non-limiting examples of native antibodies include antibodies derived from humans, primates, e.g., monkeys and apes, cows, pigs, horses, sheep, camelids (e.g., camels and llamas), chicken, goats, and rodents (e.g., rats, mice, hamsters and rabbits), including transgenic rodents genetically engineered to produce human antibodies.

Human and humanized antibodies include antibodies having variable and constant regions derived from (or having the same amino acid sequence as those derived from) human germline immunoglobulin sequences. Human antibodies may include amino acid residues not encoded by human germline immunoglobulin sequences (e.g., mutations introduced by random or site-specific mutagenesis in vitro or by somatic mutation in vivo), for example in the CDRs.

Additional modifications to the antibodies or antigen-binding fragments can be made. For example, a cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have any increased half-life *in vitro* and/or *in vivo*. Homodimeric antibodies with increased half-life *in vitro* and/or *in vivo* can also be prepared using heterobifunctional cross-linkers as described, for example, in Wolff et al. (*Cancer Res.* 53:2560-2565, 1993). Alternatively, an antibody can be engineered which has dual Fc regions (see, for example, Stevenson et al., *Anti-Cancer Drug Design* 3:219-230, 1989).

In some embodiments, a covalent modification can be made to the antibody or antigen-binding fragment thereof. These covalent modifications can be made by chemical or enzymatic synthesis, or by enzymatic or chemical cleavage. Other types of covalent modifications of the antibody or antibody fragment are introduced into the molecule by reacting targeted amino acid residues of the antibody or fragment with an organic derivatization agent that is capable of reacting with selected side chains or the N- or C-terminal residues.

Transferrin receptor 1 (TFR1)

TFR1, also known as cluster of differentiation 71 (CD71), is widely expressed and can bind to transferrin (Tf) with high affinity. Human TFR1 is a 90 kDa type II transmembrane glycoprotein consisting of 760 amino acids that is found as a dimer (180 kDa) linked by disulfide bonds on the cell surface. The TFR1 monomer is composed of a large extracellular, C-terminal domain of 671 amino acids containing the Tf-binding site, a transmembrane domain (28 amino acids), and an intracellular N-terminal domain (61 amino acids). The C-terminal extracellular domain contains three N-linked glycosylation sites at asparagine residues 251, 317, and 727 and one O-linked glycosylation site at threonine 104, which are all required for adequate function of the receptor.

Transferrin (Tf) It is an 80 kDa glycoprotein composed of two 40 kDa subunits, known as the N- and C-lobes that are separated by a short linker sequence. Each subunit is capable of binding to one free ferric iron (Fe^{3+}) and thus, Tf may have up to two atoms of iron attached. Tf in its iron free form, apo-Tf, binds Fe^{3+} with high efficiency in the blood and transports it to the cell surface for internalization through the interaction with TFR1. As a membrane protein regulating iron import, TFR1 is a member of the TFR family that shows nanomolar affinity to transferrin (Tf) bound to Fe^{3+}. The complex of Tf-TFR1 is internalized through endocytosis mediated by clathrin, and Fe^{3+} is disassociated from Tf when pH decreases to 5.5. At this pH, apo-Tf and TFR1 are still associated and recycled to cell surface with physiological pH, so the former is released.

Iron uptake by transferrin receptor is an important way for cancer cells to absorb iron, thus accumulating evidence has proven that TFR1 participated in tumor onset and progression, and its expression was dysregulated significantly in many cancers. The relationship between TFR1 and cancers has been revealed, rendering TFR1 a valuable pharmaceutical target for intervening with cancers.

TFR1 expressed on the endothelial cells of the blood-brain barrier is used also in preclinical research to allow the delivery of large molecules including antibodies into the brain.

The TFR1 targeting antibodies can cross the blood-brain barrier, without interfering with the uptake of iron.

A detailed description of TFR1, Tf, and their functions can be found, e.g., in Candelaria, P. V., et al. "Antibodies targeting the transferrin receptor 1 (TfR1) as direct anti-cancer agents." Frontiers in Immunology 12 (2021): 607692; and Shen, Y., et al. "Transferrin receptor 1 in cancer: a new sight for cancer therapy." American Journal of Cancer Research 8.6 (2018): 916; each of which is incorporated by reference in its entirety.

Heavy chain single variable domain (VHH) antibodies

Monoclonal and recombinant antibodies are important tools in medicine and biotechnology. Like all mammals, camelids (e.g., llamas) can produce conventional antibodies made of two heavy chains and two light chains bound together with disulfide bonds in a Y shape (e.g., IgG1). However, they also produce two unique subclasses of IgG: IgG2 and IgG3, also known as heavy chain IgG. These antibodies are made of only two heavy chains, which lack the CH1 region but still bear an antigen-binding domain at their N-terminus called VHH (or nanobody). Conventional Ig require the association of variable regions from both heavy and light chains to allow a high diversity of antigen-antibody interactions. Although isolated heavy and light chains still show this capacity, they exhibit very low affinity when compared to paired heavy and light chains. The unique feature of heavy chain IgG is the capacity of their monomeric antigen binding regions to bind antigens with specificity, affinity and especially diversity that are comparable to conventional antibodies without the need of pairing with another region. This feature is mainly due to a couple of major variations within the amino acid sequence of the variable region of the two heavy chains, which induce deep conformational changes when compared to conventional Ig. Major substitutions in the variable regions prevent the light chains from binding to the heavy chains, but also prevent unbound heavy chains from being recycled by the immunoglobulin binding proteins.

The single variable domain of these antibodies (designated VHH, sdAb, or nanobody) is the smallest antigen-binding domain generated by adaptive immune systems. The third Complementarity Determining Region (CDR3) of the variable region of these antibodies has been found to be twice as long as the conventional ones. This results in an increased interaction surface with the antigen as well as an increased diversity of antigen-antibody interactions, which compensates the absence of the light chains. With a long complementarity-determining region 3 (CDR3), VHHs can extend into crevices on proteins that are not accessible to conventional antibodies, including functionally interesting sites such as the active site of an enzyme or the receptor-binding canyon on a virus surface.

VHHs offer numerous other advantages compared to conventional antibodies carrying variable domains (VH and VL) of conventional antibodies, including higher stability, solubility, expression yields, and refolding capacity, as well as better *in vivo* tissue penetration and internalization. Moreover, in contrast to the VH domains of conventional antibodies, VHH do not display an intrinsic tendency to bind to light chains. Since VHH do not bind to VL domains, it is much easier to reformat VHHs into multi-specific (e.g., bispecific antibody) constructs than constructs containing conventional VH-VL pairs or single domains based on VH domains.

The disclosure provides e.g., anti-TFR1 antibodies, the modified antibodies thereof, the chimeric antibodies thereof, and the humanized antibodies thereof.

The CDR sequences for 23B8, and 23B8 derived antibodies (e.g., humanized antibodies) include CDRs of the VHH domain as set forth in SEQ ID NOs: 42, 43, and 44, respectively, as defined by Kabat numbering. The CDRs can also be defined by IMGT system. Under the IMGT numbering, the CDRs of the VHH domain are set forth in SEQ ID NOs: 54, 55, and 56, respectively.

The CDR sequences for 24A1, and 24A1 derived antibodies (e.g., humanized antibodies) include CDRs of the VHH domain as set forth in SEQ ID NOs: 45, 46, and 47, respectively, as defined by Kabat numbering. The CDRs can also be defined by IMGT system. Under the IMGT numbering, the CDRs of the VHH domain are set forth in SEQ ID NOs: 57, 58, and 59, respectively.

The CDR sequences for 24C9, and 24C9 derived antibodies (e.g., humanized antibodies) include CDRs of the VHH domain as set forth in SEQ ID NOs: 48, 49, and 50, respectively, as defined by Kabat numbering. The CDRs can also be defined by IMGT system. Under the IMGT numbering, the CDRs of the VHH domain are set forth in SEQ ID NOs: 60, 61, and 62, respectively.

The CDR sequences for 24G5, and 24G5 derived antibodies (e.g., humanized antibodies) include CDRs of the VHH domain as set forth in SEQ ID NOs: 51, 52, and 53, respectively, as defined by Kabat numbering. The CDRs can also be defined by IMGT system. Under the IMGT numbering, the CDRs of the VHH domain are set forth in SEQ ID NOs: 63, 64, and 65, respectively.

The amino acid sequence for the VHH domain of 23B8 antibody is set forth in SEQ ID NO: 66. The amino acid sequence for the VHH domain of 24A1 antibody is set forth in SEQ ID NO: 67. The amino acid sequence for the VHH domain of 24C9 antibody is set forth in SEQ ID NO: 68. The amino acid sequence for the VHH domain of 24G5 antibody is set forth in SEQ ID NO: 69.

The amino acid sequences for various modified or humanized VHH are also provided. As there are different ways to modify or humanize a heavy-chain antibody (e.g., a sequence can be modified with different amino acid substitutions), the VHH domain of a heavy-chain antibody can have more than one version of humanized sequences. In some embodiments, the humanized VHH domain is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to any sequence of SEQ ID NOS: 66-69.

Furthermore, in some embodiments, the antibodies or antigen-binding fragments thereof described herein can also contain one, two, or three VHH domain CDRs selected from the group of SEQ ID NOs: 42-44, SEQ ID NOs: 45-47, SEQ ID NOs: 48-50, SEQ ID NOs: 51-53, SEQ ID NOs: 54-56, SEQ ID NOs: 57-59, SEQ ID NOs: 60-62, and SEQ ID NOs: 63-65.

In some embodiments, the antibodies can have a heavy chain single variable domain (VHH) comprising complementarity determining regions (CDRs) 1, 2, 3, wherein the CDR1 region comprises or consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VHH CDR1 amino acid sequence, the CDR2 region comprises or consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VHH CDR2 amino acid sequence, and the CDR3 region comprises or consists of an amino acid

sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VHH CDR3 amino acid sequence. The selected VHH CDRs 1, 2, 3 amino acid sequences is shown in **FIG. 37** and **FIG. 38.**

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of VHH CDR1 with zero, one or two amino acid insertions, deletions, or substitutions; VHH CDR2 with zero, one or two amino acid insertions, deletions, or substitutions; VHH CDR3 with zero, one or two amino acid insertions, deletions, or substitutions, wherein VHH CDR1, VHH CDR2, and VHH CDR3 are selected from the CDRs in **FIG. 39**.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 42 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 43 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 44 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 45 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 46 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 47 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 48 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 49 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 50 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 51 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 52 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 53 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 54 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 55 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 56 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 57 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 58 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 59 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 60 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ

ID NO: 61 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 62 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain single variable domain (VHH) containing one, two, or three of the CDRs of SEQ ID NO: 63 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 64 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 65 with zero, one or two amino acid insertions, deletions, or substitutions.

The insertions, deletions, and substitutions can be within the CDR sequence, or at one or both terminal ends of the CDR sequence. In some embodiments, the CDR is determined based on Kabat numbering scheme. In some embodiments, the CDR is determined based on Chothia numbering scheme. In some embodiments, the CDR is determined based on a combination numbering scheme. In some embodiments, the CDR is determined based on IMGT numbering scheme.

The disclosure also provides antibodies or antigen-binding fragments thereof that bind to TFR1 (human TFR1). The antibodies or antigen-binding fragments thereof contain a heavy chain single variable region (VHH) comprising or consisting of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VHH sequence. In some embodiments, the selected VHH sequence is SEQ ID NO: 66. In some embodiments, the selected VHH sequence is SEQ ID NO: 67. In some embodiments, the selected VHH sequence is SEQ ID NO: 68. In some embodiments, the selected VHH sequence is SEQ ID NO: 69.

To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. For purposes of illustration, the comparison of sequences and determination of percent identity between two sequences can be accomplished, e.g., using a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

The disclosure also provides nucleic acid comprising a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain single variable domain (VHH). The VHH comprises CDRs as shown in **FIG. 37** and **FIG. 38**, or has sequences as shown in **FIG. 39**.

The antibodies and antigen-binding fragments can also be antibody variants (including derivatives and conjugates) of antibodies or antibody fragments and multi-specific (e.g., bi-specific) antibodies or antibody fragments. Additional antibodies provided herein are polyclonal, monoclonal, multi-specific (multimeric, e.g., bi-specific), human antibodies, chimeric antibodies (e.g., human-mouse chimera), single-chain antibodies, intracellularly-made antibodies (i.e., intrabodies), and antigen-binding fragments thereof.

In some embodiments, the antibodies or antigen-binding fragments thereof comprises an Fc domain that can be originated from various types (e.g., IgG, IgE, IgM, IgD, IgA, and IgY), class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2), or subclass. In some embodiments, the Fc domain is originated from an IgG antibody or antigen-binding fragment thereof. In some embodiments, the Fc domain comprises one, two, three, four, or more heavy chain constant regions.

The present disclosure also provides an antibody or antigen-binding fragment thereof that cross-competes with any antibody or antigen-binding fragment as described herein. The cross-competing assay is known in the art, and is described e.g., in Moore et al., "Antibody cross-competition analysis of the human immunodeficiency virus type 1 gp120 exterior envelope glycoprotein." Journal of virology 70.3 (1996): 1863-1872, which is incorporated herein reference in its entirety. In one aspect, the present disclosure also provides an antibody or antigen-binding fragment thereof that binds to the same epitope or region as any antibody or antigen-binding fragment as described herein. The epitope binning assay is known in the art, and is described e.g., in Estep et al. "High throughput solution-based measurement of antibody-antigen affinity and epitope binning." MAbs. Vol. 5. No. 2. Taylor & Francis, 2013, which is incorporated herein reference in its entirety.

In some embodiments, the antibody or antigen-binding fragment thereof comprises a heavy chain single variable domain (VHH) CDR1 selected from SEQ ID NOs: 42, 45, 48, 51, 54, 57, 60, and 63

In some embodiments, the antibody or antigen-binding fragment thereof comprises a heavy chain single variable domain (VHH) CDR2 selected from SEQ ID NOs: 43, 46, 49, 52, 55, 58, 61, and 64.

In some embodiments, the antibody or antigen-binding fragment thereof comprises a heavy chain single variable domain (VHH) CDR3 selected from SEQ ID NOs: 44, 47, 50, 53, 56, 59, 62, and 65.

Antibody characteristics

TFR1 performs a critical role in cellular iron uptake through the interaction with iron-bound TF. Iron is required for multiple cellular processes and is essential for DNA synthesis and, thus, cellular proliferation. Due to its central role in cancer cell pathology, malignant cells often overexpress TFR1 and this increased expression can be associated with poor prognosis in different types of cancer. The elevated levels of TfR1 expression on malignant cells, together with its extracellular accessibility, ability to internalize, and central role in cancer cell pathology make this receptor an attractive target for antibody-mediated therapy.

In some embodiments, the antibodies or antigen-binding fragments thereof described herein cannot block the binding between TFR1 and TF. In some embodiments, the antibodies or antigen-binding fragments thereof described herein can block the binding between TFR1 and TF. In some embodiments, the antibodies or antigen-binding fragments thereof described herein can be conjugated to anti-cancer agents that are internalized by receptor-mediated endocytosis. In some embodiments, the antibodies or antigen-binding fragments thereof described herein can disrupt the function of the receptor. In some embodiments, the antibodies or antigen-binding

fragments thereof described herein cannot induce Fc effector functions, therefore preventing or ameliorating their negative effects to normal cells.

The disclosure provides antibodies or antigen-binding fragments thereof comprising a human Fc domain, which induce Fc-dependent effector functions by at least or about at least or about 1 fold, at least or about 2 folds, at least or about 3 folds, at least or about 4 folds, at least or about 5 folds, at least or about 6 folds, at least or about 7 folds, at least or about 8 folds, at least or about 9 folds, at least or about 10 folds, at least or about 20 folds, at least or about 30 folds, at least or about 40 folds, at least or about 50 folds, or at least or about 100 folds as compared when no antibodies or antigen-binding fragments thereof as described herein are present.

The disclosure provides antibodies or antigen-binding fragments thereof comprising a human Fc domain, which induce host immune response by at least or about at least or about 1 fold, at least or about 2 folds, at least or about 3 folds, at least or about 4 folds, at least or about 5 folds, at least or about 6 folds, at least or about 7 folds, at least or about 8 folds, at least or about 9 folds, at least or about 10 folds, at least or about 20 folds, at least or about 30 folds, at least or about 40 folds, at least or about 50 folds, or at least or about 100 folds as compared when no antibodies or antigen-binding fragments thereof as described herein are present.

The disclosure provides antibodies or antigen-binding fragments thereof that can internalize into human brain cells (e.g., cortical microvascular endothelial cells) that the endocytosis rate is at least 50%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%. In some embodiments, the endocytosis rate of the antibodies or antigen-binding fragments thereof described herein is at least 1-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 50-fold, 100-fold, 500-fold, or 1000-fold as compared to that of an isotype control antibody.

In some embodiments, provided herein is an antibody or antigen-binding fragment thereof comprising a single heavy chain. In some embodiments, provided herein is an antibody or antigen-binding fragment thereof comprising a pair of heavy chains. In some embodiments, the heavy chain pair is linked by disulfide bonds. In some embodiments, the heavy chain pair comprises knob-in-hole modifications. In some embodiments, the heavy chain comprises a human IgG Fc domain. In some embodiments, the antibody or antigen-binding fragment thereof comprises in each heavy chain at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 VHH domains. In some embodiments, the VHH domains in each heavy chain specifically bind to the same epitope. In some embodiments, the VHH domains in each heavy chain specifically bind to different epitopes. In some embodiments, the VHH domains in each heavy chain bind to at least 1, 2, 3, 4, or 5 different epitopes.

In some embodiments, the antibody or antigen-binding fragment thereof is a bi-specific antibody, or tri-specific antibody. In some embodiments, the antibody or antigen-binding fragment thereof can specifically bind to at least 4, 5, or 6 antigens.

In some embodiments, the antibody (or antigen-binding fragment thereof) specifically binds to TFR1 with a dissociation rate (koff) of less than 0.1 s^{-1}, less than 0.01 s^{-1}, less than 0.001 s^{-1}, less than 0.0001 s^{-1}, or less than 0.00001 s^{-1}. In some embodiments, the dissociation rate (koff) is greater than 0.01 s^{-1}, greater than 0.001 s^{-1}, greater than 0.0001 s^{-1}, greater than 0.00001 s^{-1}, or greater than 0.000001 s^{-1}.

In some embodiments, kinetic association rates (kon) is greater than 1×10^2/Ms, greater than 1×10^3/Ms, greater than 1×10^4/Ms, greater than 1×10^5/Ms, or greater than 1×10^6/Ms. In some embodiments, kinetic association rates (kon) is less than 1×10^5/Ms, less than 1×10^6/Ms, or less than 1×10^7/Ms.

Affinities can be deduced from the quotient of the kinetic rate constants (KD=koff/kon). In some embodiments, KD is less than 1×10^{-6} M, less than 1×10^{-7} M, less than 1×10^{-8} M, less than 1×10^{-9} M, or less than 1×10^{-10} M. In some embodiments, the KD is less than 50 nM, 30 nM, 20 nM, 15 nM, 10 nM, 9 nM, 8 nM, 7 nM, 6 nM, 5 nM, 4 nM, 3 nM, 2 nM, or 1 nM. In some embodiments, KD is greater than 1 x 10^{-7} M, greater than 1 x 10^{-8} M, greater than 1 x 10^{-9} M, greater than 1 x 10^{-10} M, greater than 1 x 10^{-11} M, or greater than 1 x 10^{-12} M.

General techniques for measuring the affinity of an antibody for an antigen include, e.g., ELISA, RIA, and surface plasmon resonance (SPR). In some embodiments, the antibody binds to human TFR1, monkey TFR1, mouse TFR1, or chimeric TFR1. In some embodiments, the antibody does not bind to human TFR1, monkey TFR1, mouse TFR1, or chimeric TFR1.

In some embodiments, thermal stabilities are determined. The antibodies or antigen-binding fragments as described herein can have a Tm (melting temperature) greater than 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, or 95 °C. In some embodiments, Tm is less than 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, or 95 °C. The antibodies or antigen-binding fragments as described herein can have a Tagg (aggregation temperature, e.g., Tagg at 266 nm (Tagg266) or Tagg at 473 nm (Tagg473)) great than 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, or 95 °C. In some embodiments, Tagg is less than 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, or 95 °C.

In some embodiments, the Fc region is human IgG1, human IgG2, human IgG3, or human IgG4.

In some embodiments, the antibodies or antigen binding fragments thereof have a functional Fc region. In some embodiments, the antibodies or antigen binding fragments thereof comprise a human IgG1 Fc region. In some embodiments, the human IgG1 Fc region comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 75.

In some embodiments, the antibodies or antigen binding fragments do not have an Fc region. For example, the antibody (or antigen-binding fragment thereof) is a polypeptide comprising one or more VHH domains that are interconnected by linker peptides. In some embodiments, the antibody comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 VHH domain. In some embodiments, the VHH domains specifically bind to the same epitope. In some embodiments, the VHH domains bind to different epitopes. In some embodiments, the VHH domains bind to at least 1, 2, 3, 4, or 5 different epitopes.

In some embodiments, the antibodies or antigen binding fragments thereof do not have a functional Fc region. In some embodiments, the Fc region has LALA mutations (L234A and L235A mutations in EU numbering), or LALA-PG mutations (L234A, L235A, P329G mutations

in EU numbering). In some embodiments, the Fc region has a mutation at position 297 (e.g., N297A) according to EU numbering. In some embodiments, the mutated human IgG1 Fc region comprises an amino acid sequence that is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 76.

In some embodiments, after being administered to a subject after 6 hours, 12 hours, 24 hours, 36 hours, 48 hours, 60 hours, or 72 hours, concentration of the antibody or antigen binding fragment thereof described herein in brain (e.g., whole brain or parenchyma) can be greater than 10%, 20%, 30%, 40%, 50%, 60%, 70%, or 80% of its concentration immediately (e.g., 0.5 hour) after administration. In some embodiments, after being administered to a subject after 6 hours, 12 hours, 24 hours, 36 hours, 48 hours, 60 hours, or 72 hours, concentration of the antibody or antigen binding fragment thereof described herein in brain (e.g., whole brain or parenchyma) can be at least 1-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold, 60-fold, 70-fold, 80-fold, 90-fold, 100-fold, 200-fold, 300-fold, 400-fold, 500-fold, 1000-fold, 2000-fold, 5000-fold, or 10000-fold of the concentration of a control antibody (e.g., hIgG1 or JR141-N), or the concentration in serum of the subject.

Methods of making anti-TFR1 antibodies

Variants of the antibodies or antigen-binding fragments described herein can be prepared by introducing appropriate nucleotide changes into the DNA encoding a human, humanized, or chimeric antibody, or antigen-binding fragment thereof described herein, or by peptide synthesis. Such variants include, for example, deletions, insertions, or substitutions of residues within the amino acids sequences that make-up the antigen-binding site of the antibody or an antigen-binding domain. In a population of such variants, some antibodies or antigen-binding fragments will have increased affinity for the target protein, e.g., TFR1. Any combination of deletions, insertions, and/or combinations can be made to arrive at an antibody or antigen-binding fragment thereof that has increased binding affinity for the target. The amino acid changes introduced into the antibody or antigen-binding fragment can also alter or introduce new post-translational modifications into the antibody or antigen-binding fragment, such as changing (e.g., increasing or decreasing) the number of glycosylation sites, changing the type of glycosylation site (e.g., changing the amino acid sequence such that a different sugar is attached by enzymes present in a cell), or introducing new glycosylation sites. In some embodiments, the heavy-chain antibody or antigen-binding fragment thereof described herein is obtained by immunizing any of the genetically modified animals (e.g., mice that are homozygous for heavy chain *Mutant allele 3* genotype) described herein.

Humanized antibodies include antibodies having variable and constant regions derived from (or having the same amino acid sequence as those derived from) human germline immunoglobulin sequences of human immunoglobulin scaffold sequences. Humanized antibodies may include amino acid residues not encoded by human germline immunoglobulin sequences (e.g., mutations introduced by random or site-specific mutagenesis *in vitro* or by somatic mutation *in vivo*). Accordingly, "humanized" antibodies are chimeric antibodies wherein sequences from a non-human species are substituted by the corresponding human sequences.

Ordinarily, amino acid sequence variants of the human, humanized, or chimeric anti-TFR1 antibody will contain an amino acid sequence having at least 75%, 80%, 85%, 90%, 95%,

96%, 97%, 98%, or 99% percent identity with a sequence present in the VHH domain of the original antibody.

Identity or homology with respect to an original sequence is usually the percentage of amino acid residues present within the candidate sequence that are identical with a sequence present within the human, humanized, or chimeric anti-TFR1 antibody or fragment, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity.

Additional modifications to the anti-TFR1 antibodies or antigen-binding fragments can be made. For example, a cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have any increased half-life *in vitro* and/or *in vivo*. Homodimeric antibodies with increased half-life *in vitro* and/or *in vivo* can also be prepared using heterobifunctional cross-linkers as described, for example, in Wolff et al. Wolff et al. ("Monoclonal antibody homodimers: enhanced antitumor activity in nude mice." Cancer research 53.11 (1993): 2560-2565). Alternatively, an antibody can be engineered which has dual Fc regions.

In some embodiments, a covalent modification can be made to the anti-TFR1 antibody or antigen-binding fragment thereof. These covalent modifications can be made by chemical or enzymatic synthesis, or by enzymatic or chemical cleavage. Other types of covalent modifications of the antibody or antibody fragment are introduced into the molecule by reacting targeted amino acid residues of the antibody or fragment with an organic derivatization agent that is capable of reacting with selected side chains or the N- or C-terminal residues.

In some embodiments, antibody variants are provided having a carbohydrate structure that lacks fucose attached (directly or indirectly) to an Fc region. For example, the amount of fucose in such antibody composition may be from 1% to 80%, from 1% to 65%, from 5% to 65% or from 20% to 40%. The amount of fucose is determined by calculating the average amount of fucose within the sugar chain at Asn297, relative to the sum of all glycostructures attached to Asn 297 (e.g. complex, hybrid and high mannose structures) as measured by MALDI-TOF mass spectrometry, as described in WO 2008/077546, for example. Asn297 refers to the asparagine residue located at about position 297 in the Fc region (Eu numbering of Fc region residues; or position 314 in Kabat numbering); however, Asn297 may also be located about ±3 amino acids upstream or downstream of position 297, i.e., between positions 294 and 300, due to minor sequence variations in antibodies. Such fucosylation variants may have improved ADCC function. In some embodiments, to reduce glycan heterogeneity, the Fc region of the antibody can be further engineered to replace the Asparagine at position 297 with Alanine (N297A).

The present disclosure also provides recombinant vectors (e.g., an expression vectors) that include an isolated polynucleotide disclosed herein (e.g., a polynucleotide that encodes a polypeptide disclosed herein), host cells into which are introduced the recombinant vectors (i.e., such that the host cells contain the polynucleotide and/or a vector comprising the polynucleotide), and the production of recombinant antibody polypeptides or fragments thereof by recombinant techniques.

As used herein, a “vector” is any construct capable of delivering one or more polynucleotide(s) of interest to a host cell when the vector is introduced to the host cell. An

"expression vector" is capable of delivering and expressing the one or more polynucleotide(s) of interest as an encoded polypeptide in a host cell into which the expression vector has been introduced. Thus, in an expression vector, the polynucleotide of interest is positioned for expression in the vector by being operably linked with regulatory elements such as a promoter, enhancer, and/or a poly-A tail, either within the vector or in the genome of the host cell at or near or flanking the integration site of the polynucleotide of interest such that the polynucleotide of interest will be translated in the host cell introduced with the expression vector.

A vector can be introduced into the host cell by methods known in the art, e.g., electroporation, chemical transfection (e.g., DEAE-dextran), transformation, transfection, and infection and/or transduction (e.g., with recombinant virus). Thus, non-limiting examples of vectors include viral vectors (which can be used to generate recombinant virus), naked DNA or RNA, plasmids, cosmids, phage vectors, and DNA or RNA expression vectors associated with cationic condensing agents.

In some implementations, a polynucleotide disclosed herein (e.g., a polynucleotide that encodes a polypeptide disclosed herein) is introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus, or may use a replication defective virus. In the latter case, viral propagation generally will occur only in complementing virus packaging cells. Suitable systems are disclosed, for example, in Fisher-Hoch et al., 1989, Proc. Natl. Acad. Sci. USA 86:317-321; Flexner et al., 1989, Ann. N.Y. Acad Sci. 569:86-103; Flexner et al., 1990, Vaccine, 8:17-21; U.S. Pat. Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Pat. No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner-Biotechniques, 6:616-627, 1988; Rosenfeld et al., 1991, Science, 252:431-434; Kolls et al., 1994, Proc. Natl. Acad. Sci. USA, 91:215-219; Kass-Eisler et al., 1993, Proc. Natl. Acad. Sci. USA, 90:11498-11502; Guzman et al., 1993, Circulation, 88:2838-2848; and Guzman et al., 1993, Cir. Res., 73:1202-1207. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., 1993, Science, 259:1745-1749, and Cohen, 1993, Science, 259:1691-1692. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads that are efficiently transported into the cells.

For expression, the DNA insert comprising an antibody-encoding or polypeptide-encoding polynucleotide disclosed herein can be operatively linked to an appropriate promoter (e.g., a heterologous promoter), such as the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters are known to the skilled artisan. In some embodiments, the promoter is a cytomegalovirus (CMV) promoter. The expression constructs can further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs may include a translation initiating at the beginning and a termination codon (UAA, UGA, or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors can include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing in *E. coli* and other bacteria.

Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces*, and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, Bowes melanoma, and HK 293 cells; and plant cells. Appropriate culture mediums and conditions for the host cells described herein are known in the art.

Non-limiting vectors for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Non-limiting eukaryotic vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Non-limiting bacterial promoters suitable for use include the *E. coli* lacI and lacZ promoters, the T3 and T7 promoters, the gpt promoter, the lambda PR and PL promoters and the trp promoter. Suitable eukaryotic promoters include the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus (RSV), and metallothionein promoters, such as the mouse metallothionein-I promoter.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis *et al.*, Basic Methods In Molecular Biology (1986), which is incorporated herein by reference in its entirety.

Transcription of DNA encoding an antibody of the present disclosure by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at base pairs 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide (e.g., antibody) can be expressed in a modified form, such as a fusion protein (*e.g.*, a GST-fusion) or with a histidine-tag, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties can be added to the polypeptide to facilitate purification. Such regions can be removed prior to final preparation of the polypeptide. The

addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

The disclosure also provides a nucleic acid sequence that 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%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to any nucleotide sequence as described herein, and an amino acid sequence that 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%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identical to any amino acid sequence as described herein. In some embodiments, the disclosure relates to nucleotide sequences encoding any peptides that are described herein, or any amino acid sequences that are encoded by any nucleotide sequences as described herein. In some embodiments, the nucleic acid sequence is less than 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 150, 200, 250, 300, 350, 400, 500, or 600 nucleotides. In some embodiments, the amino acid sequence is less than 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, or 200 amino acid residues.

In some embodiments, the amino acid sequence (i) comprises an amino acid sequence; or (ii) consists of an amino acid sequence, wherein the amino acid sequence is any one of the sequences as described herein.

In some embodiments, the nucleic acid sequence (i) comprises a nucleic acid sequence; or (ii) consists of a nucleic acid sequence, wherein the nucleic acid sequence is any one of the sequences as described herein.

In some embodiments, the antibody or antigen-binding fragment thereof is expressed in yeast, insect cells, or mammalian cells (e.g., CHO cells).

Methods of treatment and diagnosis

The anti-TFR1 antibodies or antibody or antigen-binding fragments thereof of the present disclosure can be used for various therapeutic purposes. In one aspect, the disclosure provides methods for treating a brain disease (e.g., brain cancer, dementia, or Alzheimer's disease) in a subject, methods of identifying a subject having a brain disease (e.g., brain cancer, dementia, or Alzheimer's disease), , methods of reducing the risk of developing a brain disease, or methods of reducing the risk of developing an additional symptoms in a subject. In some embodiments, the treatment can halt, slow, retard, or inhibit progression of a brain disease (e.g., brain cancer, dementia, or Alzheimer's disease). In some embodiments, the treatment can result in the reduction of in the number, severity, and/or duration of one or more symptoms of a brain disease (e.g., brain cancer, dementia, or Alzheimer's disease) in a subject.

In one aspect, the disclosure features methods that include administering a therapeutically effective amount of an antibody or antigen-binding fragment thereof disclosed herein to a subject in need thereof (e.g., a subject having, or identified or diagnosed as having, a brain disease).

In one aspect, the disclosure features methods that carry a therapeutic agent to cross the blood brain barrier. In some embodiments, the antibodies or antigen-binding fragments thereof as described herein are linked to the therapeutic agent. In some embodiments, the therapeutic agent is an antibody, an antigen binding fragment thereof, a small molecule, or an antibody-drug conjugate.

In some embodiments, the compositions and methods disclosed herein can be used for treatment of patients at risk for a brain disease (e.g., brain cancer, dementia, or Alzheimer's disease). Patients with a brain disease (e.g., brain cancer, dementia, or Alzheimer's disease) can be identified with various methods known in the art.

In some embodiments, the brain disease is a brain cancer.

In one aspect, the disclosure is related to methods of decreasing the rate of tumor growth, including contacting a tumor cell with an effective amount of a composition including the antibody or antigen-binding fragment thereof, or the antibody-drug conjugate described herein. In one aspect, the disclosure is related to methods of killing a tumor cell, including contacting a tumor cell with an effective amount of a composition including the antibody or antigen-binding fragment thereof, or the antibody-drug conjugate described herein.

As used herein, by an "effective amount" is meant an amount or dosage sufficient to effect beneficial or desired results including halting, slowing, retarding, or inhibiting progression of a disease, e.g., a cancer. An effective amount will vary depending upon, e.g., an age and a body weight of a subject to which the antibody, antigen binding fragment, antibody-encoding polynucleotide, vector comprising the polynucleotide, and/or compositions thereof is to be administered, a severity of symptoms and a route of administration, and thus administration can be determined on an individual basis.

An effective amount can be administered in one or more administrations. By way of example, an effective amount of an antibody or an antigen binding fragment thereof is an amount sufficient to ameliorate, stop, stabilize, reverse, inhibit, slow and/or delay progression of a disease in a patient. As is understood in the art, an effective amount of an antibody or antigen binding fragment may vary, depending on, *inter alia*, patient history as well as other factors such as the type (and/or dosage) of antibody used.

Effective amounts and schedules for administering the antibodies, antibody-encoding polynucleotides, and/or compositions disclosed herein may be determined empirically, and making such determinations is within the skill in the art. Those skilled in the art will understand that the dosage that must be administered will vary depending on, for example, the mammal that will receive the antibodies, antibody-encoding polynucleotides, and/or compositions disclosed herein, the route of administration, the particular type of antibodies, antibody-encoding polynucleotides, antigen binding fragments, and/or compositions disclosed herein used and other drugs being administered to the mammal. Guidance in selecting appropriate doses for antibody or antigen binding fragment can be found in the literature on therapeutic uses of antibodies and antigen binding fragments, e.g., Handbook of Monoclonal Antibodies, Ferrone et al., eds., Noges Publications, Park Ridge, N.J., 1985, ch. 22 and pp. 303-357; Smith et al., Antibodies in Human Diagnosis and Therapy, Haber et al., eds., Raven Press, New York, 1977, pp. 365-389.

A typical daily dosage of an effective amount of an antibody is 0.01 mg/kg to 100 mg/kg. In some embodiments, the dosage can be less than 100 mg/kg, 10 mg/kg, 9 mg/kg, 8 mg/kg, 7 mg/kg, 6 mg/kg, 5 mg/kg, 4 mg/kg, 3 mg/kg, 2 mg/kg, 1 mg/kg, 0.5 mg/kg, or 0.1 mg/kg. In some embodiments, the dosage can be greater than 10 mg/kg, 9 mg/kg, 8 mg/kg, 7 mg/kg, 6 mg/kg, 5 mg/kg, 4 mg/kg, 3 mg/kg, 2 mg/kg, 1 mg/kg, 0.5 mg/kg, 0.1 mg/kg, 0.05 mg/kg, or 0.01 mg/kg. In some embodiments, the dosage is about 10 mg/kg, 9 mg/kg, 8 mg/kg, 7 mg/kg, 6

mg/kg, 5 mg/kg, 4 mg/kg, 3 mg/kg, 2 mg/kg, 1 mg/kg, 0.9 mg/kg, 0.8 mg/kg, 0.7 mg/kg, 0.6 mg/kg, 0.5 mg/kg, 0.4 mg/kg, 0.3 mg/kg, 0.2 mg/kg, or 0.1 mg/kg.

In any of the methods described herein, the at least one antibody, antigen-binding fragment thereof, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding fragments, or pharmaceutical compositions described herein) and, optionally, at least one additional therapeutic agent can be administered to the subject at least once a week (e.g., once a week, twice a week, three times a week, four times a week, once a day, twice a day, or three times a day). In some embodiments, at least two different antibodies and/or antigen-binding fragments are administered in the same composition (e.g., a liquid composition). In some embodiments, at least one antibody or antigen-binding fragment and at least one additional therapeutic agent are administered in the same composition (e.g., a liquid composition). In some embodiments, the at least one antibody or antigen-binding fragment and the at least one additional therapeutic agent are administered in two different compositions (e.g., a liquid composition containing at least one antibody or antigen-binding fragment and a solid oral composition containing at least one additional therapeutic agent). In some embodiments, the at least one additional therapeutic agent is administered as a pill, tablet, or capsule. In some embodiments, the at least one additional therapeutic agent is administered in a sustained-release oral formulation.

In some embodiments, the one or more additional therapeutic agents can be administered to the subject prior to, or after administering the at least one antibody, antigen-binding antibody fragment, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding antibody fragments, or pharmaceutical compositions described herein). In some embodiments, the one or more additional therapeutic agents and the at least one antibody, antigen-binding antibody fragment, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding antibody fragments, or pharmaceutical compositions described herein) are administered to the subject such that there is an overlap in the bioactive period of the one or more additional therapeutic agents and the at least one antibody or antigen-binding fragment (e.g., any of the antibodies or antigen-binding fragments described herein) in the subject.

In some embodiments, the subject can be administered the at least one antibody, antigen-binding antibody fragment, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding antibody fragments, or pharmaceutical compositions described herein) over an extended period of time (e.g., over a period of at least 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 12 months, 1 year, 2 years, 3 years, 4 years, or 5 years). A skilled medical professional may determine the length of the treatment period using any of the methods described herein for diagnosing or following the effectiveness of treatment (e.g., the observation of at least one symptom of the disease). As described herein, a skilled medical professional can also change the identity and number (e.g., increase or decrease) of antibodies or antigen-binding antibody fragments (and/or one or more additional therapeutic agents) administered to the subject and can also adjust (e.g., increase or decrease) the dosage or frequency of administration of at least one antibody or antigen-binding antibody fragment (and/or one or more additional therapeutic agents) to the subject based on an assessment of the effectiveness of the treatment (e.g., using any of the methods described herein and known in the art).

In some embodiments, one or more additional therapeutic agents can be administered to the subject. The additional therapeutic agent can comprise one or more inhibitors selected from the group consisting of an inhibitor of B-Raf, an EGFR inhibitor, an inhibitor of a MEK, an inhibitor of ERK, an inhibitor of K-Ras, an inhibitor of c-Met, an inhibitor of anaplastic lymphoma kinase (ALK), an inhibitor of a phosphatidylinositol 3-kinase (PI3K), an inhibitor of an Akt, an inhibitor of mTOR, a dual PI3K/mTOR inhibitor, an inhibitor of Bruton's tyrosine kinase (BTK), and an inhibitor of Isocitrate dehydrogenase 1 (IDH1) and/or Isocitrate dehydrogenase 2 (IDH2).

In some embodiments, the additional therapeutic agent can comprise one or more inhibitors selected from the group consisting of an inhibitor of HER3, an inhibitor of LSD1, an inhibitor of MDM2, an inhibitor of BCL2, an inhibitor of CHK1, an inhibitor of activated hedgehog signaling pathway, and an agent that selectively degrades the estrogen receptor.

In some embodiments, the additional therapeutic agent can comprise one or more therapeutic agents selected from the group consisting of Trabectedin, nab-paclitaxel, Trebananib, Pazopanib, Cediranib, Palbociclib, everolimus, fluoropyrimidine, IFL, regorafenib, Reolysin, Alimta, Zykadia, Sutent, temsirolimus, axitinib, everolimus, sorafenib, Votrient, Pazopanib, IMA-901, AGS-003, cabozantinib, Vinflunine, an Hsp90 inhibitor, Ad-GM-CSF, Temazolomide, IL-2, IFNa, vinblastine, Thalomid, dacarbazine, cyclophosphamide, lenalidomide, azacytidine, lenalidomide, bortezomid, amrubicine, carfilzomib, pralatrexate, and enzastaurin.

In some embodiments, the additional therapeutic agent can comprise one or more therapeutic agents selected from the group consisting of an adjuvant, a TLR agonist, tumor necrosis factor (TNF) alpha, IL-1, HMGB1, an IL-10 antagonist, an IL-4 antagonist, an IL-13 antagonist, an IL-17 antagonist, an HVEM antagonist, an ICOS agonist, a treatment targeting CX3CL1, a treatment targeting CXCL9, a treatment targeting CXCL10, a treatment targeting CCL5, an LFA-1 agonist, an ICAM1 agonist, and a Selectin agonist.

In some embodiments, carboplatin, nab-paclitaxel, paclitaxel, cisplatin, pemetrexed, gemcitabine, FOLFOX, or FOLFIRI are administered to the subject.

In some embodiments, the additional therapeutic agent is an anti-PD1 antibody, an anti-PD-L1 antibody, an anti-LAG-3 antibody, an anti-TIGIT antibody, an anti-BTLA antibody, an anti-CTLA-4 antibody, or an anti-GITR antibody.

Pharmaceutical Compositions and Routes of Administration

Also provided herein are pharmaceutical compositions that contain at least one (e.g., one, two, three, or four) of the antibodies or antigen-binding fragments described herein. Two or more (e.g., two, three, or four) of any of the antibodies or antigen-binding fragments described herein can be present in a pharmaceutical composition in any combination. The pharmaceutical compositions may be formulated in any manner known in the art.

Pharmaceutical compositions are formulated to be compatible with their intended route of administration (e.g., intravenous, intraarterial, intramuscular, intradermal, subcutaneous, or intraperitoneal). The compositions can include a sterile diluent (e.g., sterile water or saline), a fixed oil, polyethylene glycol, glycerine, propylene glycol or other synthetic solvents, antibacterial or antifungal agents, such as benzyl alcohol or methyl parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like, antioxidants, such as ascorbic acid or sodium

bisulfite, chelating agents, such as ethylenediaminetetraacetic acid, buffers, such as acetates, citrates, or phosphates, and isotonic agents, such as sugars (e.g., dextrose), polyalcohols (e.g., mannitol or sorbitol), or salts (e.g., sodium chloride), or any combination thereof. Liposomal suspensions can also be used as pharmaceutically acceptable carriers (see, e.g., U.S. Patent No. 4,522,811). Preparations of the compositions can be formulated and enclosed in ampules, disposable syringes, or multiple dose vials. Where required (as in, for example, injectable formulations), proper fluidity can be maintained by, for example, the use of a coating, such as lecithin, or a surfactant. Absorption of the antibody or antigen-binding fragment thereof can be prolonged by including an agent that delays absorption (e.g., aluminum monostearate and gelatin). Alternatively, controlled release can be achieved by implants and microencapsulated delivery systems, which can include biodegradable, biocompatible polymers (e.g., ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid; Alza Corporation and Nova Pharmaceutical, Inc.).

Compositions containing one or more of any of the antibodies or antigen-binding fragments described herein can be formulated for parenteral (e.g., intravenous, intraarterial, intramuscular, intradermal, subcutaneous, or intraperitoneal) administration in dosage unit form (i.e., physically discrete units containing a predetermined quantity of active compound for ease of administration and uniformity of dosage).

Toxicity and therapeutic efficacy of compositions can be determined by standard pharmaceutical procedures in cell cultures or experimental animals (e.g., monkeys). One can, for example, determine the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population): the therapeutic index being the ratio of LD50:ED50. Agents that exhibit high therapeutic indices are preferred. Where an agent exhibits an undesirable side effect, care should be taken to minimize potential damage (i.e., reduce unwanted side effects). Toxicity and therapeutic efficacy can be determined by other standard pharmaceutical procedures.

Data obtained from cell culture assays and animal studies can be used in formulating an appropriate dosage of any given agent for use in a subject (e.g., a human). A therapeutically effective amount of the one or more (e.g., one, two, three, or four) antibodies or antigen-binding fragments thereof (e.g., any of the antibodies or antibody fragments described herein) will be an amount that treats the disease in a subject in a subject, or a subject identified as being at risk of developing the disease, decreases the severity, frequency, and/or duration of one or more symptoms of a disease in a subject (e.g., a human). The effectiveness and dosing of any of the antibodies or antigen-binding fragments described herein can be determined by a health care professional or veterinary professional using methods known in the art, as well as by the observation of one or more symptoms of disease in a subject (e.g., a human). Certain factors may influence the dosage and timing required to effectively treat a subject (e.g., the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and the presence of other diseases).

Exemplary doses include milligram or microgram amounts of any of the antibodies or antigen-binding fragments described herein per kilogram of the subject's weight (e.g., about 1 μg/kg to about 500 mg/kg; about 100 μg/kg to about 500 mg/kg; about 100 μg/kg to about 50 mg/kg; about 10 μg/kg to about 5 mg/kg; about 10 μg/kg to about 0.5 mg/kg; or about 1 μg/kg to

about 50 μg/kg). While these doses cover a broad range, one of ordinary skill in the art will understand that therapeutic agents, including antibodies and antigen-binding fragments thereof, vary in their potency, and effective amounts can be determined by methods known in the art. Typically, relatively low doses are administered at first, and the attending health care professional or veterinary professional (in the case of therapeutic application) or a researcher (when still working at the development stage) can subsequently and gradually increase the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, and the half-life of the antibody or antibody fragment *in vivo*.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration. The disclosure also provides methods of manufacturing the antibodies or antigen binding fragments thereof for various uses as described herein.

ADDITIONAL EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the claims. Other aspects, advantages, and modifications are within the scope of the claims. Also provided are additional embodiments.

Embodiment 1 is a genetically modified rodent comprising a modified immunoglobulin heavy chain locus, wherein the modified immunoglobulin heavy chain locus comprises an IgG constant region gene, wherein the IgG constant region gene encodes an IgG heavy chain constant region lacking a CH1 domain, wherein the genetically modified rodent expresses a heavy-chain antibody.

Embodiment 2 is the genetically modified rodent of embodiment 1, wherein the rodent comprises exactly one IgG constant region gene.

Embodiment 3 is the genetically modified rodent of embodiment 1 or 2, wherein the IgG heavy chain constant region gene is IGHG1.

Embodiment 4 is the genetically modified rodent of any one of embodiments 1-3, wherein the IgG heavy chain constant region comprises or consists of a CH2 domain and a CH3 domain, and optionally a hinge region.

Embodiment 5 is a genetically modified rodent whose genome comprises a germline genetic modification comprising a deletion of IGHG3, IGHG2b, and IGHG2c genes and a deletion of the CH1 exon of IGHG1 gene at a rodent immunoglobulin heavy chain gene locus.

Embodiment 6 is the rodent of embodiment 5, wherein the germline genetic modification further comprises a deletion of rodent IGHE gene at the rodent immunoglobulin heavy chain gene locus.

Embodiment 7 is the rodent of embodiment 5 or 6, wherein the genetic modification further comprises a deletion of rodent Sγ2b, Sγ2c, and Sε switch regions at the rodent immunoglobulin heavy chain gene locus.

Embodiment 8 is the rodent of any one of embodiments 5-7, wherein the modified immunoglobulin heavy chain gene locus comprises a modified IGHG1 gene lacking a sequence encoding a CH1 domain, wherein the modified IGHG1 gene comprises a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 1.

Embodiment 9 is the rodent of any one of embodiments 5-8, wherein the genetic modification further comprises a deletion of rodent Sγ3 switch region at the rodent immunoglobulin heavy chain gene locus.

Embodiment 10 is the rodent of any one of embodiments 5-9, wherein the rodent's genome comprises rodent Sμ, Sγ1, Sα switch regions, the modified rodent IGHG1 gene lacking a sequence encoding a CH1 domain, and rodent IGHM, IGHδ, IGHA genes.

Embodiment 11 is the rodent of any one of embodiments 5-9, wherein the genetic modification further comprises a deletion of rodent IGHM and IGHδ genes at the rodent immunoglobulin heavy chain gene locus.

Embodiment 12 is the rodent of any one of embodiments 5-9 and 11, wherein the rodent genome comprises rodent Sμ, Sγ1, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and the rodent IGHA gene.

Embodiment 13 is the rodent of embodiment 12, wherein the Sμ and Sγ1 switch regions are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 8.

Embodiment 14 is the rodent of any one of embodiments 5-9, wherein the genetic modification further comprises a deletion of the CH1 coding sequence of IGHM gene at the rodent immunoglobulin heavy chain gene locus.

Embodiment 15 is the rodent of any one of embodiments 5-9 and 14, wherein the rodent's genome comprises rodent Sμ, Sγ1, Sα switch regions, a modified rodent IGHM gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and rodent IGHδ, IGHA genes.

Embodiment 16 is the rodent of embodiment 15, wherein the Sμ switch region and the modified IGHM gene are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 10.

Embodiment 17 is the rodent of any one of embodiments 5-9, wherein the genetic modification further comprises a deletion of the CH1 exon of IGHM gene and a deletion of IGHδ gene at the rodent immunoglobulin heavy chain gene locus.

Embodiment 18 is the rodent of any one of embodiments 5-9 and 17, wherein the rodent's genome comprises rodent Sμ, Sγ1, Sα switch regions, a modified IGHM gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and a rodent IGHA gene.

Embodiment 19 is the rodent of any one of embodiments 5-9, wherein the genetic modification further comprises a deletion of the CH1 exon of IGHM gene and a deletion of the CH1 coding sequence of IGHδ gene at the rodent immunoglobulin heavy chain gene locus.

Embodiment 20 is the rodent of any one of embodiments 5-9 and 19, wherein the rodent's genome comprises rodent Sμ, Sγ1, Sα switch regions, a modified IGHM gene lacking a sequence encoding a CH1 domain, a modified IGHδ gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and a rodent IGHA gene.

Embodiment 21 is the rodent of embodiment 19 or 20, wherein the modified IGHM gene are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 10, and the modified IGHδ gene comprises a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 41.

Embodiment 22 is the rodent of any one of embodiments 14-21, wherein the modified IGHM gene comprises a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 13.

Embodiment 23 is the rodent of any one of embodiments 5-8, wherein the genetic modification further comprises a deletion of rodent Sγ1 switch region at the rodent immunoglobulin heavy chain gene locus.

Embodiment 24 is the rodent of any one of embodiments 5-8 and 23, wherein the rodent's genome comprises rodent Sμ, Sγ3, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and rodent IGHM, IGHδ, IGHA genes.

Embodiment 25 is the rodent of any one of embodiments 5-8 and 23, wherein the genetic modification further comprises a deletion of rodent Sγ3 switch region at the rodent immunoglobulin heavy chain gene locus.

Embodiment 26 is the rodent of any one of embodiments 5-8, 23 and 25, wherein the genetic modification further comprises a deletion of rodent IGHM and IGHδ genes at the rodent immunoglobulin heavy chain gene locus.

Embodiment 27 is the rodent of any one of embodiments 5-8, 23, 25 or 26, wherein the rodent's genome comprises rodent Sμ, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and the rodent IGHA gene.

Embodiment 28 is the rodent of embodiment 27, wherein the Sμ switch region and the modified IGHG1 gene are linked with a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 9.

Embodiment 29 is the rodent of any one of embodiments 5-10, 23 and 24, wherein the modified genome comprises a functional IGHM gene.

EXAMPLES

The invention is further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLE 1: Overview

Immunoglobulin heavy chain locus of non-human animals was modified by gene editing. For example, in order to obtain mice that can express heavy-chain antibodies, the immunoglobulin heavy chain constant region locus within mouse chromosome 12 was modified. The genetically-modified mice can express heavy-chain antibodies as shown in **FIG. 1**. Furthermore, all endogenous VDJ sequences in the heavy chain variable region locus on chromosome 12 was replaced by the human VDJ sequences, so that the variable regions in the heavy-chain antibodies expressed by the mice had fully humanized sequences. In some experiments, the genetic modification on immunoglobulin heavy chain constant region locus was performed on mice with human VDJ sequences.

EXAMPLE 2: Modification of mouse immunoglobulin heavy chain constant region locus

As shown in **FIG. 2**, the mouse (C57BL/6) immunoglobulin constant region genes include (as shown in the following order): immunoglobulin heavy constant mu (IGHM, or Cμ), immunoglobulin heavy constant delta (IGHδ, or Cδ), immunoglobulin heavy constant gamma 3 (IGHG3, or Cγ3), immunoglobulin heavy constant gamma 1 (IGHG1, or Cγ1), immunoglobulin heavy constant gamma 2b (IGHG2b, or Cγ2b), immunoglobulin heavy constant gamma 2c (IGHG2c, or Cγ2c), immunoglobulin heavy constant epsilon (IGHE, or Cε), and immunoglobulin heavy constant alpha (IGHA, or Cα) genes. As shown in **FIG. 3**, switch regions (e.g., Sμ, Sγ3, Sγ1, Sγ2b, Sγ2c, Sε, and Sα) and their respective promoters are located upstream of the corresponding constant region genes. Particularly, the Cγ1 gene includes sequences that encode, from N-terminus to C-terminus, CH1, H (hinge), CH2, CH3, M1 and M2 regions of IgG1.

The mouse immunoglobulin heavy chain constant region locus was modified by various methods discussed as follows. The experiments were performed in mice with fully humanized VDJ sequences. Details of the VDJ region humanized mice can be found, e.g., in WO2020169022A1 and US20200390073A1; each of which is incorporated herein by reference in its entirety.

In order to obtain mice expressing only heavy-chain antibodies, the Cγ3, Cγ1, Cγ2b, Cγ2c, and Cε gene loci were modified to only keep a truncated Cγ1 sequence without the CH1 coding region (CγlΔCH1), to obtain a mutated allele shown as *Mutant allele 1* or *Mutant allele 1'* in **FIG. 4A**. On the basis of *Mutant allele 1*, different modifications were made to the switch region (e.g., Sμ and Sγ1), Cμ and/or Cδ sequences. For example, the entire Cμ and Cδ sequences were knocked out, and the resultant alleles are shown as *Mutant allele 2* and *Mutant allele 2'* (further missing Sγ1). Alternatively, the region being knocked out was selected from: the CH1 coding region of Cμ (with resultant allele shown as *Mutant allele 3*); the CH1 coding region of Cμ and the entire Cδ (with resultant allele shown as *Mutant allele 4*); and both the CH1 coding region of Cμ and the CH1 coding region of Cδ (with resultant alleles shown as *Mutant allele 5*).

EXAMPLE 3: Construction of targeting vectors

Targeting vectors, e.g., V1, V2, V3, V4, V5, V6, and V7 were used in the modifications.

Vector V1

Vector V1 included the CγlΔCH1 knock-in sequence (SEQ ID NO: 1) as shown in **FIG. 5**. The CγlΔCH1 sequence did not have the coding sequence of CH1. The CH1 coding region was either deleted or replaced with a Neo cassette. The Neo cassette includes a Neo gene sequence flanked by two Frt (or LoxP) sequences. Flp transgenic mice were bred with the mice carrying the Neo cassette to remove the cassette.

The targeting vector V1 included the following feature sequences in a 5' to 3' order: an upstream homologous arm (5' homologous arm), the mouse Sγ1 promoter sequence, the mouse Sγ1, the CγlΔCH1 sequence, and a downstream homologous arm (3' homologous arm). The targeting vector can further include an antibiotic resistance gene for positive clone screening (e.g., neomycin phosphotransferase gene, or Neo), and two Frt recombination sites flanking the

antibiotic resistance gene. In addition, a coding gene with a negative selectable marker (a gene encoding diphtheria toxin A subunit (DTA)) can also be inserted to the targeting vector.

A 16076 bp sequence comprising the mouse Sγ1 promoter, the mouse Sγ1, and the Cγ1ΔCH1 knock-in sequence (SEQ ID NO: 1) was cloned from mouse bacterial artificial chromosome (RP23-38K22 or RP23-265P18). The following two sets of primer pairs were used to obtain a sequence comprising mouse Sγ1 promoter, the mouse Sγ1, and the Cγ1ΔCH1 knock-in sequence for the construction of the V1 vector. The sequence amplified from V1-F1 and V1-R1 and the sequence amplified from V1-F2 and V1-R2 were then ligated.

V1-F1 (SEQ ID NO: 2): 5'-GTGGTTCTGGCTACAAGATAGAGCTCTGTCAATGATGTTTGCAGAGACTACA-3'

V1-R1 (SEQ ID NO: 3): 5'-CTCCCTATACGTCCTCTCACCTACAAGAAAAAGTATATGTGATTACACTGTCAGACAG-3'

V1-F2 (SEQ ID NO: 4): 5'-GTGTAATCACATATACTTTTTCTTGTAGGTGAGAGGACGTATAGGGAGGAGGGGTTC-3'

V1-R2 (SEQ ID NO: 5): 5'-CGTCTAGTCCTTGCCCACGTGTCGACCCCATAGGGAGGACAGACTGAGG-3'

As shown in **FIG. 6**, the V1 vector was used to replace a 100883 bp sequence (nucleic acids 113232142 to 113333024 of NCBI Reference Sequence NC_000078.7) of the mouse heavy chain constant region locus (from Sγ3 to Cε) in a single step.

Vector V2

As shown in **FIG. 7**, the V2 vector was used to replace a 92859 bp sequence (nucleic acids 113232142 to 113325000 of NCBI Reference Sequence NC_000078.7) of the mouse heavy chain constant region (spanning from Cγ3 to Cε) in a single step.

The method also included using a pair of primers to clone from mouse bacterial artificial chromosome (BAC) to obtain a knock-in sequence comprising the Cγ1ΔCH1 knock-in sequence (SEQ ID NO: 1). The following primers were used:

V2-F1 (SEQ ID NO: 6):
5'-TCTGAACTACTTCGTCGACGTGAGAGGACGTATAGGGAGGAGGG-3'

V2-R1 (SEQ ID NO: 7):
5'-CACGTGGATCCGCGGCCGCCCATAGGGAGGACAGACTGAGGAC-3'

Vector V3

As shown in **FIG. 8**, the V3 vector was used to knock out a total of 16434 bp nucleotides from the chromosome including Cμ and Cδ. The linkage sequence between Sμ and Sγ1 in the recombined *Mutant allele 2* is shown in SEQ ID NO: 8.

Vector V4

As shown in **FIG. 9**, the targeting vector V4 included the following feature sequences in a 5' to 3' order: an upstream homologous arm (5' homologous arm), a portion of the mouse Sμ, and a downstream homologous arm (3' homologous arm). The targeting vector further included

an antibiotic resistance gene for positive clone screening (neomycin phosphotransferase gene, or Neo), and two Frt recombination sites flanking the antibiotic resistance gene. In addition, a coding gene with a negative selectable marker (a gene encoding diphtheria toxin A subunit (DTA)) was also inserted to the targeting vector. Specifically, Sμ was directly linked to Cγ1ΔCH1' in the recombined *Mutant allele 2'* (i.e., there was no other switch region or immunoglobulin gene sequences in between). The linkage sequence is shown in SEQ ID NO: 9.

Vector V5

As shown in **FIG. 10**, the V5 vector was used to knock out the CH1 coding sequence in Cμ. Upon deletion of the CH1 coding sequence in Cμ, the linkage sequence is shown in SEQ ID NO: 10.

Vector V6

As shown in **FIG. 11**, the targeting vector V6 included the following feature sequences in a 5' to 3' order: an upstream homologous arm (5' homologous arm), a portion of the mouse Sμ region, a mouse CμΔCH1 sequence and a downstream homologous arm (3' homologous arm). The targeting vector further included an antibiotic resistance gene for positive clone screening (neomycin phosphotransferase gene, or Neo), and two Frt recombination sites flanking the antibiotic resistance gene. In addition, a coding gene with a negative selectable marker (a gene encoding diphtheria toxin A subunit (DTA)) was also inserted to the targeting vector.

A sequence including the mouse CμΔCH1 sequence were cloned from mouse somatic cells. The following primer pair was used to obtain a sequence including the mouse CμΔCH1 (or CμΔCH1 knock-in sequence; SEQ ID NO: 13) for the construction of the V6 vector.

V6-F1 (SEQ ID NO: 11):
5'-ATCCCTCTCTGGTCCTAACCAAACCCTCCCAGCAGGGGTG-3'
V6-R1 (SEQ ID NO: 12):
5'-TTGACCCATCTCAGTTTACATGGTGAATGACTACAATATATCTGGAATTTGG-3'

Vector V7

As shown in **FIG. 12**, the targeting vector V7 included the following feature sequences in a 5' to 3' order: an upstream homologous arm (5' homologous arm), a CμΔCH1 sequence, a CδΔCH1 sequence (or CδΔCH1 knock-in sequence; SEQ ID NO: 41) and a downstream homologous arm (3' homologous arm). The targeting vector further included an antibiotic resistance gene for positive clone screening (neomycin phosphotransferase gene, or Neo), and two Frt recombination sites flanking the antibiotic resistance gene. In addition, a coding gene with a negative selectable marker (a gene encoding diphtheria toxin A subunit (DTA)) was also inserted to the targeting vector.

A sequence including the mouse Sμ switch region, the CμΔCH1 sequence and the CδΔCH1 sequence were cloned from mouse genome. The following primer pair was used to obtain a knock-in sequence including the mouse CμΔCH1 and nucleotides upstream of CδΔCH1 for the construction of the V7 vector.

V7-F1 (SEQ ID NO: 14):
5'-ATCCCTCTCTGGTCCTAACCAAACCCTCCCAGCAGGGGTG -3'

V7-R1 (SEQ ID NO: 15):
5'-TTCTGCATGGTCCAGGGATTGATCAGACAGATAGTGAAGTTCTGAGGACA-3'

EXAMPLE 4: Verification of genetic modifications

Mutant Allele 1

PCR and Southern Blot were used to detect the genotype of *Mutant Allele 1*. First, positive clone cells were identified by PCR using two sets of primer pairs L-GT-F1/L-GT-R1 and R-GT-F2/R-GT-R2. Exemplary results are shown in **FIGS. 13A-13B**, respectively. Sequences of the primers are shown in the table below.

Table 7

Primer	Sequence (5' to 3')	Fragment size (bp)
L-GT-F1	TAGGAAGTGGCTCAGAAGTCTGCCT (SEQ ID NO: 16)	Mut: 10653 bp
L-GT-R1	TCTCCAGCCCCATTTATCCCAGCTA (SEQ ID NO: 17)	
R-GT-F2	TGGGATAGCTATGTGGGGAGACCAG (SEQ ID NO: 18)	Mut: 10526 bp
R-GT-R2	AGTTCTGTCATCAGGCCAACTACGC (SEQ ID NO: 19)	

Next, the positive clones were verified by Southern Blot (digested with BclI, ScaI, XmnI, and BglII, respectively, followed by hybridization with 4 corresponding probes) to screen out correct positive clone cells. The Southern Blot detection strategy (including restriction enzymes, probes, and target fragment sizes) and the probe primers are shown in the tables below.

Table 8

Restriction enzyme	Probe	WT size	Targeted size
BclI	LR Probe	6.1 kb	18.4 kb
ScaI	3' Probe	9.7 kb	12.8 kb
XmnI	A Probe	9.7 kb	16.6 kb
BglII	5' Probe	7.2 kb	4.3 kb

Table 9

Probe	Sequence (5' -3')
LR Probe	TCCTTGCCCCTCAGTCAGTCTTGA (SEQ ID NO: 20)
	TCCTCCATGGCTTGGTGCCTTCT (SEQ ID NO: 21)
3' Probe	GCCAGCTCTTGAGGACCTGCTCC (SEQ ID NO: 22)
	TGGCTCAGACAAGGAGAGCAAGACC (SEQ ID NO: 23)
A Probe	ATTAGCCTGCCTAAACCAACCAGGC (SEQ ID NO: 24)
	TCAATACCTGGGCCACCCTCTAACA (SEQ ID NO: 25)
5'Probe	GAAGTCTGCCTGACCCTCTACATG (SEQ ID NO: 26)
	CACAGATGTGTAATACATGTACTTGC (SEQ ID NO: 27)

Exemplary detection results are shown in **FIGS. 14A-14D**. According to the PCR and Southern Blot results, mice numbered F1-012, F1-017, F1-018 and F1-019 were identified as

Mutant Allele 1 positive heterozygous mice. No random insertion in *Mutant Allele 1* was detected.

Mutant Allele 2

Primers DE-F1 and DE-R1 were used to confirm knock-out of the sequence from Cμ to Cδ in *Mutant Allele 2*. The detection result is shown in **FIG. 15**. Mice numbered F1-1, F1-2, F1-3, F1-4, F1-5, F1-6, F1-7 and F1-8 were identified as positive heterozygous mice. Sequences of the primers are shown in the table below.

Table 10

Primer	Sequence (5' to 3')	Fragment size (bp)
DE-F1	TGTGTGCTGATACAGAAGCCTGAGC (SEQ ID NO: 28)	Mut: ~792 bp
DE-R1	TCCCATCACCCTGTGTTATCATCTGT (SEQ ID NO: 29)	

Mutant Allele 2'

Primers GT-Mut-F, GT-Mut-R and GT-WT-R were used to confirm knock-out of the sequence from Cμ to Sγ1 in *Mutant Allele 2'*. The detection result is shown in **FIG. 16**. Mice numbered F1-2, and F1-4 were identified as positive heterozygous mice. Sequences of the primers are shown in the table below.

Table 11

Primer	Sequence (5' to 3')	Fragment size (bp)
GT-Mut-F	ACCATATCTAGCCCTGCCTGCCTTA (SEQ ID NO: 30)	WT: 330 bp Mut: 590 bp
GT-Mut-R	TCAATACCTGGGCCACCCTCTAACA (SEQ ID NO: 31)	
GT-WT-R	ACTTGCCCCCTGTCCTCAGTG (SEQ ID NO: 32)	

Mutant Allele 3

Primers GT-3F and GT-3R were used to confirm knock-out of the CH1 coding sequence of Cμ in *Mutant Allele 3*. The detection result is shown in **FIG. 17**. Mice numbered F1-2, F1-3 and F1-6 were identified as positive heterozygous mice. Sequences of the primers are shown in the table below.

Table 12

Primer	Sequence (5' to 3')	Fragment size (bp)
GT-3F	GTGGCTAGAAGGCAGCTCCA (SEQ ID NO: 33)	Mut: 361 bp WT: 684 bp
GT-3R	CAGAGAAGCCATCCCGTGGT (SEQ ID NO: 34)	

Mutant Allele 4

Primers Mut-F and Mut-R were used to confirm the sequence of CμΔCH1 in *Mutant Allele 4*. Primers F4 and R4 were used to confirm the absence of Cδ in *Mutant Allele 4*. The detection results are shown in **FIGS. 18A-18B**, respectively. Mice numbered F1-1, F1-2, F1-3, F1-4, F1-5, F1-6, F1-7, F1-8, F1-9 and F1-10 were identified as positive heterozygous mice. Sequences of the primers are shown in the table below.

Table 13

Primer	Sequence (5' to 3')	Fragment size (bp)
Mut-F	TACCATATCTAGCCCTGCCTGCCTT (SEQ ID NO: 35)	Mut: 342 bp WT: 665 bp
Mut-R	TTCGTGGCCTCGCAGATGAGTTTAG (SEQ ID NO: 36)	
F4	GCCTCTTTGGGCTCATATCCTGGAC (SEQ ID NO: 39)	Mut: 5594 bp
R4	GAGTGTGTCACATGACTCCACTGGG (SEQ ID NO: 40)	

Mutant Allele 5

Primers Mut-F and Mut-R were used to confirm the sequence of CμΔCH1, and primers F3 and R3 were used to confirm the sequence of CδΔCH1 in *Mutant Allele 5*. The detection results are shown in **FIGS. 19A-19B**, respectively. Mice numbered F1-1, F1-2, F1-3, F1-4, F1-5, F1-6, F1-7 and F1-8 were identified as positive heterozygous mice. Sequences of the primers are shown in the table below.

Table 14

Primer	Sequence (5' to 3')	Fragment size (bp)
Mut-F	TACCATATCTAGCCCTGCCTGCCTT (SEQ ID NO: 35)	Mut: 342 bp WT: 665 bp
Mut-R	TTCGTGGCCTCGCAGATGAGTTTAG (SEQ ID NO: 36)	
F3	CAGACCTCACCTTTGTCCCCAGTTC (SEQ ID NO: 37)	Mut: 323 bp WT: 611 bp
R3	GGTACTGCTATGGAAGGGACCATGC (SEQ ID NO: 38)	

EXAMPLE 5: Modification of mouse immunoglobulin heavy chain variable region locus

Experiments were performed to introduce human immunoglobulin genes into the mouse genome to produce mice expressing humanized antibodies. **FIG. 20** shows the methods of making the humanized mice. The methods first involve modifying the human immunoglobulin region on the human chromosome. The modified human chromosomes were then introduced into the mouse recipient cell.

Sequences of the mouse immunoglobulin variable region locus were replaced with sequences of the human immunoglobulin variable region locus by direct replacement (e.g., homologous recombination, or Cre-mediated recombination). In some cases, the human immunoglobulin variable region genes can be introduced into the mouse genome by a stepwise approach. Then, the recipient cells were screened for the correct replacement. The cells were then injected to blastocysts to prepare chimeric mice. Subsequent breeding was performed to obtain mice containing human or humanized immunoglobulin variable region locus sequences.

The immunoglobulin heavy chain locus is located on mouse chromosome 12. Two recombination sites were introduced on both sides of the immunoglobulin heavy chain variable region locus.

Experiments were also performed to generate a modified human chromosome. Two recombination sites were introduced on both sides of the variable region of the immunoglobulin heavy chain locus. The modified human chromosome was then introduced into mouse cell. The cells were then screened, and those containing only one human chromosome were selected. Cre recombinase then mediated the replacement of V, D, J regions on mouse chromosome with the V, D, J regions on human chromosome (**FIG. 21**).

The positive clone cells were injected into the blastocysts of BALB/c mice by microinjection. The embryo microinjection was carried out according to the method described, e.g., in A. Nagy, et al., "Manipulating the Mouse Embryo: A Laboratory Manual (Third Edition)," Cold Spring Harbor Laboratory Press, 2003. The injected fertilized eggs were then transferred to a culture medium for a short time culture, and then transplanted into the oviduct of the recipient mouse to produce the genetically-modified humanized mice (F0 generation). The mice were then bred with mice having C57BL/6 background. PCR analysis was performed on the DNA obtained from the tail of the mice. The mice were further bred with mice with BALB/c background several times (e.g., at least 5 times) to obtain heavy chain immunoglobulin locus humanized heterozygous mice with BALB/c background.

The heterozygous mice were then bred with each other to obtain homozygous mice. A detailed description of how to make immunoglobulin heavy chain locus humanized homozygous mice is provided in WO2020169022A1 and US20200390073A1; each of which is incorporated herein by reference in its entirety.

EXAMPLE 6: Modification of mouse immunoglobulin light chain loci

Kappa (κ) light chain locus knock-out mice

The immunoglobulin kappa light chain (κ) locus is located on mouse chromosome 6. The mouse chromosome 6 was modified by knocking out the entire sequence of the immunoglobulin kappa light chain variable region locus. Detailed knock-out methods can be found, e.g., in WO2020169022A1 and US20200390073A1; Zou, X., et al. "Subtle differences in antibody responses and hypermutation of λ light chains in mice with a disrupted χ constant region." *European Journal of Immunology* 25.8 (1995): 2154-2162; Zou, Y. R., et al. "Gene targeting in the Ig kappa locus: efficient generation of lambda chain-expressing B cells, independent of gene rearrangements in Ig kappa." *The EMBO Journal* 12.3 (1993): 811-820; Takeda, S., et al. "Deletion of the immunoglobulin kappa chain intron enhancer abolishes kappa chain gene rearrangement in cis but not lambda chain gene rearrangement in trans." *The EMBO Journal* 12.6 (1993): 2329-2336; each of which is incorporated herein by reference in its entirety.

Lambda (λ) light chain locus knock-out mice

The immunoglobulin lambda light chain (λ) is located on mouse chromosome 16. The mouse chromosome 16 was modified by knocking out the entire sequence of the immunoglobulin lambda light chain variable region locus. Detailed knocking out methods can be found, e.g., in Zou, X., et al. "Block in development at the pre-B-II to immature B cell stage in mice without Igκ and Igλ light chain." *The Journal of Immunology* 170.3 (2003): 1354-1361, which is incorporated herein by reference in its entirety.

The mice described herein can be bred with each other to obtain mice having human immunoglobulin heavy chain VDJ regions, modified mouse immunoglobulin heavy chain constant region locus (lacking CH1 coding region of Cγ1), and lacking all or part of mouse immunoglobulin light chain loci.

EXAMPLE 7: Heavy chain antibodies from mice carrying a modified IgG1 gene

Mice identified above lacking CH1 coding region of Cγ1 (mice with *Mutant allele 2'*, *Mutant allele 3*, and *Mutant allele 4* genotypes, respectively) and a wild-type (WT) mouse were bled to obtain serum samples. The serum samples were prepared for Western blotting analysis to identify any expressed IgG in the serum using an anti-mIgG1 antibody (Cat#: ab190481, Abcam). As shown in **FIG. 32**, the results revealed a mixture of bands: one band of about 75 kD (the expected size for a dimer IgG1 lacking a CH1 domain), and one band of about 150 kD (the expected size for a wild-type IgG). The results indicate that the mice produced using the methods described herein can express an IgG1 lacking a CH1 domain in the peripheral blood.

Mice can be immunized by injecting immunogens, and then screened for antibodies with specific binding by various methods (e.g., hybridoma, phage screening, single-cell technologies by 10x Genomics, or the Beacon® Optofluidic System). Preliminary results showed that mice prepared by the methods described herein can be used to obtain antibodies with high affinity, high diversity, good function (e.g., high endocytic activity), and good developability (e.g., high hydrophilicity and good thermal stability).

Mice with Mutant Allele 3 (Mut3)

Humanized mice (heterozygous for heavy chain *Mutant allele 3* genotype, kappa light chain locus knockout, and lambda light chain locus not knockout) were immunized with antigen A (5 mice). After 3 immunizations, serum titers increased by 10^4 folds as detected by FACS. Beacon® Optofluidic System was then used to isolate the plasma cells that can produce antigen-specific monoclonal antibodies. Expression vector for each antibody was constructed and transfer into a host cell. The number of positive cells confirmed by FACS was 63. Further analysis of the antibody sequences showed that the CDR3 length of the heavy chain variable region (defined by IMGT) was between 6-23 (**FIG. 22**). In addition, 94% of the clones (59/63) had CDR3 lengths greater than or equal to 12, with a total of 35 unique CDR3 sequences. **FIG. 23** shows the germline gene usage of the variable region genes. Further, the affinity of some antibodies was also tested. As shown in **FIG. 24**, the KD of these antibodies against antigen A reached 10^{-9} M, indicating good binding affinity.

In another experiment, humanized mice (4 mice were homozygous for heavy chain *Mutant allele 3* genotype, homozygous for the deletion of kappa light chain locus, and heterozygous for the deletion of lambda light chain locus; 3 mice were homozygous for heavy chain *Mutant allele 3* genotype, homozygous for the deletion of kappa light chain locus, and homozygous for the deletion of lambda light chain locus) were immunized with human 4-1BB (Cat#: 41B-H5258, ACROBiosystems). After 4 immunizations, serum titers increased by 10^4 folds as detected by FACS. The number of antigen-specific clones confirmed by FACS was 67. Further analysis of the antibody sequences showed that the CDR3 length of the heavy chain variable region (defined by IMGT) was between 7-19 (**FIG. 33**). **FIG. 34** shows the germline gene usage of the variable region genes. Further, the affinity of some antibodies was also tested.

In another experiment, humanized mice (19 mice were homozygous for heavy chain *Mutant allele 3* genotype, homozygous for the deletion of kappa light chain locus, and heterozygous for the deletion of lambda light chain locus) were immunized with antigen human CD3ED and cynomolgus CD3ED. After 4 immunizations, serum titers increased by 10^5 folds as detected by FACS. The number of antigen-specific clones confirmed by FACS was 273. Further

analysis of the antibody sequences showed that the CDR3 length of the heavy chain variable region (defined by IMGT) was between 6-25 (**FIG. 35**). **FIG. 36** shows the germline gene usage of the variable region genes. Further, the affinity of some antibodies was also tested.

Humanized mice (homozygous for heavy chain *Mutant Allele 3* genotype, kappa light chain locus and lambda light chain locus knockout) were immunized with antigen Human Serum Albumin (10 mice). Beacon® Optofluidic System was then used isolate the plasma cells that can produce antigen-specific monoclonal antibodies. The number of positive cells confirmed by FACS was 84. Further analysis of the antibody sequences showed that the CDR3 length of the heavy chain variable region (defined by IMGT) was between 12-17. **FIG. 43** shows the germline gene usage of the variable region genes.

Mice with Mutant Allele 2 (Mut2)

Humanized mice (homozygous for heavy chain *Mutant Allele 2* genotype, kappa light chain locus knockout, lambda light chain locus not knockout) were immunized with antigen A (10 mice). Beacon® Optofluidic System was then used to isolate the plasma cells that can produce antigen-specific monoclonal antibodies. The number of positive cells confirmed by FACS was 40. Further analysis of the antibody sequences showed a total of 14 unique CDR3 sequences. In addition, the affinity of some antibodies was also tested. As shown in **FIG. 25**, the KD of these antibodies against antigen A reached 10^{-8} M, indicating good binding affinity.

Further, spleen tissues of the immunized *Mut2* mice were collected to extract total RNA from splenocytes. The immunoglobulin variable region locus can be cloned by PCR, and then inserted to phage plasmid to construct a phage recombinant plasmid library. In one experiment, after 2 rounds of panning and screening of the constructed library, a total of 202 ELISA positive clones were obtained. After removing redundant sequences, the number of positive cells confirmed by FACS was 94. Further analysis of the sequences of these antibodies showed that the CDR3 length of the heavy chain variable region was between 8-18 (defined by IMGT). **FIG. 26** shows the germline gene usage of the variable region genes.

Mice with Mutant Allele 2'(Mut2')

Humanized mice (heterozygous for heavy chain Mutant Allele 2' genotype, kappa light chain and lambda light chain locus knockout) were immunized with antigen Human Serum Albumin (12mice). Beacon® Optofluidic System was then used to isolate the plasma cells that can produce antigen-specific monoclonal antibodies. The number of positive cells confirmed by FACS was 96. Further analysis of the antibody sequences showed that the CDR3 length of the heavy chain variable region (defined by IMGT) was between 5-22. In addition, 78.1% of the clones (75/96) had CDR3 lengths greater than or equal to 12. **FIG. 44** shows the germline gene usage of the variable region genes. Further, the affinity of some antibodies was also tested. As shown in **FIG. 45**, the KD of some antibodies against Human Serum Albumin reached 10^{-9} M, indicating good binding affinity.

Mice with Mutant Allele 4 (Mut4)

Humanized mice (homozygous for heavy chain *Mutant Allele 4* genotype, kappa light chain and lambda light chain locus knockout) were immunized with antigen Human Serum Albumin (10 mice). Beacon® Optofluidic System was then used to isolate the plasma cells that can produce antigen-specific monoclonal antibodies. The number of positive cells confirmed by FACS was 41. Further analysis of the antibody sequences showed that the CDR3 length of the heavy chain variable region (defined by IMGT) was between 8-23, 65.9% of the clones (27/41) had CDR3 lengths greater than or equal to 12. **FIG. 46** shows the germline gene usage of the variable region genes.

Mice with Mutant Allele 5 (Mut5)

Humanized mice (homozygous for heavy chain *Mutant Allele 5* genotype, kappa light chain and lambda light chain locus knockout) were immunized with antigen A. Beacon® Optofluidic System was then used to isolate the plasma cells that can produce antigen-specific monoclonal antibodies. The number of positive cells confirmed by FACS was 20. Further analysis of sequences showed that the CDR3 length of the heavy chain variable region (defined by IMGT) was between 10-18. **FIG. 47** shows the germline gene usage of the variable region genes. In addition, the affinity of some antibodies was also tested. The KD of these antibodies against antigen A reached 10^{-9} M, indicating good binding affinity.

B-cell development

Experiments were performed to compare the immune systems of the modified IgG1 mice and the wild-type mice. 6-8 week old RenMab mice (humanized heavy chain immunoglobulin locus), and modified IgG1 mice were selected. Among them, the modified IgG1 mice had similar body weight, appearance and vitality as compared to the RenMab mice. Peripheral blood, spleen, lymph nodes and bone marrow tissues of these mice were obtained, and no obvious anatomical changes were discovered.

EXAMPLE 8. Generating human anti-TFR1 antibodies

The mice (5 mice were homozygous for heavy chain *Mutant allele 3* genotype, homozygous for the deletion of kappa light chain locus, and heterozygous for the deletion of lambda light chain locus) were immunized with His-tagged human TFR1 (transferrin receptor 1) protein (hTFR1-His, ACROBiosystems, Cat#: CD1-H5243) to obtain anti-TFR1 antibodies. Before immunization, retro-orbital blood was collected as a negative control. Freund's complete adjuvant (CFA) was used for the first immunization and Freund's incomplete adjuvant (IFA) was used for the second and third immunizations. A total of three immunizations (bi-weekly) were performed. One week after the third immunization, retro-orbital blood was collected, and the antibody titer of serum was detected by FACS.

Procedures to enhance immunization were also performed at least fourteen days after the previous immunizations. TFR1 protein was injected by intraperitoneal injection, and the CHO-S cells expressing human TFR1 antigen was injected through the tail vein.

Antigen-specific immune cells were isolated from the immunized mice to further obtain anti-TFR1 antibodies or to obtain the heavy chain variable region sequences of the anti-TFR1 antibodies. For example, single cell technology (e.g., using Beacon® Optofluidic System,

Berkeley Lights Inc.) was used to screen and find plasma cells that secrete antigen-specific monoclonal antibodies. Reverse transcription and PCR sequencing were used to obtain antibody variable region sequences. The obtained variable region sequences were used for antibody expression to verify the binding affinity to TFR1 using FACS. Because the lack of the CH1 domain, the heavy chain variable region (VH) of the obtained antibodies is also referred to as a heavy chain single variable domain (VHH).

Specifically, the obtained VHH sequences were respectively connected to a human IgG1 constant region (e.g., the hinge region, CH2 domain and CH3 domain). Exemplary antibodies obtained by this method included: 23B8, 24A1, 24C9 and 24G5. The heavy chain CDR1-3 sequences are shown in **FIG. 37** and **FIG. 38**. The VHH region sequences of 23B8, 24A1, 24C9 and 24G5 are shown in **FIG. 39**.

The constant region of the antibodies can be further engineered to replace the Asparagine at position 297 with Alanine (N297A). For example, when N297A mutation is introduced into the constant region of 24G5, the resulting antibody is named as 24G5-N.

EXAMPLE 9. Cross-species binding of anti-TFR1 antibodies

CHO-S-hTFR1 cells or CHO-S-fasTFR1 cells were transferred to a 96-well plate at a density of 10^5 cells/well respectively. Serially diluted sample anti-TFR1 antibodies were added to the 96-well plate, and incubated at 4°C for 30 minutes. PBS was used as a negative control (NC). Then, the cells were incubated with the secondary antibody anti-hIgG-Fc-Alex Flour™ 647 (Jackson ImmunoResearch Laboratories, Cat#: 109-606-170) at 4°C in the dark for 15 minutes before flow cytometry analysis.

CHO-S-hTFR1 cells or CHO-S-fasTFR1 cells were obtained by transfecting CHO-S cells with vectors expressing human TFR1 (hTFR1, SEQ ID NO: 70) or *Macaca fascicularis* (crab-eating macaque) TFR1 amino acid sequence (fasTFR1, SEQ ID NO: 71), respectively. The test results are shown in the table below.

JR141, a humanized IgG1 antibody targeting human TFR1 conjugated to human iduronate-2-sulfatase, was first approved in March 2021 in Japan for the intravenous treatment of mucopolysaccharidosis type II. The VH and VL sequences of JR141 are set forth in SEQ ID NO: 72 and SEQ ID NO: 73, respectively. For positive control (JR141-N), the VH and VL of JR141 were connected to a human IgG1 constant region with N297A mutation.

Table 15

Antibody	Cells	Percentage of Positive Cells	Evaluation
NC	CHO-S-hTFR1	0.1%	No Binding
	CHO-S-fasTFR1	0.2%	No Binding
JR141-N	CHO-S-hTFR1	45.0%	Binding
	CHO-S-fasTFR1	37.2%	Binding
23B8	CHO-S-hTFR1	38.8%	Binding
	CHO-S-fasTFR1	39.6%	Binding
24A1	CHO-S-hTFR1	56.4%	Binding
	CHO-S-fasTFR1	50.9%	Binding
24C9	CHO-S-hTFR1	45.8%	Binding

	CHO-S-fasTFR1	0.35%	No Binding
24G5	CHO-S-hTFR1	54.3%	Binding
	CHO-S-fasTFR1	51.2%	Binding

EXAMPLE 10. Binding affinity of anti-TFR1 antibodies

The binding affinity of the anti-TFR1 antibodies to His-tagged TFR1 protein of human (hTFR1-His, ACROBiosystems, Cat#: CD1-H5243) or monkey (fasTFR1-His, ACROBiosystems, Cat#: TFR-C524a) were verified using surface plasmon resonance (SPR) on Biacore™ (Biacore, Inc., Piscataway N.J.) 8K biosensor equipped with pre-immobilized Protein A sensor chips.

Purified anti-TFR1 antibodies was captured on the Protein A chip (Series S Sensor Chip Protein A) for the detection. Purified anti-TFR1 antibodies (1 μg/mL) was loaded at 10 μL/min to bind to hTFR1-His and fasTFR1-His (200 nM). The flow rate was 30 μL/min. The binding and dissociation time were set to 180 seconds and 600 seconds, respectively. The chip was regenerated after the last injection of each titration with a glycine solution (pH 2.0) at 30 μL/min for 30 seconds.

Kinetic association rates (kon) and dissociation rates (koff) were obtained simultaneously by fitting the data globally to a 1:1 Langmuir binding model (Karlsson, R. Roos, H. Fagerstam, L. Petersson, B., 1994. *Methods Enzymology* 6. 99-110) using Biacore™ 8K Evaluation Software 3.0. Affinities were deduced from the quotient of the kinetic rate constants (KD=koff/kon).

As a person of ordinary skill in the art would understand, the same method with appropriate adjustments for parameters (e.g., antibody concentration) was performed for each tested antibody. The results for the tested antibodies are summarized in the table below. The results show that all four anti-TFR1 antibodies can bind to human and monkey TFR1 with high affinity.

Table 16

Protein	Analyte 1 Solution	kon (1/Ms)	koff (1/s)	KD (M)
23B8	hTFR1-His	1.51E+04	1.32E-04	8.73E-09
	fasTFR1-His	4.10E+05	7.99E-03	1.95E-08
24A1	hTFR1-His	2.62E+04	7.68E-05	2.93E-09
	fasTFR1-His	4.36E+04	1.01E-04	2.32E-09
24C9	hTFR1-His	2.68E+04	3.83E-04	1.43E-08
	fasTFR1-His	2.78E+04	3.68E-04	1.32E-08
24G5	hTFR1-His	4.24E+04	8.68E-05	2.05E-09
	fasTFR1-His	6.33E+04	8.79E-05	1.39E-09

EXAMPLE 11. Epitope analysis of anti-TFR1 antibodies

Relative positions of target protein epitope between a pair of purified anti-TFR1 antibodies were analyzed by Biolayer Interferometry (BLI) using ForteBio Octet system at 30 °C. 1× HBS-EP+ buffer (10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 150 mM NaCl, 3 mM ethylenediaminetetraacetic acid (EDTA) and 0.05% P20, pH7.4) diluted from HBS-EP+ buffer (10×) was used as the running buffer throughout the experiment. About 10

μg/mL of hTFR1-His protein was captured by HIS1K (Anti-Penta-HIS) for 200 seconds, and 200 nM of antibody (analyte 1) was injected at a flow rate of 30 μL/min to bind the ligand. Another antibody (analyte 2) was injected under the same conditions to determine whether the binding of different antibodies interfered with each other. The binding time was 300 seconds for each antibody.

The binding value of each antibody was obtained using Data Analysis HT 12.0. To quantify the interference of one antibody binding to another, a binding ratio was calculated to compare each pair of antibodies. The binding ratio is defined as the binding value of the second antibody (analyte 2), divided by the binding value of the first antibody (analyte 1). The binding ratio of each antibody pair is summarized in the matrix table below. Specifically, the binding ratio was between 0.0 to 0.5, if analyte 1 exhibited a blocking effect to analyte 2; the binding ratio was between 0.5-1.1, if analyte 1 did not exhibit a blocking effect to analyte 2. In general, antibody pairs that interfere with each other have the same or overlapping epitopes.

The epitope binding assay results indicate that 24A1 and 24G5 can recognize the same epitope, and 23B8, 24C9 and JR141-N can recognize different epitopes.

Table 17

Analyte 2 \ Analyte 1	23B8	24A1	24G5	24C9	JR141-N
23B8	0.1	0.2	0.3	1.4	1.1
24A1	0.0	0.1	0.1	0.1	0.3
24G5	0.0	0.0	0.0	0.0	0.2
24C9	0.6	0.1	0.2	0.0	0.8
JR141-N	0.7	0.0	0.0	1.1	0.0

EXAMPLE 12. Internalization of anti-TFR1 antibodies

Anti-TFR1 antibodies together with the pHAb-Goat anti-human IgG secondary antibody were added to human cortical microvascular endothelial cells (hCMEC/D cells), and incubated for 3 hours. After incubation, the cells were centrifuged and washed with FACS buffer. Mean fluorescence intensity (MFI) was measured using a flow cytometer. Endocytosis rates of antibodies were calculated. Human IgG1 protein (CrownBio, Cat#: C0001) was used as an isotype control (ISO). The results are shown in the following table, indicating that all four antibodies exhibited good endocytosis activity in human cortical microvascular endothelial cells.

Table 18

Antibody	MFI	Positive Population
ISO	4288	0.9%
23B8	11392	67.0%
24A1	26972	96.6%
24C9	37086	96.0%
24G5	26406	96.0%

EXAMPLE 13. Stability analysis of anti-TFR1 antibodies

Stability of anti-TFR1 antibodies 23B8, 24A1, 24C9 and 24G5 was evaluated. Specifically, the following tests were performed: (1) observing the solution appearance and presence of visible non-soluble objects; (2) detecting the purity changes of antibodies by Size-Exclusion Ultra Performance Liquid Chromatography (SEC-UPLC) (indicated as the percentage of the main peak area to the sum of all peak areas (Purity, %)); (3) detecting changes in the apparent hydrophobicity of the antibodies using the Hydrophobic Interaction Chromatography-High Performance Liquid Chromatography (HIC-HPLC) method (indicated as the retention time of the main peak (HIC, min); (4) detecting charge variants in the antibodies by Capillary Isoelectric Focusing (cIEF) (indicated as the percentages of the main component, acidic component, and alkaline component); and (5) detecting the thermal stability of antibodies via the UNcle system (indicated as the melting temperature (Tm) and aggregation temperature (Tagg)).

In the SEC-UPLC experiments, the Agilent 1290 chromatograph system (connected with XBridge™ Protein BEH SEC column (200 Å, Waters Corporation)) was used. The antibody samples were diluted to 1 mg/mL with purified water. The following parameters were used: mobile phase: 25 mM phosphate buffer (PB) (pH 6.8) + 0.3 M NaCl; flow rate: 1.8 mL/min; column temperature: 25 °C; detection wavelength: 280 nm; injection volume: 10 μL; sample tray temperature: 6 °C; and running time: 7 minutes.

In the HIC-HPLC experiments, the Agilent 1260 chromatograph system (connected with ProPac™ HIC-10 column (4.6 × 100 mm, Thermo Scientific)) was used, and samples were diluted using mobile phase A to 0.5 mg/mL. The following parameters were used: mobile phase A: 0.9 M ammonium sulfate, 0.1 M PB, 10% acetonitrile pH 6.5; mobile phase B: 0.1 M PB, 10% acetonitrile pH 6.5; flow rate: 0.8 mL/min; gradient: 0 min 100% A, 2 min 100% A, 32 min 100% B, 34 min 100% B, 35 min 100% A, and 45 min 100% A; column temperature: 30 °C; detection wavelength: 280 nm; injection volume: 10 μg; sample tray temperature: about 6 °C; and running time: 45 minutes.

In the cIEF experiments, the Maurice cIEF Method Development Kit (Protein Simple, Cat#: PS-MDK01-C) was used for sample preparation. Specifically, 40 μg protein sample was mixed with the following reagents in the kit: 1 μL Maurice cIEF pI Marker-4.05, 1 μL Maurice cIEF pI Marker-9.99, 35 μL 1% Methyl Cellulose Solution, 2 μL Maurice cIEF 500 mM Arginine, 4 μL Ampholytes (Pharmalyte pH ranges 3-10), and water (added to make a final volume of 100 μL). On the Maurice analyzer (Protein Simple, Santa Clara, CA), Maurice cIEF Cartridges (PS-MC02-C) were used to generate imaging capillary isoelectric focusing spectra. The sample was focused for a total of 10 minutes. The analysis software installed on the instrument was used to analyze the absorbance of the 280 nm-focused protein.

In the thermal stability experiments, antibody solutions at 60 mg/mL were heated from 25 to 95 °C using 1 °C increments, with an equilibration time of 1 minute before each measurement.

Detailed results are shown in the table below. The results indicate that all four antibodies have good stability and physical and chemical properties.

Table 19

Antibody	SEC Purity (%)	HIC (min)	PI	Tm			Tagg (°C)
				Tm1 (°C)	Tm2 (°C)	Tm3(°C)	
23B8	92.82	19.909	7.376	58.87	67.80	84.13	56.48

24A1	98.74	2.486	6.795	64.80	73.80	82.20	52.73
24C9	98.20	3.740	8.595	65.38	69.90	78.20	33.63
24G5	98.86	3.829	6.495	60.86	77.66	85.30	76.03

EXAMPLE 14. Pharmacokinetics (PK) analysis

A humanized TFR1 mouse model (hTFR1 mice) was engineered to express a chimeric TFR1 protein (SEQ ID NO: 74) in which the extracellular region of mouse TFR1 protein was replaced with the corresponding human TFR1 extracellular region. A detailed description regarding the humanized TFR1 mouse model can be found in PCT Application No. PCT/CN2022/105924, which is incorporated herein by reference in its entirety.

The concentrations of the anti-TFR1 antibodies were determined in hTFR1 mice. Specifically, the mice were placed into different groups (8 mice per group), and administered with an approximately equal molar dosage of JR141-N (G2), 23B8-N (G3), 24A1-N (G4), 24G5-N (G5), or 24C9-N (G6), by intravenous (i.v.) injection. The control group (G1) mice were administered with human IgG1 (hIgG1). Details of the administration scheme are shown in the table below.

Table 20

Group	No. of mice	Antibodies	Dosage	Route	Frequency
G1	8	hIgG1	18.4 mg/kg	i.v.	single dose
G2	8	JR141-N	18.4 mg/kg	i.v.	single dose
G3	8	23B8-N	10.0 mg/kg	i.v.	single dose
G4	8	24A1-N	10.0 mg/kg	i.v.	single dose
G5	8	24G5-N	10.0 mg/kg	i.v.	single dose
G6	8	24C9-N	10.0 mg/kg	i.v.	single dose

Blood and brain samples were collected 0.5, 6, 24, and 72 hours after the administration. Two mice were sampled at each time point, and the mice were anesthetized after retro-orbital blood was collected. To avoid interference from the residual blood in the brain, the mice were perfused by saline for 10 minutes at room temperature. Specifically, saline was perfused via systemic circulation from the left ventricle to the right ventricle. Brain samples were excised and divided into two hemibrains by the sagittal plane. The left hemibrain was subjected to quantification of the injected antibody, while the right hemibrain was fixed by formalin and embedded with paraffin for serial sections. The brain samples were cut into pieces and homogenized with DPBS (Dulbecco's phosphate-buffered saline) containing 1× mixed protease inhibitors. The brain homogenate was aliquoted for protein extraction followed by antibody quantification by electrochemiluminescenc. For the rest homogenate, capillaries were depleted by gradient density centrifugation at 5400 g for 15 minutes, using 15% dextran. After centrifugation, the fraction at the top of the centrifuge tube was saved as parenchyma and subjected to protein extraction and antibody quantification as well. **FIGS. 40A-40D** show the antibody concentration in total brain protein (**FIG. 40A**), the ratio of antibody concentration in brain total protein to serum antibody concentration (**FIG. 40B**), the antibody concentration in brain parenchyma (**FIG. 40C**), and the ratio of antibody concentration in brain parenchyma to

serum antibody concentration at each time point (**FIG. 40D**). These results demonstrated that 24G5-N (group G5) was the most abundant either in parenchyma or in the whole brain.

In a similar experiment, hTFR1 mice were placed into five groups (3 mice per group) and administered with 18.4 mg/kg JR141-N (G2), 10 mg/kg 23B8-N (G3), 10 mg/kg 24A1-N (G4), or 10 mg/kg 24G5-N (G5) by intravenous injection (1 administration in total). The control group (G1) mice were administered with hIgG1 (G1). 24 hours after administration, brain samples were collected to determine the concentrations of the anti-TFR1 antibodies. **FIGS. 41A-41B** show the antibody concentration test results in brain parenchyma and brain total protein, respectively. All tested antibodies showed a higher concentration in brain than hIgG1 (G1), indicating that compared with the positive control JR141-N (G2), 23B8-N (G3) and 24G5-N (G5) can better pass across the blood-brain barrier and enter the brain parenchyma.

In another similar experiment, hTFR1 mice were placed into seven groups (6 mice per group) and administered with JR141-N (G2-G4) or 24G5-N (G5-G7)) by intravenous (i.v.) injection. The control group (G1) mice were administered with hIgG1. Details of the administration scheme are shown in the table below.

Table 21

Group	No. of mice	Antibodies	Dosage (mg/kg)	Route	Frequency
G1	6	hIgG1	5.52 mg/kg	i.v.	single dose
G2	6	JR141-N	1.84 mg/kg	i.v.	single dose
G3	6	JR141-N	5.52 mg/kg	i.v.	single dose
G4	6	JR141-N	18.4 mg/kg	i.v.	single dose
G5	6	24G5-N	1 mg/kg	i.v.	single dose
G6	6	24G5-N	3 mg/kg	i.v.	single dose
G7	6	24G5-N	10 mg/kg	i.v.	single dose

6 hours and 24 hours after the administration, blood and brain samples were collected using the methods described above. Three mice were sampled at each time point. Treatment of the tissues and quantification of antibodies were also performed as described above. The measurement results of the concentration of humanized anti-TFR1 antibodies in the brain parenchyma are shown in **FIG. 42.** The results show that the antibody concentration accumulated in the brain parenchyma of 24G5-N was significantly higher than that of hIgG1 under each dosage condition. Furthermore, the concentration of both JR141-N and 24G5-N exhibited a dose-dependent trend in brain parenchyma.

To detect the distribution of the humanized anti-TFR1 antibody 24G5-N in mouse brain, immunofluorescence assays were performed by staining hIgG, hTFR1, and mCD31, respectively, on right hemibrain sections of the mice used in the experiment above. The results showed that, mCD31 was well-labeled in the microvessels. hTFR1 was also detected on microvessels, which was co-localized with mCD31. Besides, hTFR1 expression was also detected on several neurons in parenchyma. In particular, the anti-TFR1 antibody 24G5-N was stained by a secondary anti-IgG antibody conjugated with DyLight® 488. Similar to hTFR1, 24G5-N was detected in microvessels and parenchyma, and its signal overlapped with the hTFR1 signal. Thus, either for the quantification of 24G5-N in the whole brain or parenchyma, or the visual evidence of

immunofluorescence of 24G5-N in the parenchyma, the results indicate that humanized anti-TFR1 antibody 24G5-N can pass across the blood-brain barrier (BBB) efficiently.

EXAMPLE 15. Blocking assay

Blockade of TFR1 binding to TF (transferrin) by anti-TFR1 antibodies 23B8, 24A1, 24C9 and 24G5 were tested by Biolayer Interferometry (BLI) using ForteBio Octet® system at 30 ℃. Specifically, 1× HBS-EP+ buffer (10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 150 mM NaCl, 3 mM EDTA and 0.05% Surfactant P20, pH 7.4) diluted from HBS-EP+ buffer (10×) was used as the running buffer throughout the experiment. About 10 μg/mL of antibodies were captured by AHC (Anti-Human IgG Fc Capture) for 200 seconds, and 800 nM of hTFR1-His (ACROBiosystems, Cat#: CD1-H5243) and hTF-His (human transferrin protein, Kactus Biosystems, Cat#: TFN-HM101) were injected to bind the ligand. The binding time was 300 seconds for each antibody. The binding value of each antibody was obtained using Data Analysis HT 12.0. The results showed that these four antibodies did not block the binding of TFR1 to TF. Therefore, such non-blocking antibodies are not likely to interfere with TFR1-TF interactions in normal cells.

CLAIM

1. A genetically modified non-human animal comprising a modified immunoglobulin heavy chain locus, wherein the modified immunoglobulin heavy chain locus comprises an IgG constant region gene, wherein the IgG constant region gene encodes an IgG heavy chain constant region lacking a CH1 domain, wherein the genetically modified non-human animal expresses a heavy-chain antibody.

2. The genetically modified non-human animal of claim 1, wherein the animal comprises exactly one IgG constant region gene.

3. The genetically modified non-human animal of claim 1 or 2, wherein the IgG heavy chain constant region gene is IGHG1.

4. The genetically modified non-human animal of any one of claims 1-3, wherein the IgG heavy chain constant region comprises or consists of a CH2 domain and a CH3 domain, and optionally a hinge region.

5. A genetically modified non-human animal whose genome comprises a germline genetic modification comprising a deletion of IGHG3, IGHG2b, and IGHG2c genes and a deletion of the CH1 exon of IGHG1 gene at an endogenous immunoglobulin heavy chain gene locus.

6. The animal of claim 5, wherein the germline genetic modification further comprises a deletion of endogenous IGHE gene at the endogenous immunoglobulin heavy chain gene locus.

7. The animal of claim 5 or 6, wherein the genetic modification further comprises a deletion of endogenous Sγ2b, Sγ2c, and Sε switch regions at the endogenous immunoglobulin heavy chain gene locus.

8. The animal of any one of claims 5-7, wherein the modified immunoglobulin heavy chain gene locus comprises a modified IGHG1 gene lacking a sequence encoding a CH1 domain, wherein the modified IGHG1 gene comprises a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 1.

9. The animal of any one of claims 5-8, wherein the genetic modification further comprises a deletion of endogenous Sγ3 switch region at the endogenous immunoglobulin heavy chain gene locus.

10. The animal of any one of claims 5-9, wherein the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, the modified rodent IGHG1 gene lacking a sequence encoding a CH1 domain, and endogenous IGHM, IGHδ, IGHA genes.

11. The animal of any one of claims 5-9, wherein the genetic modification further comprises a deletion of endogenous IGHM and IGHδ genes at the endogenous immunoglobulin heavy chain gene locus.

12. The animal of any one of claims 5-9 and 11, wherein the animal genome comprises endogenous Sμ, Sγ1, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and the endogenous IGHA gene.

13. The animal of claim 12, wherein the Sμ and Sγ1 switch regions are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 8.

14. The animal of any one of claims 5-9, wherein the genetic modification further comprises a deletion of the CH1 coding sequence of IGHM gene at the endogenous immunoglobulin heavy chain gene locus.

15. The animal of any one of claims 5-9 and 14, wherein the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, a modified endogenous IGHM gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and endogenous IGHδ, IGHA genes.

16. The animal of claim 15, wherein the Sμ switch region and the modified IGHM gene are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 10.

17. The animal of any one of claims 5-9, wherein the genetic modification further comprises a deletion of the CH1 exon of IGHM gene and a deletion of IGHδ gene at the endogenous immunoglobulin heavy chain gene locus.

18. The animal of any one of claims 5-9 and 17, wherein the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, a modified IGHM gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and an endogenous IGHA gene.

19. The animal of any one of claims 5-9, wherein the genetic modification further comprises a deletion of the CH1 exon of IGHM gene and a deletion of the CH1 coding sequence of IGHδ gene at the endogenous immunoglobulin heavy chain gene locus.

20. The animal of any one of claims 5-9 and 19, wherein the animal's genome comprises endogenous Sμ, Sγ1, Sα switch regions, a modified IGHM gene lacking a sequence encoding a CH1 domain, a modified IGHδ gene lacking a sequence encoding a CH1 domain, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and an endogenous IGHA gene.

21. The animal of claim 19 or 20, wherein the modified IGHM gene are linked with a sequence that is at least 80%, 90%, 95% or 99% identical to SEQ ID NO: 10, and the modified IGHδ gene comprises a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 41.

22. The animal of any one of claims 14-21, wherein the modified IGHM gene comprises a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 13.

23. The animal of any one of claims 5-8, wherein the genetic modification further comprises a deletion of endogenous Sγ1 switch region at the endogenous immunoglobulin heavy chain gene locus.

24. The animal of any one of claims 5-8 and 23, wherein the animal's genome comprises endogenous Sμ, Sγ3, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and endogenous IGHM, IGHδ, IGHA genes.

25. The animal of any one of claims 5-8 and 23, wherein the genetic modification further comprises a deletion of endogenous Sγ3 switch region at the endogenous immunoglobulin heavy chain gene locus.

26. The animal of any one of claims 5-8, 23 and 25, wherein the genetic modification further comprises a deletion of endogenous IGHM and IGHδ genes at the endogenous immunoglobulin heavy chain gene locus.

27. The animal of any one of claims 5-8, 23, 25 or 26, wherein the animal's genome comprises endogenous Sμ, Sα switch regions, the modified IGHG1 gene lacking a sequence encoding a CH1 domain, and the endogenous IGHA gene.

28. The animal of claim 27, wherein the Sμ switch region and the modified IGHG1 gene are linked with a sequence that is at least 80%, 90%, 95%, or 99% identical to SEQ ID NO: 9.

29. The animal of any one of claims 5-10, 23 and 24, wherein the modified genome comprises a functional IGHM gene.

30. A genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene; wherein the elements are operably linked.

31. The animal of claim 30, wherein the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

32. A genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene, IGHδ gene, Sγ3 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene; wherein the elements are operably linked.

33. The animal of claim 32, wherein the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene, IGHδ gene, Sγ3 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

34. A genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene; wherein the elements are operably linked.

35. The animal of claim 34, wherein the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

36. A genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene; wherein the elements are operably linked.

37. The animal of claim 36, wherein the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

38. A genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene; wherein the elements are operably linked.

39. The animal of claim 38, wherein the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

40. A genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene

lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene; wherein the elements are operably linked.

41. The animal of claim 40, wherein the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

42. A genetically-modified non-human animal whose genome comprises the following elements, at an endogenous immunoglobulin heavy chain gene locus in a 5' to 3' order: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene; wherein the elements are operably linked.

43. The animal of claim 42, wherein the endogenous immunoglobulin heavy chain constant region gene locus consists of the following functional genes and switch regions: Sμ switch region, IGHM gene lacking a sequence encoding a CH1 domain, IGHδ gene lacking a sequence encoding a CH1 domain, Sγ1 switch region, IGHG1 gene lacking a sequence encoding a CH1 domain, Sα switch region, and IGHA gene.

44. The animal of any one of claims 1-43, wherein the animal expresses a heavy-chain antibody comprising an IgG heavy chain constant region lacking the CH1 domain.

45. The animal of claim 44, wherein the heavy-chain antibody binds to it target antigen with a KD of less than 10^{-7} M, less than 10^{-8} M, or less than 10^{-9} M.

46. The animal of claim 44 or 45, wherein the heavy-chain antibody comprises or consists of a variable region, a CH2 domain and a CH3 domain.

47. The animal of any one of claims 44-46, wherein the heavy-chain antibody further comprises a transmembrane domain and/or a cytoplasmic domain.

48. The animal of any one of claims 1-47, wherein the genetically modified non-human animal does not express IgG antibodies comprising light chains.

49. The animal of any one of claims 1-48, wherein the animal expresses IgM, IgD, and/or IgA (e.g., functional IgM, IgD, and/or IgA).

50. The genetically modified non-human animal of any one of claims 1-49, wherein the animal comprises at an endogenous immunoglobulin heavy chain gene locus, one or more human IGHV genes, one or more human IGHD genes, and one or more human IGHJ genes, wherein the human IGHV genes, the human IGHD genes, and the human IGHJ genes are operably linked and can undergo VDJ rearrangement.

51. The genetically modified non-human animal of claim 50, wherein the animal comprises at least 150 human IGHV genes selected from Table 1, at least 20 human IGHD genes selected from Table 2, and at least 5 human IGHJ genes selected from Table 3.

52. The genetically modified non-human animal of claim 50, wherein the animal comprises all human IGHV genes, all human IGHD genes, and all human IGHJ genes at the endogenous immunoglobulin heavy chain gene locus of human chromosome 14 of a human subject.

53. The genetically modified non-human animal of claim 50, wherein the animal comprises all human IGHV genes, all human IGHD genes, and all human IGHJ genes at the endogenous immunoglobulin heavy chain gene locus of human chromosome 14 of a human cell.

54. The genetically modified non-human animal of any one of claims 50, wherein the animal is a mouse and the genetic modification in the animal's endogenous immunoglobulin heavy chain gene locus comprises a deletion of one or more mouse IGHV genes in Table 4, one or more mouse IGHD genes in Table 5, and/or one or more mouse IGHJ genes in Table 6.

55. The genetically modified non-human animal of claim 54, wherein the animal is a mouse and the genetic modification in the animal's endogenous heavy chain immunoglobulin gene locus comprises a deletion of a contiguous sequence starting from mouse IGHV1-85 gene to mouse IGHJ4 gene.

56. The genetically modified non-human animal of any one of claims 50, wherein the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus, wherein the unmodified human sequence is at least 800 kb.

57. The genetically modified non-human animal of claim 50, wherein the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHV1-2.

58. The genetically modified non-human animal of claim 50, wherein the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHV6-1.

59. The genetically modified non-human animal of claim 50, wherein the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHD1-1 to human IGHJ6.

60. The genetically modified non-human animal of claim 50, wherein the animal comprises an unmodified human sequence derived from a human heavy chain immunoglobulin gene locus starting from human IGHV(III)-82 to human IGHJ6.

61. The genetically modified non-human animal of any one of claims 1-49, whose genome comprises, at the endogenous immunoglobulin heavy chain locus: a replacement of one or more endogenous IGHV, endogenous IGHD, and endogenous IGHJ genes with one or more human IGHV, human IGHD, and human IGHJ genes, wherein human IGHV, human IGHD, and human IGHJ genes are operably linked to one or more of endogenous IGHM, IGHδ, IGHG1 lacking a sequence encoding the CH1 domain, and IGHA genes.

62. The genetically modified non-human animal of claim 61, wherein one or more endogenous IGHV, endogenous IGHD, and endogenous IGHJ genes are replaced by at least 150 human IGHV genes in Table 1, at least 20 human IGHD genes in Table 2, and at least 5 human IGHJ genes in Table 3.

63. The genetically modified non-human animal of claim 61 or 62, wherein the animal is a mouse, and at least 180 mouse IGHV genes in Table 4, all mouse IGHD genes in Table 5, and all mouse IGHJ genes in Table 6 are replaced.

64. The genetically modified non-human animal of any one of claims 1-63, wherein the animal is homozygous with respect to the immunoglobulin heavy chain gene locus.

65. The genetically modified non-human animal of any one of claims 1-63, wherein the animal is heterozygous with respect to the immunoglobulin heavy chain gene locus.

66. The genetically modified non-human animal of any one of claims 1-65, wherein the animal comprises an endogenous light chain immunoglobulin gene locus.

67. The genetically modified non-human animal of any one of claims 1-65, wherein the animal comprises a disruption in the endogenous immunoglobulin light chain gene locus.

68. The genetically modified non-human animal of any one of claims 1-67, wherein the animal lacks an endogenous immunoglobulin heavy chain variable region locus that is capable of rearranging and forming a nucleic acid sequence that encodes an endogenous heavy chain variable domain.

69. The genetically modified non-human animal of any one of claims 1-68, wherein the animal can produce a humanized antibody.

70. The genetically modified non-human animal of any one of claims 1-53, 56-62, and 64-69, wherein the animal is a mammal.

71. The genetically modified non-human animal of claims 11-53, 56-62, and 64-70, wherein the animal is a rodent.

72. The genetically modified non-human animal of claims 1-53, 56-62, and 64-71, wherein the animal is a mouse.

73. The genetically modified non-human animal of any one of claims 1-72, wherein the animal has substantially normal B cell development and maturation.

74. A cell obtained from the genetically modified non-human animal of any one of claims 1-73.

75. The cell of claim 74, wherein the cell is a B cell that expresses a chimeric immunoglobulin heavy chain comprising an immunoglobulin heavy chain variable domain that is derived from a rearrangement of one or more human IGHV genes, one or more human IGHD genes, and one or more human IGHJ genes, wherein the immunoglobulin heavy chain variable domain is operably linked to a non-human heavy chain constant region.

76. The cell of claim 74 or 75, wherein the cell is an embryonic stem (ES) cell.

77. A method of making an antibody that specifically binds to an antigen, the method comprising
    a) exposing the genetically modified non-human animal of any one of claims 1-73 to the antigen;
    b) producing a hybridoma from a cell collected from the animal; and
    c) collecting a heavy-chain antibody produced by the hybridoma.

78. The method of claim 77, wherein the method further comprises sequencing the genome of the hybridoma.

79. A method of making an antibody that specifically binds to an antigen, the method comprising
    a) exposing the genetically modified non-human animal of any one of claims 1-73 to the antigen;
    b) sequencing nucleic acids encoding human immunoglobulin heavy chain variable regions in a cell that expresses a heavy-chain antibody that specifically binds to the antigen; and
    c) operably linking in a cell the nucleic acid encoding the human immunoglobulin heavy chain variable region with a nucleic acid encoding a human immunoglobulin heavy chain constant region.

80. A method of making an antibody that specifically binds to an antigen, the method comprising
    a) obtaining a nucleic acid sequence encoding human immunoglobulin heavy chain variable regions in a cell that expresses a heavy-chain antibody that specifically binds to the antigen, wherein the cell is obtained by exposing the genetically modified non-human animal of any one of claims 1-73 to the antigen;
    b) operably linking the nucleic acid encoding the human immunoglobulin heavy chain variable region with a nucleic acid encoding a human immunoglobulin heavy chain constant region; and
    c) expressing the nucleic acid in a cell, thereby obtaining the antibody.

81. A method of obtaining a nucleic acid that encodes an antibody binding domain that specifically binds to an antigen, the method comprising
    a) exposing the genetically modified non-human animal of any one of claims 1-73 to the antigen; and
    b) sequencing nucleic acids encoding human immunoglobulin heavy chain variable regions in a cell that expresses a heavy-chain antibody that specifically binds to the antigen.

82. A method of making an antibody that specifically binds to antigen, the method comprising
    a) exposing the genetically modified non-human animal of any one of claims 1-73 to the antigen;
    b) constructing a phage plasmid library using RNA prepared from immune cells (e.g., splenocytes) of the animal;
    c) screening the phage plasmid library; and
    d) sequencing nucleic acids encoding human immunoglobulin heavy chain variable regions from a phage plasmid that encodes a heavy-chain antibody that specifically binds to the antigen.

83. The method of claim 82, wherein screening comprises isolating phages expressing immunoglobulin heavy chain variable regions based on binding affinity to the antigen.

84. A method of obtaining a sample, the method comprising
    a) exposing the genetically modified non-human animal of any one of claims 1-73 to the antigen; and
    b) collecting the sample from the animal.

85. The method of claim 84, wherein the sample is an immune cell, a lymphoid tissue, a spleen tissue, a spleen cell, or a B cell.

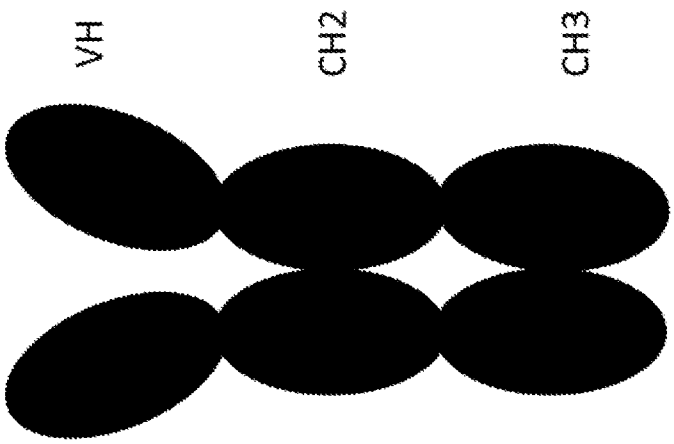


FIG. 1

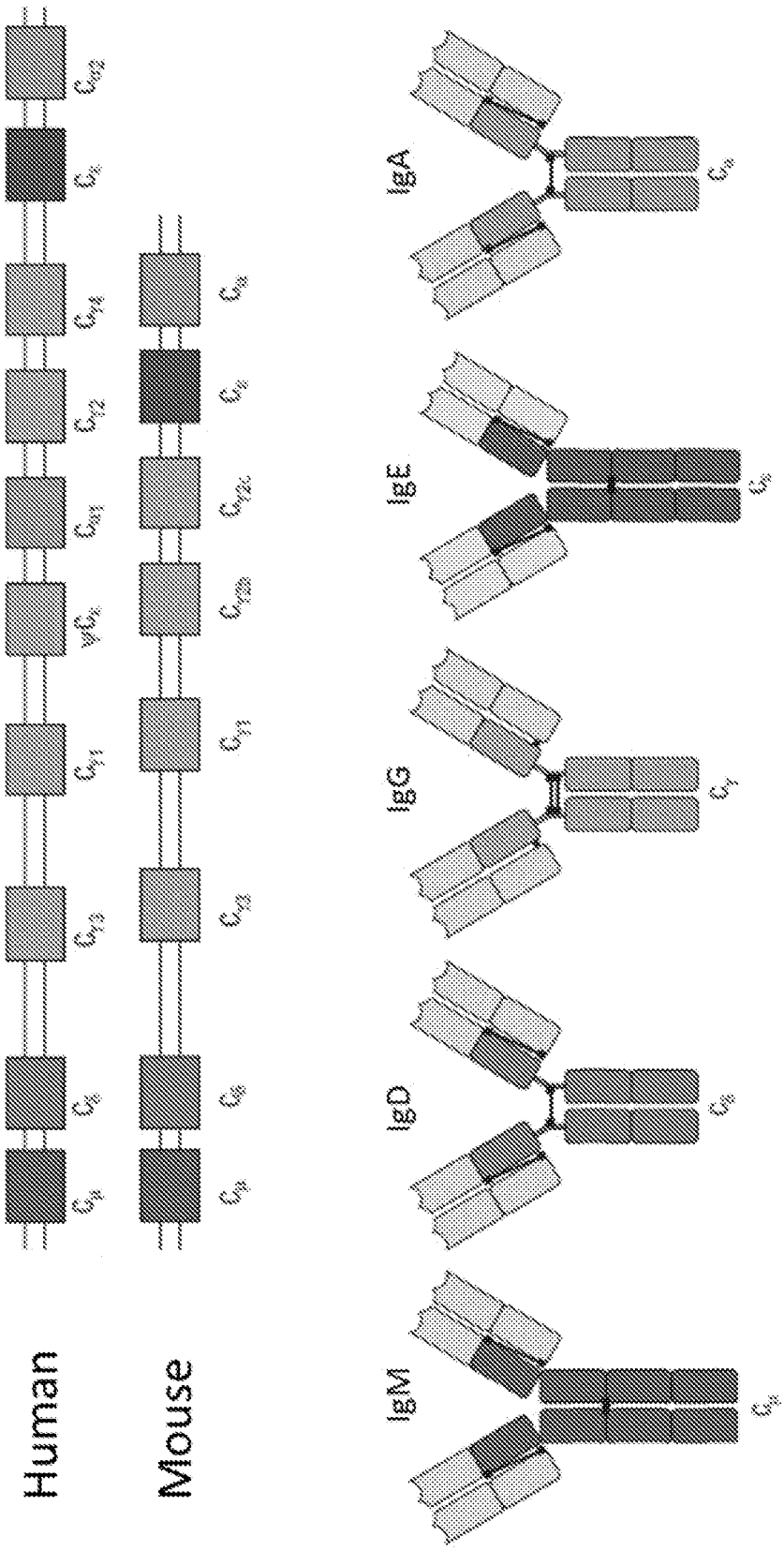


FIG. 2

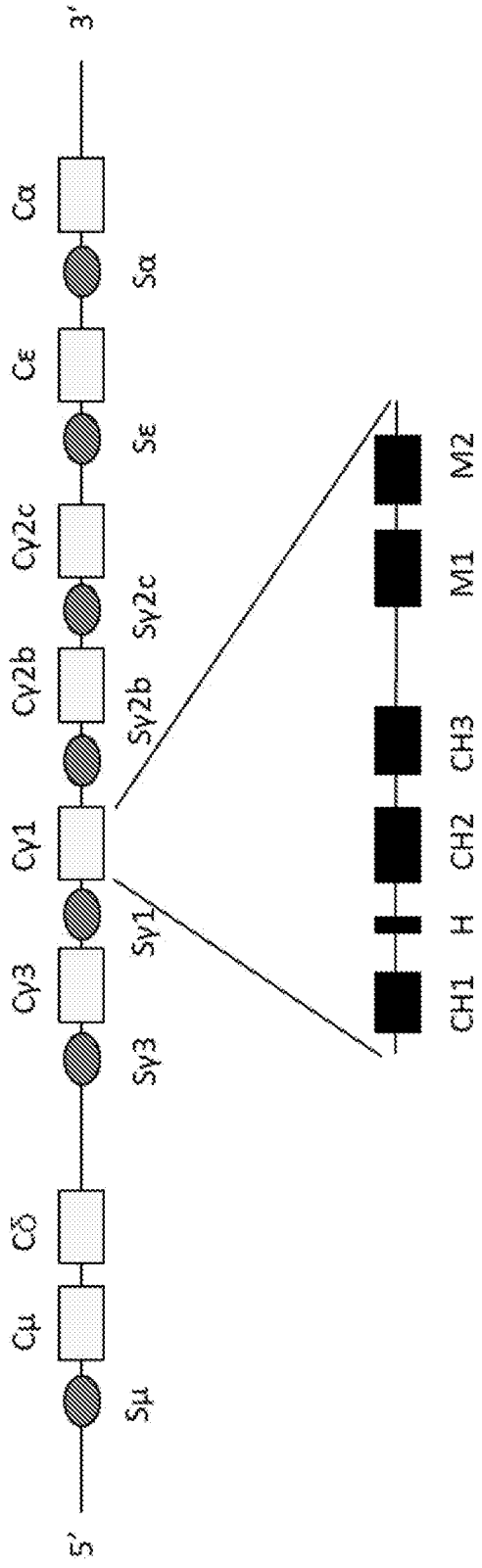

5'
Sμ
Cμ
Cδ
Sγ3
Cγ3
Sγ1
Cγ1
Sγ2b
Cγ2b
Sγ2c
Cγ2c
Sε
Cε
Sα
Cα
3'
CH1
H
CH2
CH3
M1
M2

FIG. 3

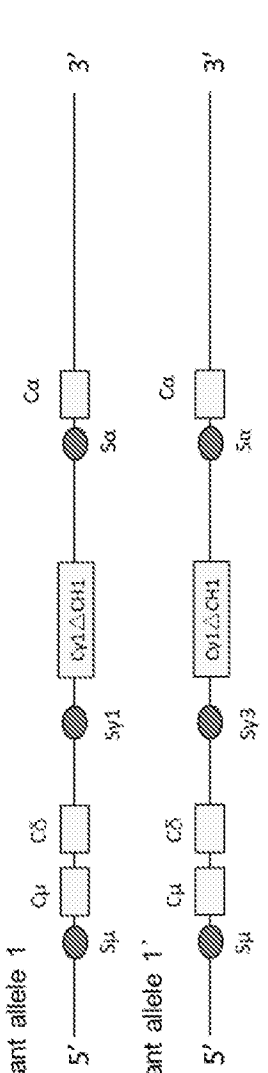

Mutant allele 1
5'
Cμ
Cδ
Sμ
Cγ1△CH1
Sγ1
Cα
Sα
3'
Mutant allele 1'
5'
Cμ
Cδ
Sμ
Cγ1△CH1
Sγ3
Cα
Sα
3'

FIG. 4A

Mutant allele 2
5'
Sμ
Cγ1△CH1
Sγ1
Cα
Sα
3'
Mutant allele 2'
5'
Sμ
Cγ2△CH1'
Cα
Sα
3'
Mutant allele 3
5'
Cμ △CH1
Cδ
Sμ
Cγ1△CH1
Sγ1
Cα
Sα
3'
Mutant allele 4
5'
Cμ △CH1
Sμ
Cγ1△CH1
Sγ1
Cα
Sα
3'
Mutant allele 5
5'
Cμ △CH1
Cδ△CH1
Sμ
Cγ1△CH1
Sγ1
Cα
Sα
3'

FIG. 4B

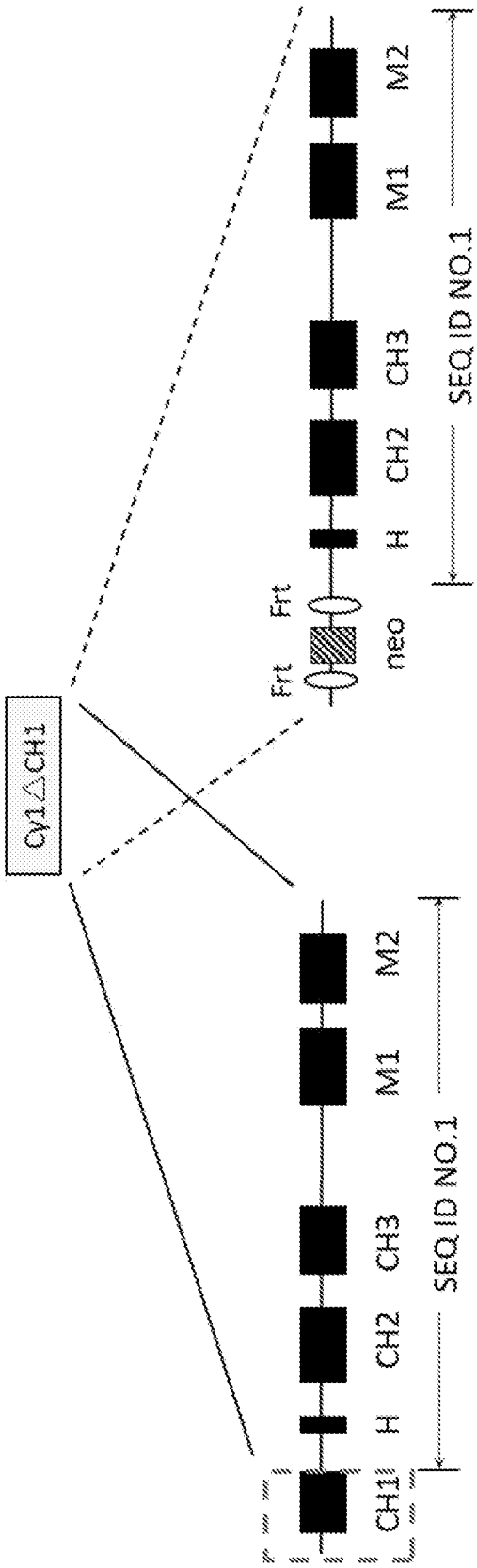

Cγ1△CH1
Frt
Frt
neo
H
CH2
CH3
M1
M2
SEQ ID NO.1
CH1
H
CH2
CH3
M1
M2
SEQ ID NO.1

FIG. 5

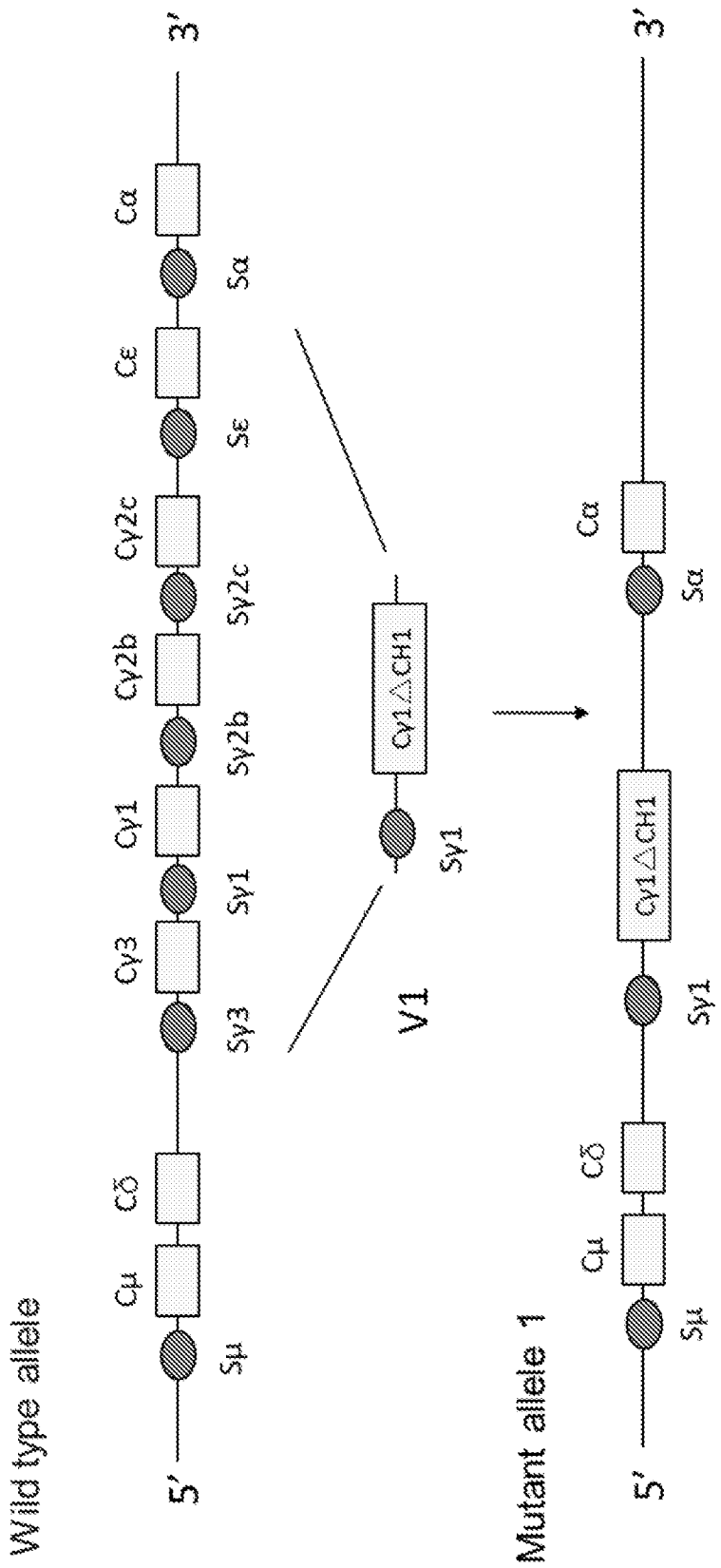

Wild type allele
5'
Cμ
Cδ
Cγ3
Cγ1
Cγ2b
Cγ2c
Cε
Cα
3'
Sμ
Sγ3
Sγ1
Sγ2b
Sγ2c
Sε
Sα
V1
Cγ1△CH1
Sγ1
Mutant allele 1
5'
Cμ
Cδ
Cγ1△CH1
Cα
Sμ
Sγ1
Sα
3'

FIG. 6

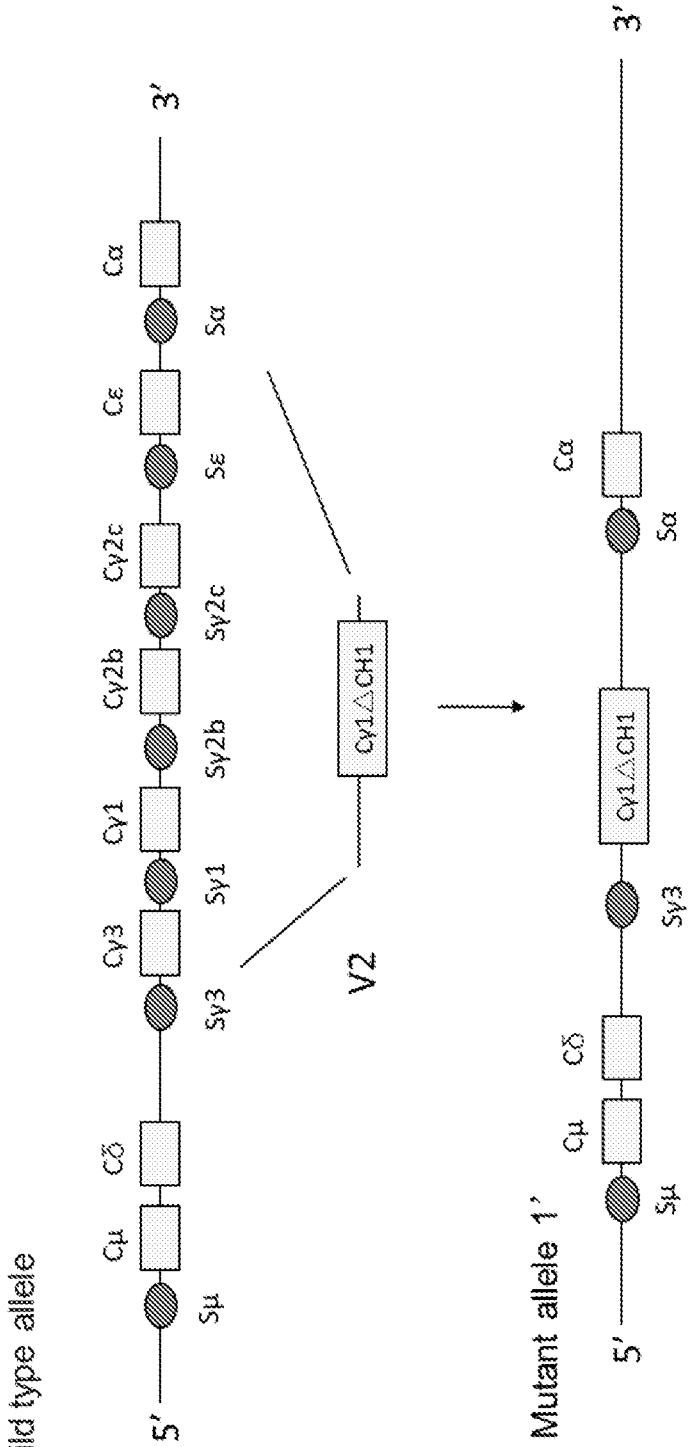

Wild type allele
5'
Cμ
Cδ
Cγ3
Cγ1
Cγ2b
Cγ2c
Cε
Cα
3'
Sμ
Sγ3
Sγ1
Sγ2b
Sγ2c
Sε
Sα
V2
Cγ1△CH1
Mutant allele 1'
5'
Cμ
Cδ
Cγ1△CH1
Cα
3'
Sμ
Sγ3
Sα

FIG. 7

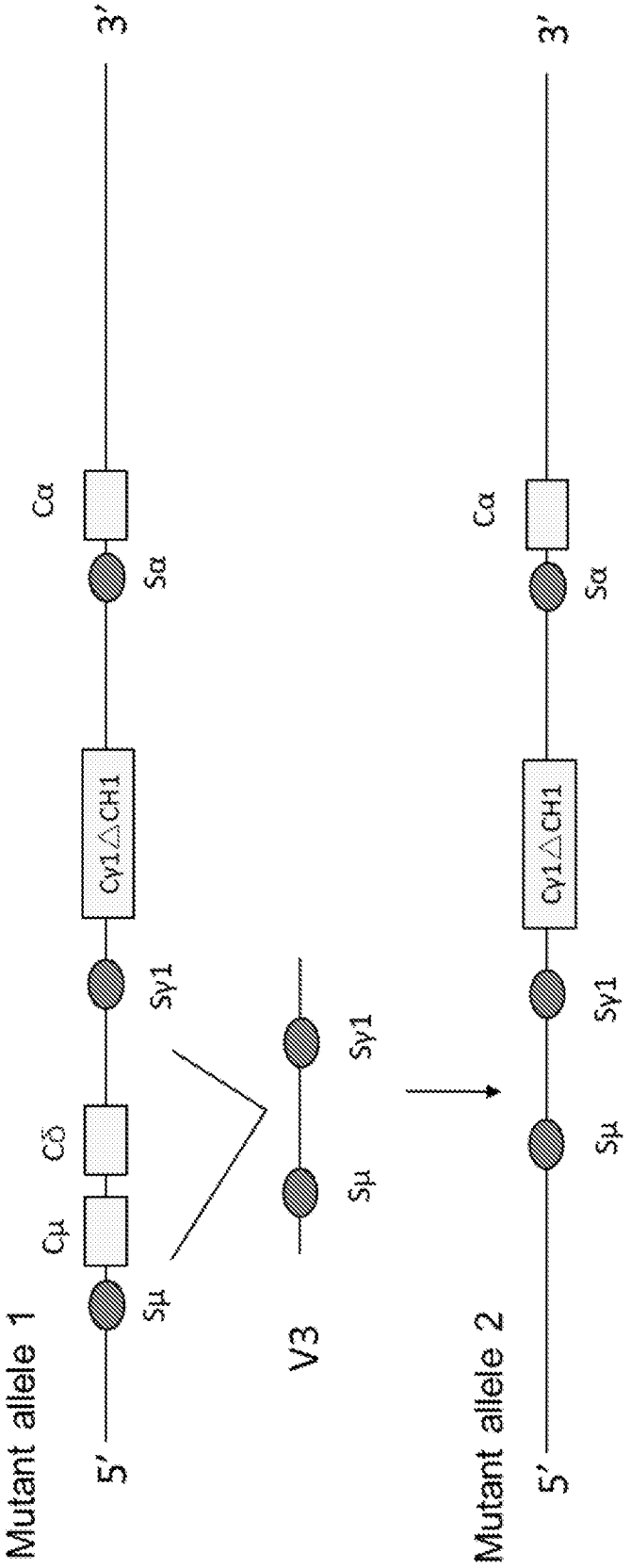

Mutant allele 1
5'
3'
Cμ
Cδ
Cγ1△CH1
Cα
Sμ
Sγ1
Sα
V3
Sγ1
Sμ
Mutant allele 2
5'
3'
Sμ
Sγ1
Cγ1△CH1
Sα
Cα

FIG. 8

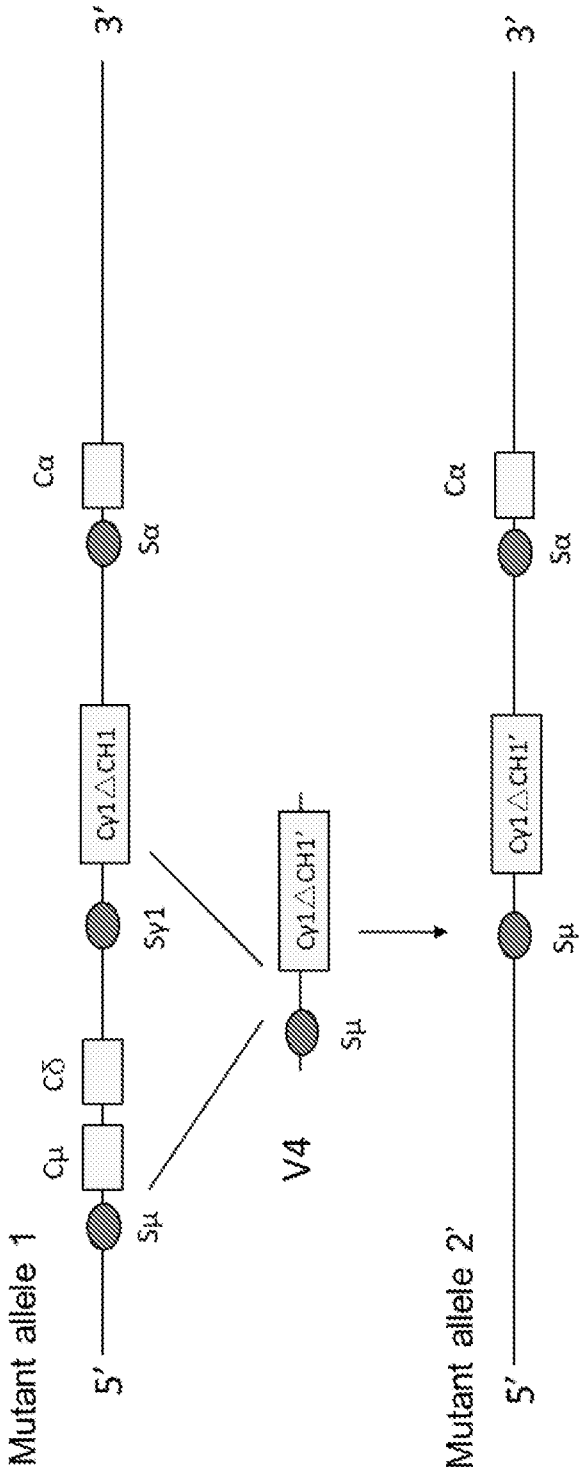

Mutant allele 1
5'
3'
Cμ
Cδ
Cγ1ΔCH1
Cα
Sμ
Sγ1
Sα
V4
Cγ1ΔCH1'
Sμ
Mutant allele 2'
5'
3'
Cγ1ΔCH1'
Cα
Sμ
Sα

FIG. 9

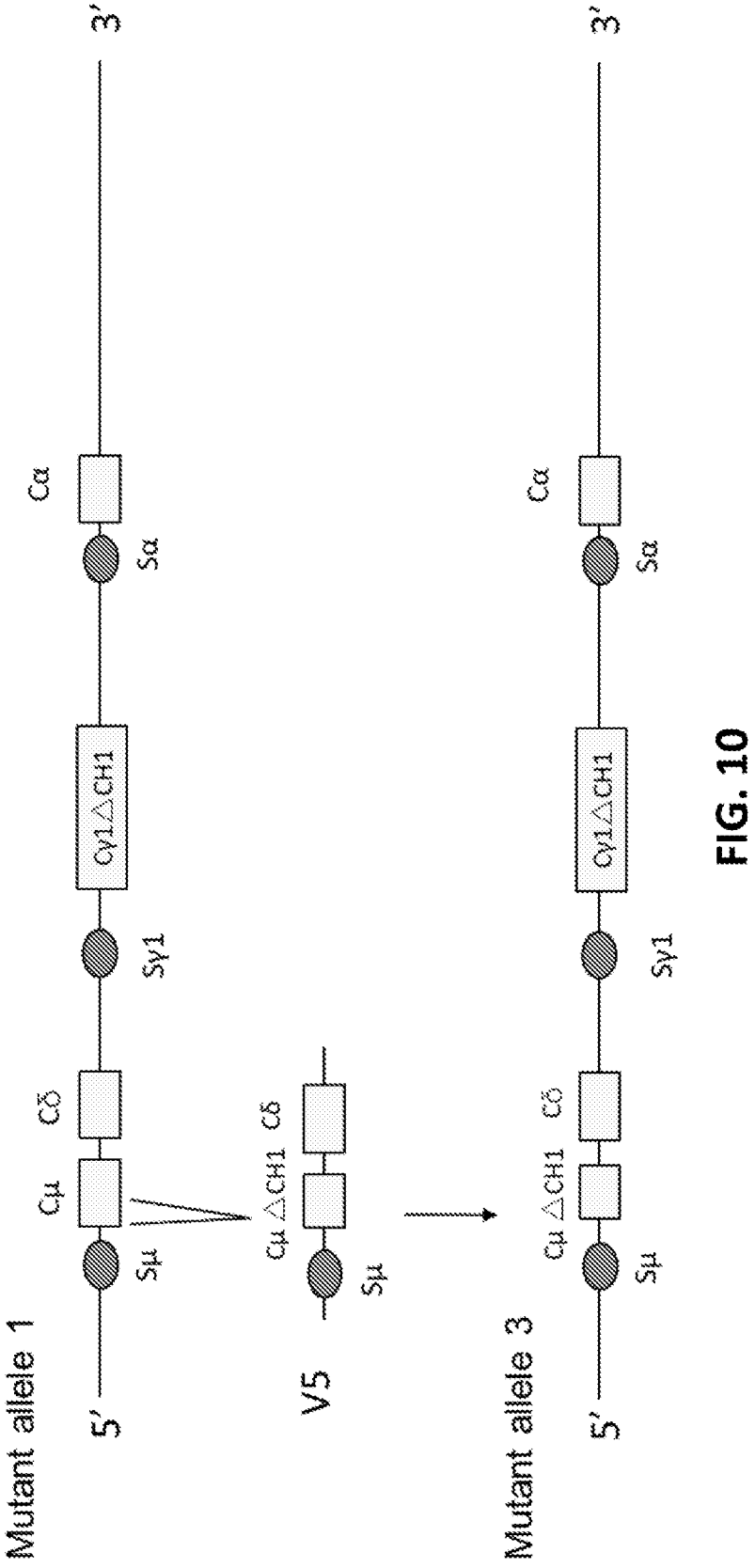

Mutant allele 1
5'
Sμ
Cμ
Cδ
Sγ1
Cγ1△CH1
Sα
Cα
3'
V5
Cμ △CH1 Cδ
Sμ
Mutant allele 3
5'
Sμ
Cμ △CH1
Cδ
Sγ1
Cγ1△CH1
Sα
Cα
3'

FIG. 10

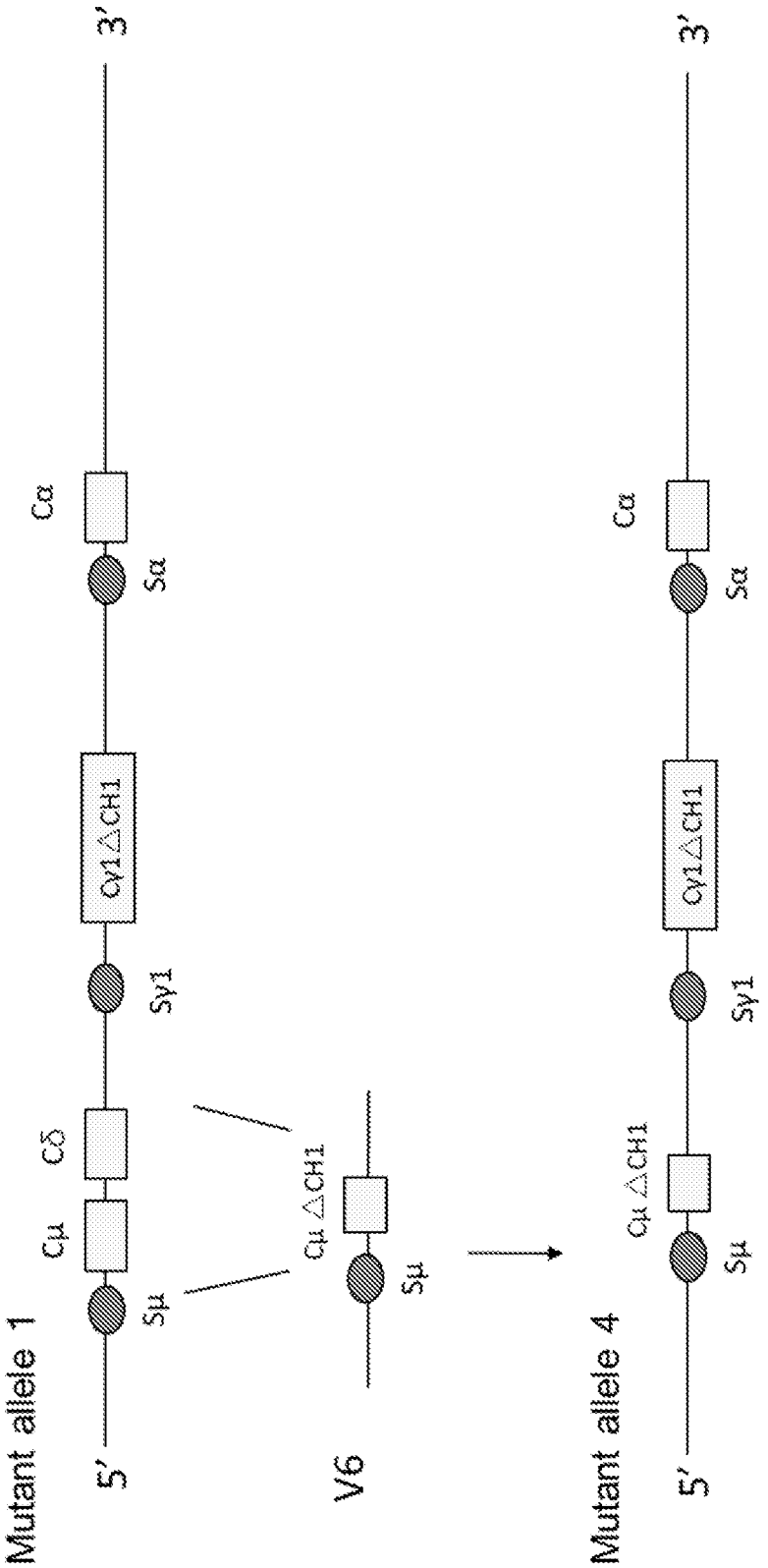

Mutant allele 1
5'
3'
Cμ
Cδ
Cγ1△CH1
Cα
Sμ
Sγ1
Sα
Cμ△CH1
Sμ
V6
Mutant allele 4
5'
3'
Cμ△CH1
Sμ
Sγ1
Cγ1△CH1
Cα
Sα

FIG. 11

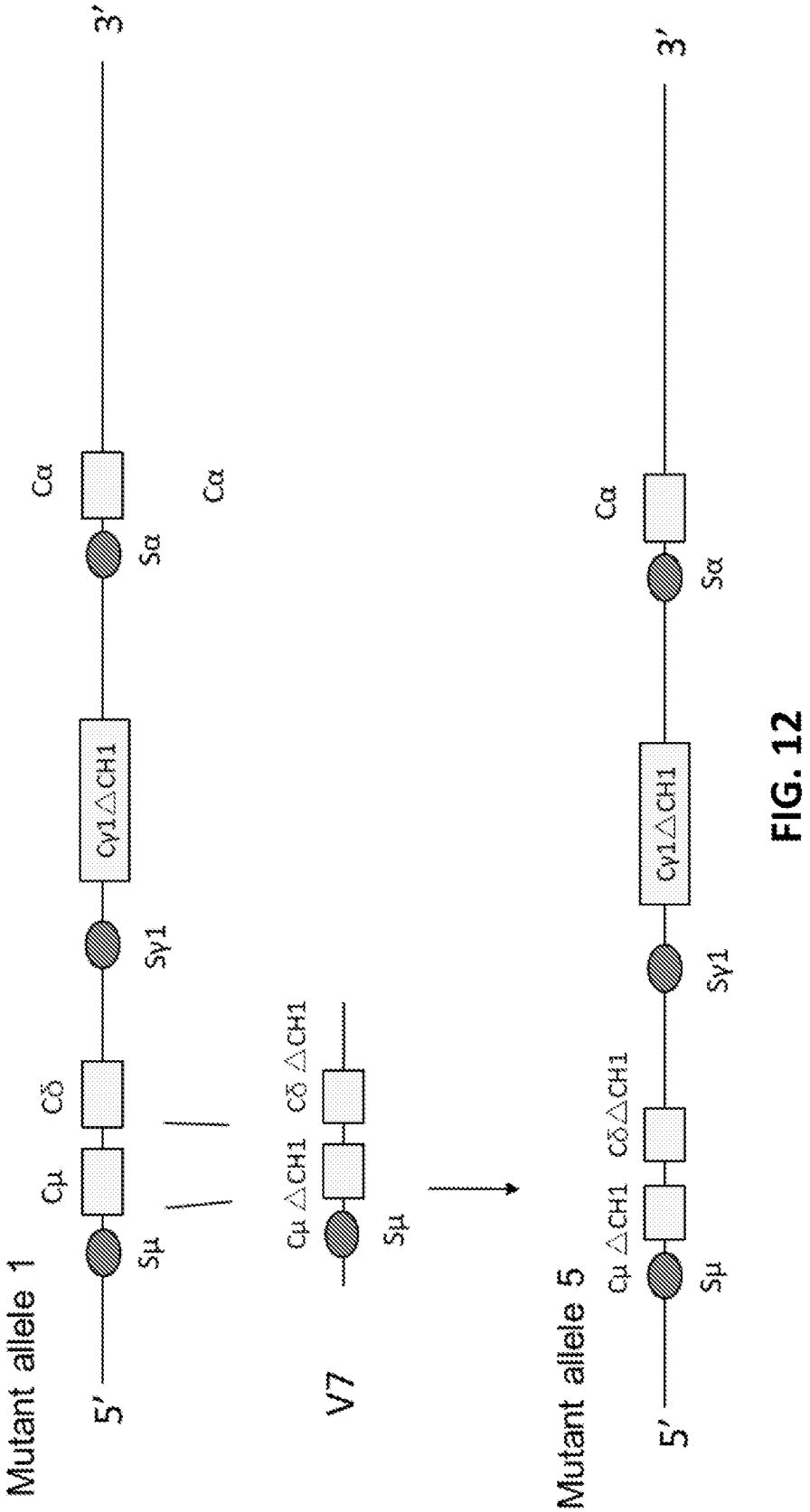
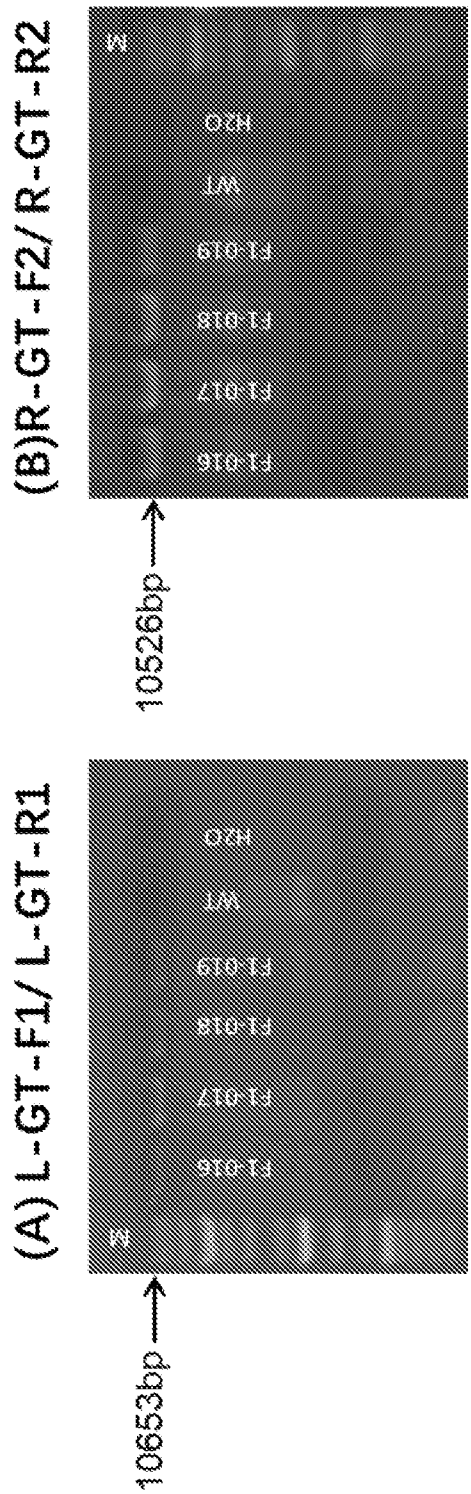
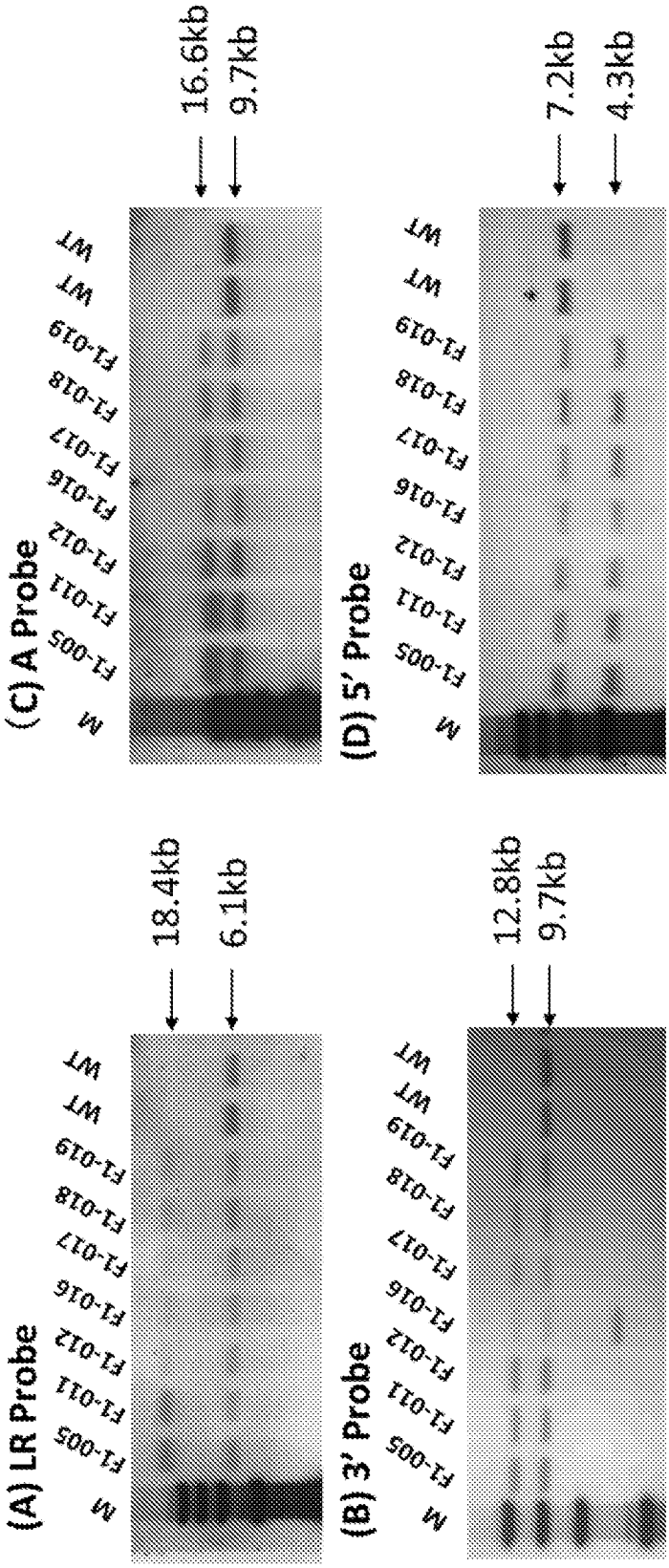

Mutant allele 1
5'
3'
Sμ
Cμ
Cδ
Sγ1
Cγ1△CH1
Sα
Cα
Cα
V7
Cμ△CH1
Cδ△CH1
Sμ
Mutant allele 5
5'
3'
Sμ
Cμ△CH1
Cδ△CH1
Sγ1
Cγ1△CH1
Sα
Cα

FIG. 12


(A) L-GT-F1/ L-GT-R1
10653bp
M
H2O
WT
F1-019
F1-018
F1-017
F1-016
(B)R-GT-F2/ R-GT-R2
10526bp
M
H2O
WT
F1-019
F1-018
F1-017
F1-016

FIGS. 13A-13B


(A) LR Probe
18.4kb
6.1kb
M
F1-005
F1-011
F1-012
F1-016
F1-017
F1-018
F1-019
WT
WT
(B) 3' Probe
12.8kb
9.7kb
M
F1-005
F1-011
F1-012
F1-016
F1-017
F1-018
F1-019
WT
WT
(C) A Probe
16.6kb
9.7kb
M
F1-005
F1-011
F1-012
F1-016
F1-017
F1-018
F1-019
WT
WT
(D) 5' Probe
7.2kb
4.3kb
M
F1-005
F1-011
F1-012
F1-016
F1-017
F1-018
F1-019
WT
WT

FIGS. 14A-14D

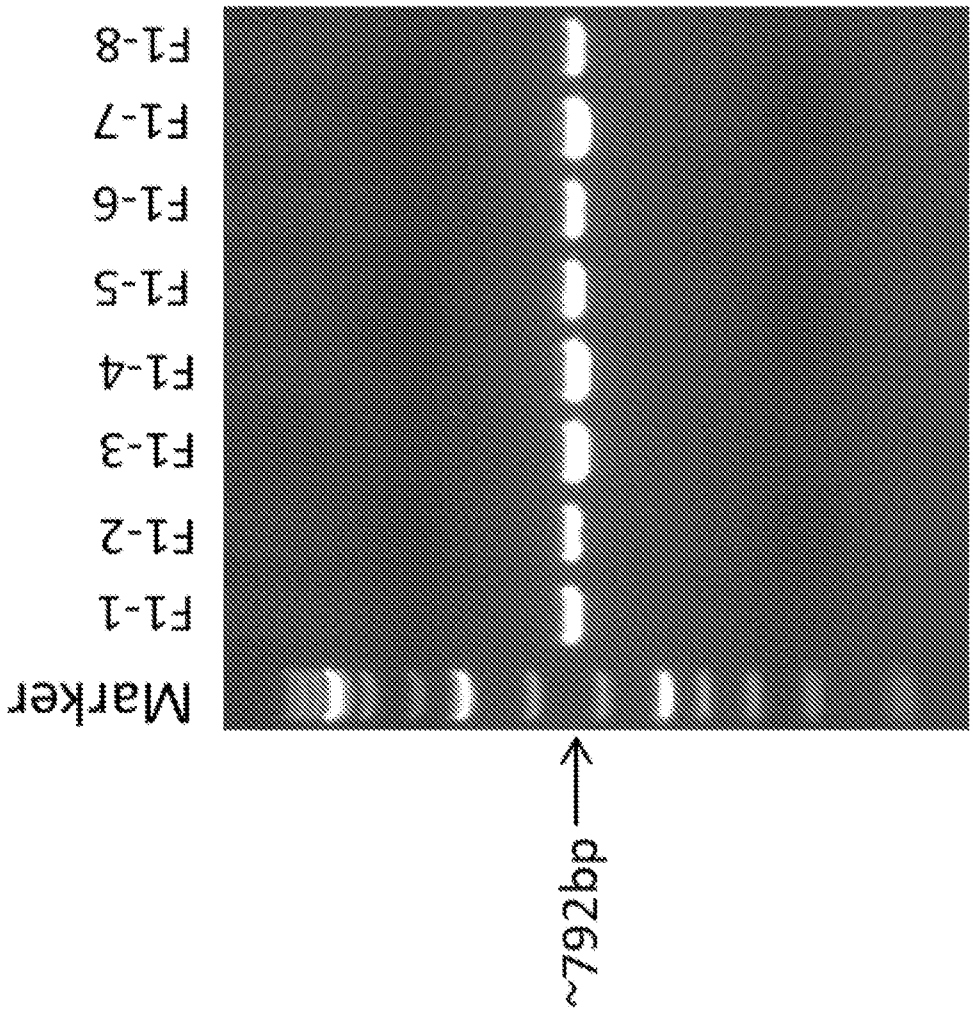


FIG. 15

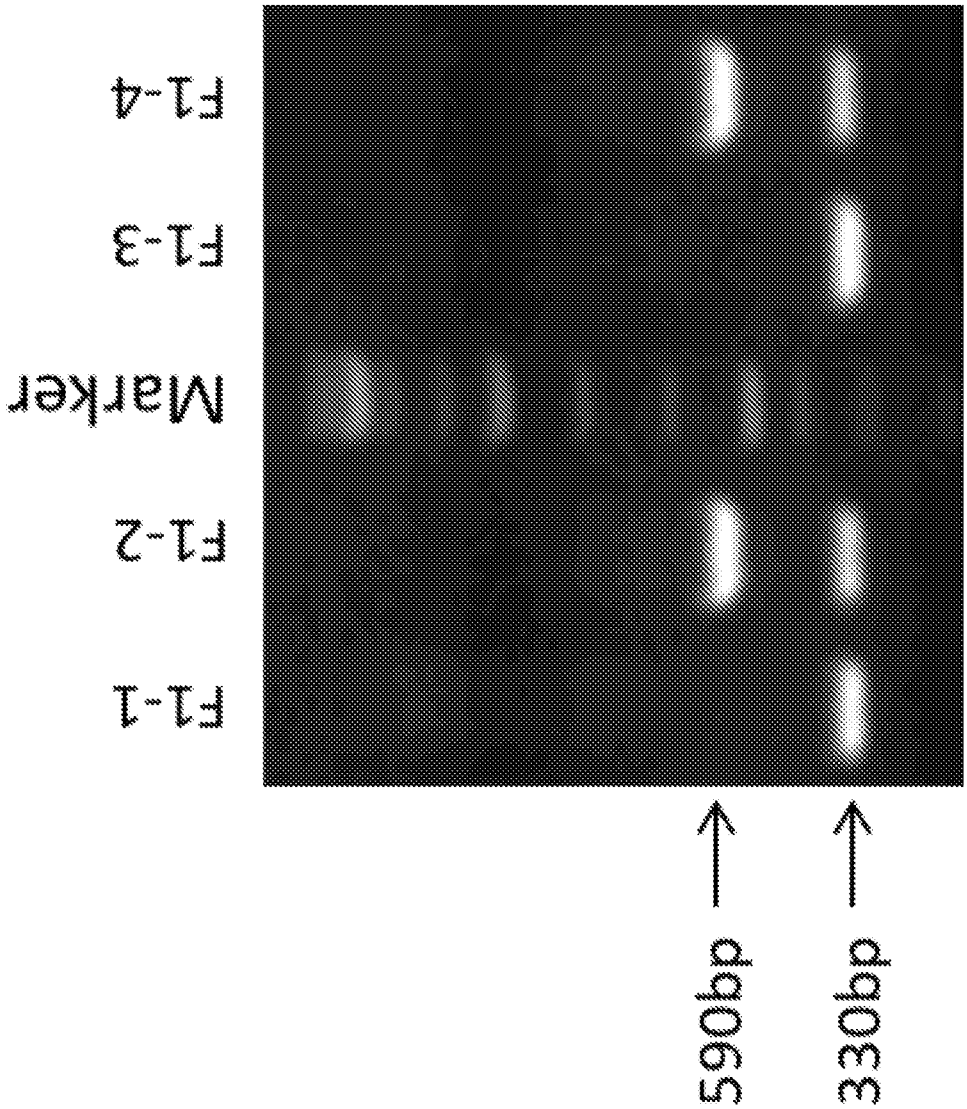


FIG. 16

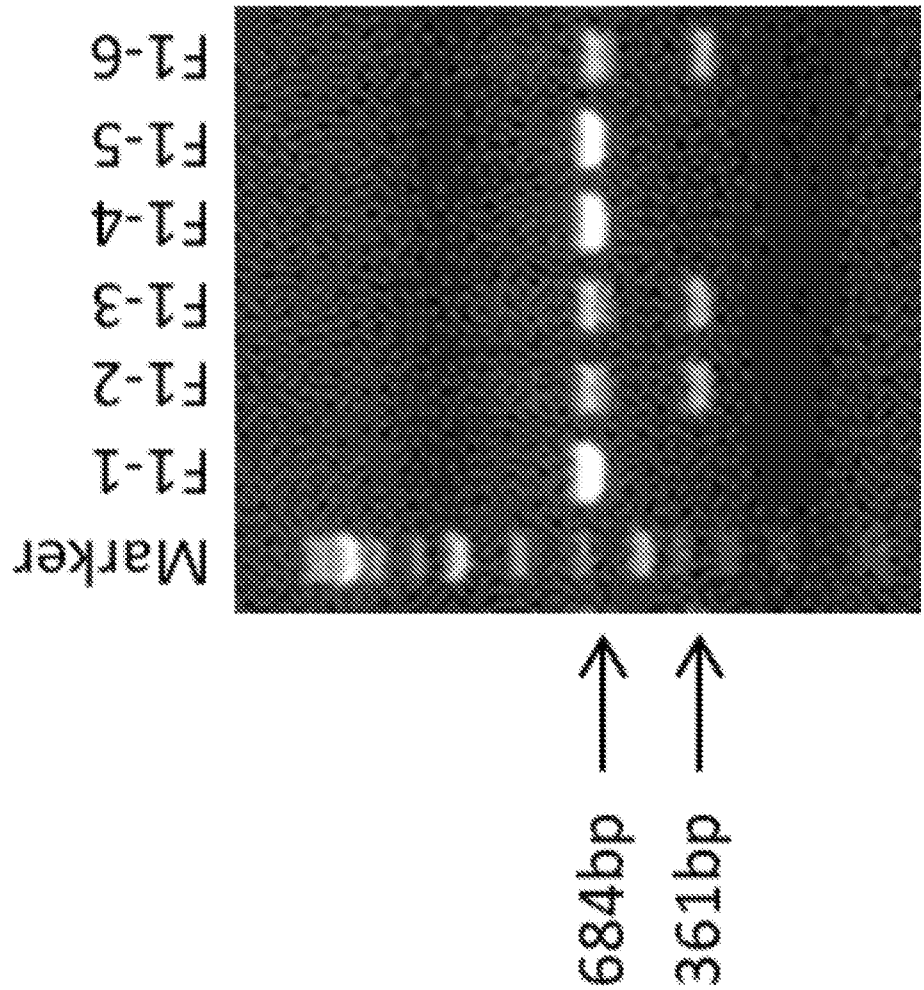


FIG. 17

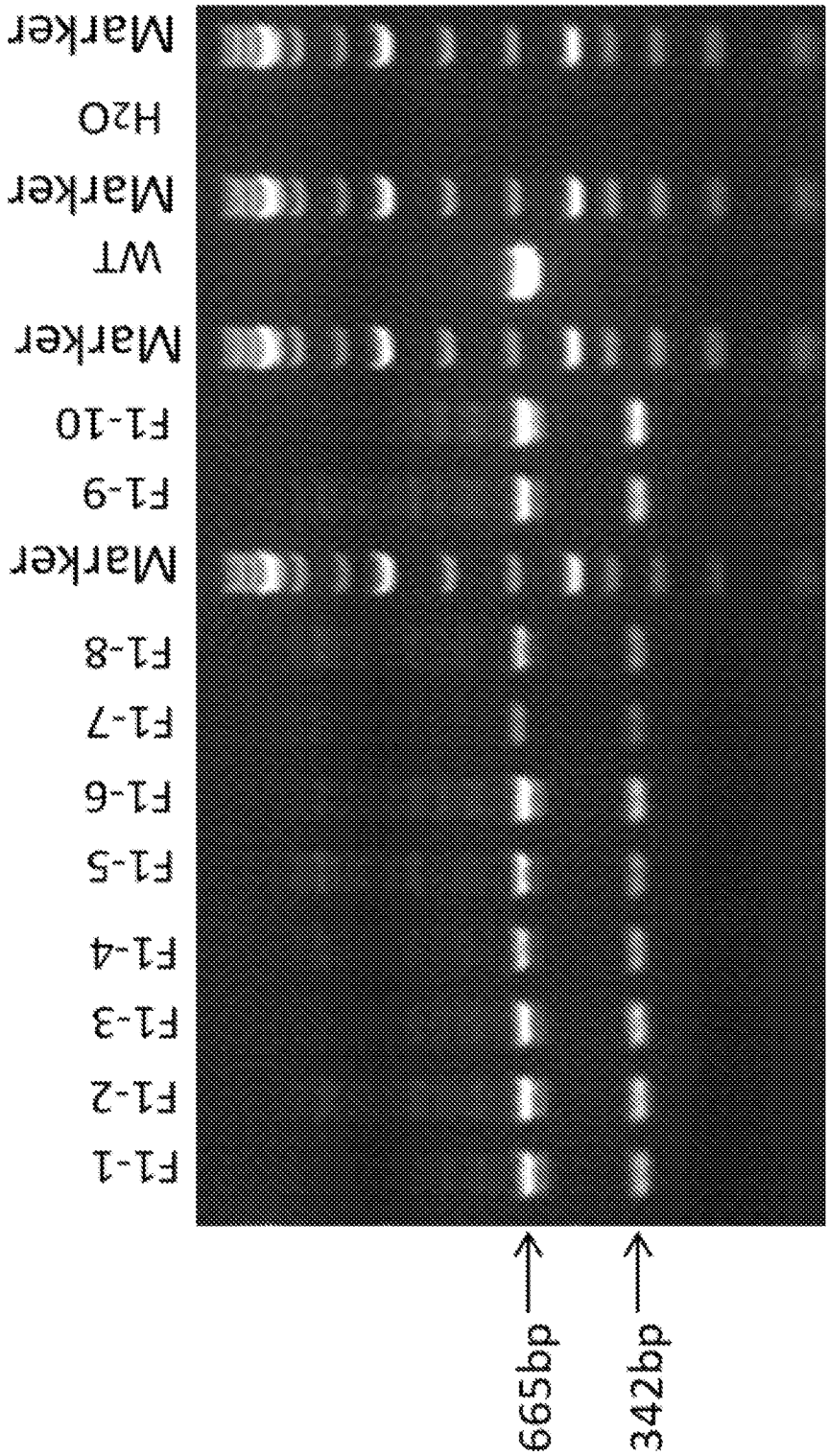

665bp
342bp
F1-1
F1-2
F1-3
F1-4
F1-5
F1-6
F1-7
F1-8
Marker
F1-9
F1-10
Marker
WT
Marker
H2O
Marker

FIG. 18A

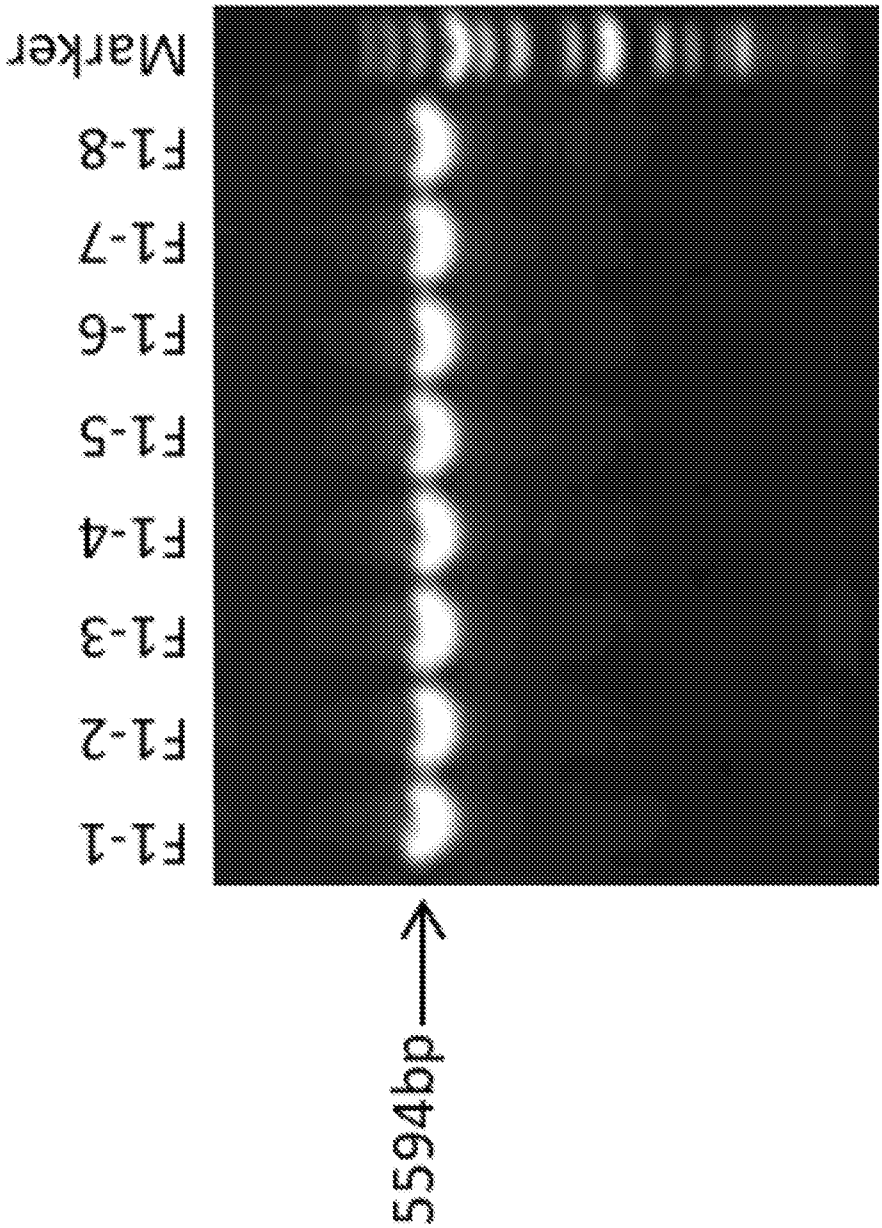
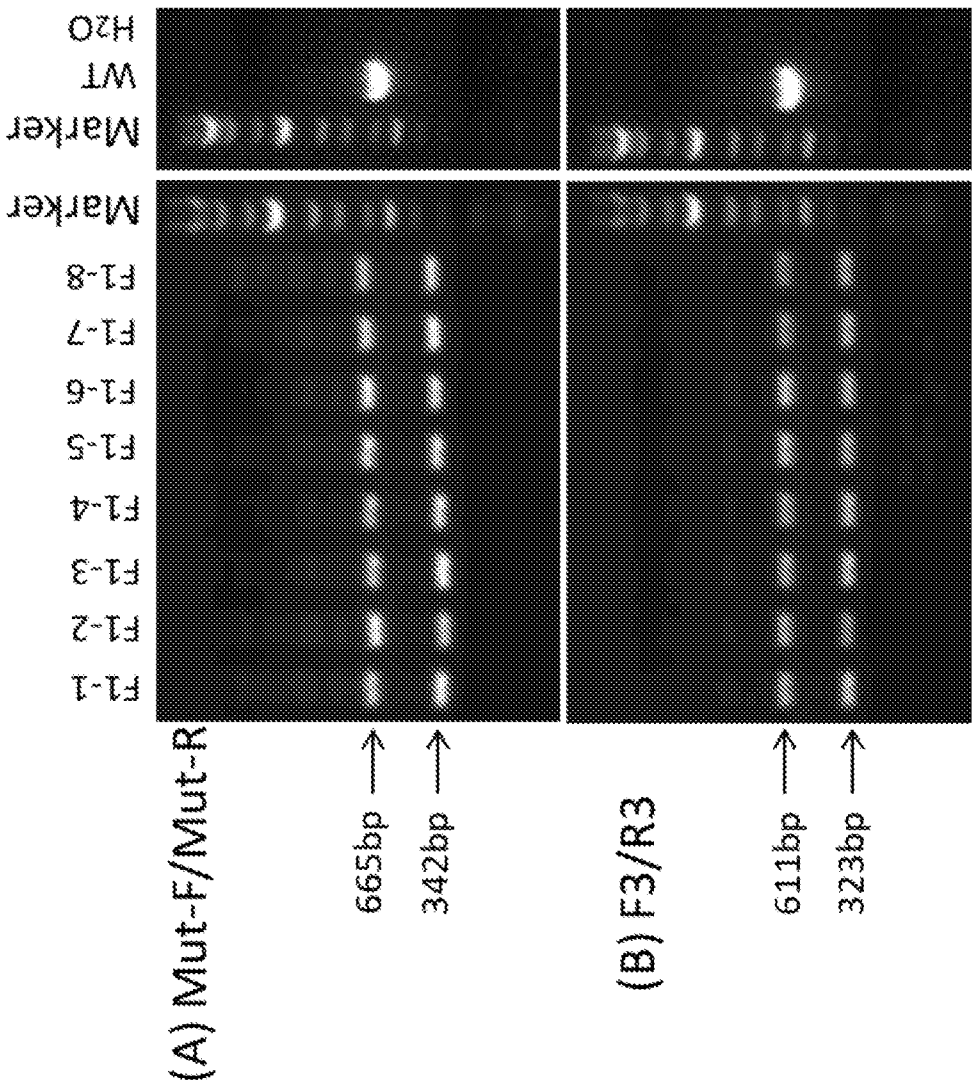


FIG. 18B


(A) Mut-F/Mut-R
665bp
342bp
(B) F3/R3
611bp
323bp
F1-1
F1-2
F1-3
F1-4
F1-5
F1-6
F1-7
F1-8
Marker
Marker
WT
H2O

FIGS. 19A-19B

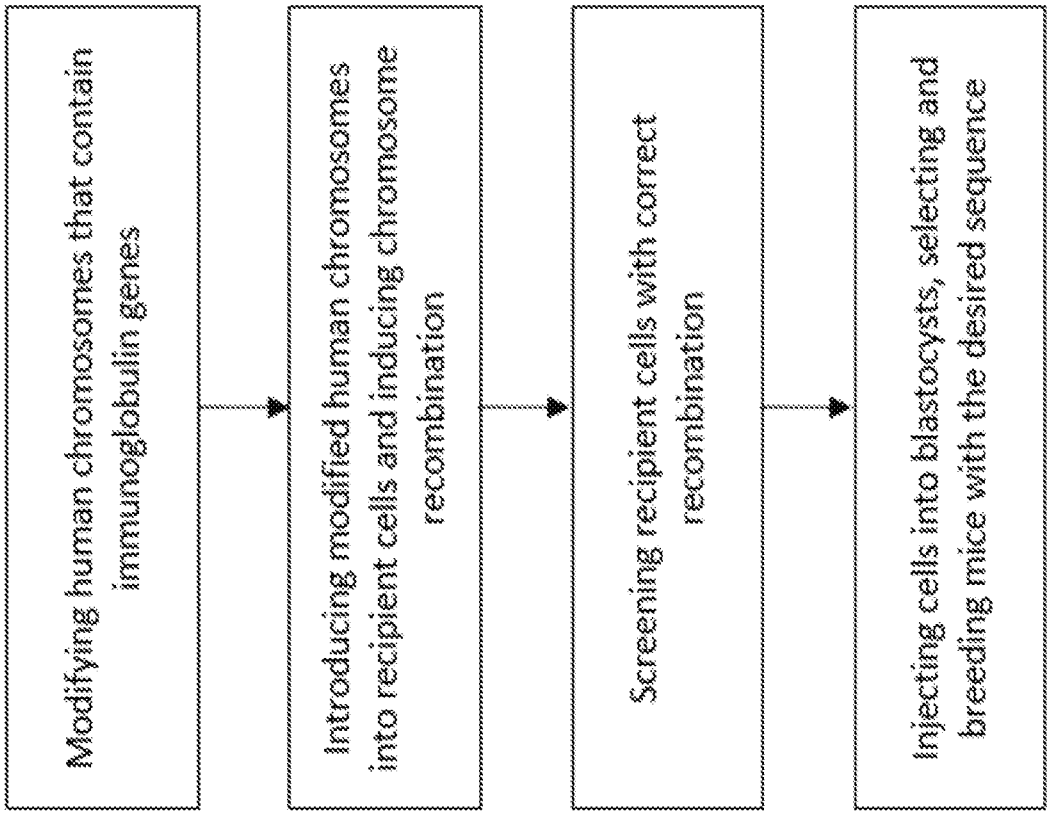

Modifying human chromosomes that contain immunoglobulin genes
Introducing modified human chromosomes into recipient cells and inducing chromosome recombination
Screening recipient cells with correct recombination
Injecting cells into blastocysts, selecting and breeding mice with the desired sequence

FIG. 20

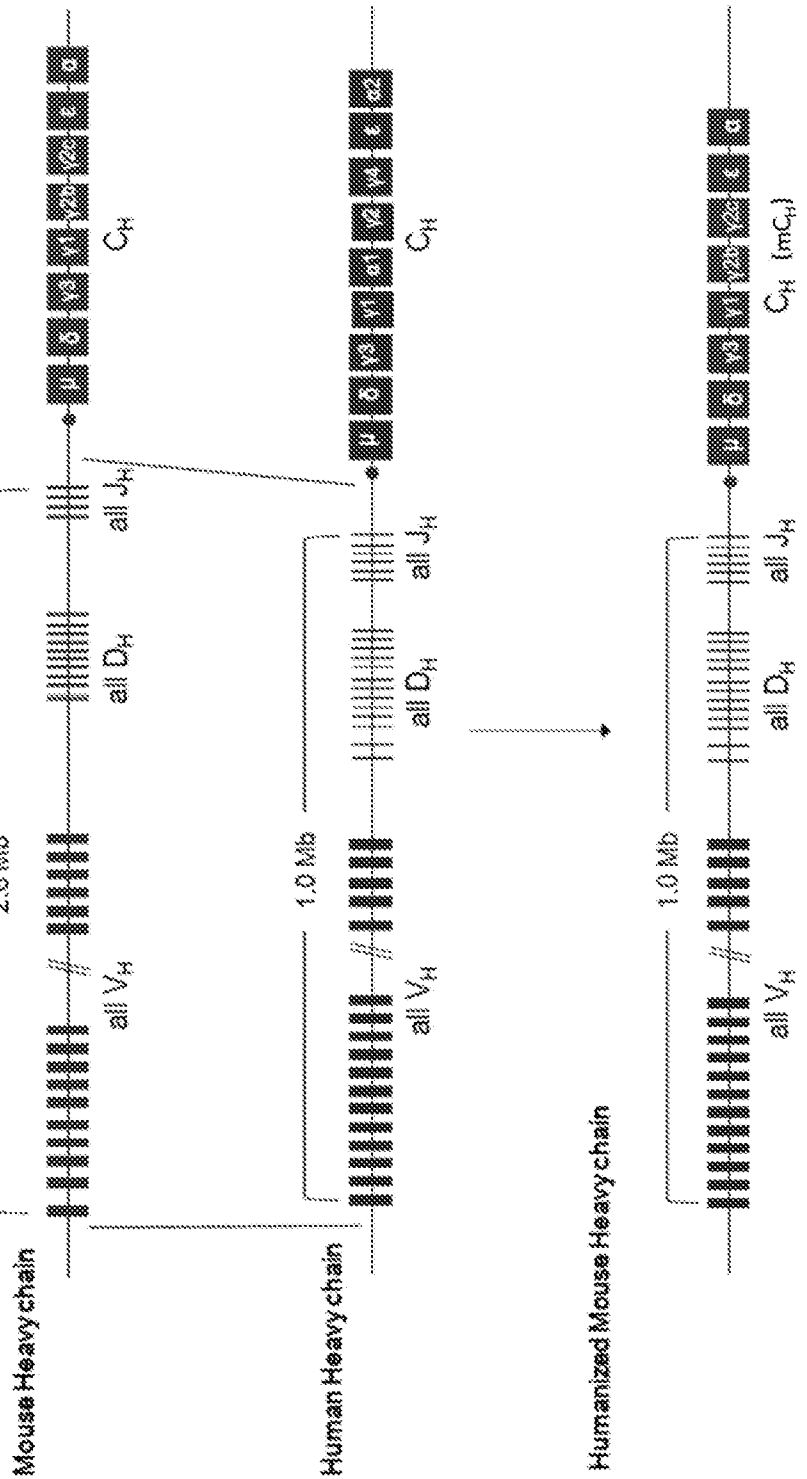

Mouse Heavy chain
2.6 Mb
all V_H
all D_H
all J_H
C_H
Human Heavy chain
1.0 Mb
all V_H
all D_H
all J_H
C_H
Humanized Mouse Heavy chain
1.0 Mb
all V_H
all D_H
all J_H
C_H (mC_H)

FIG. 21

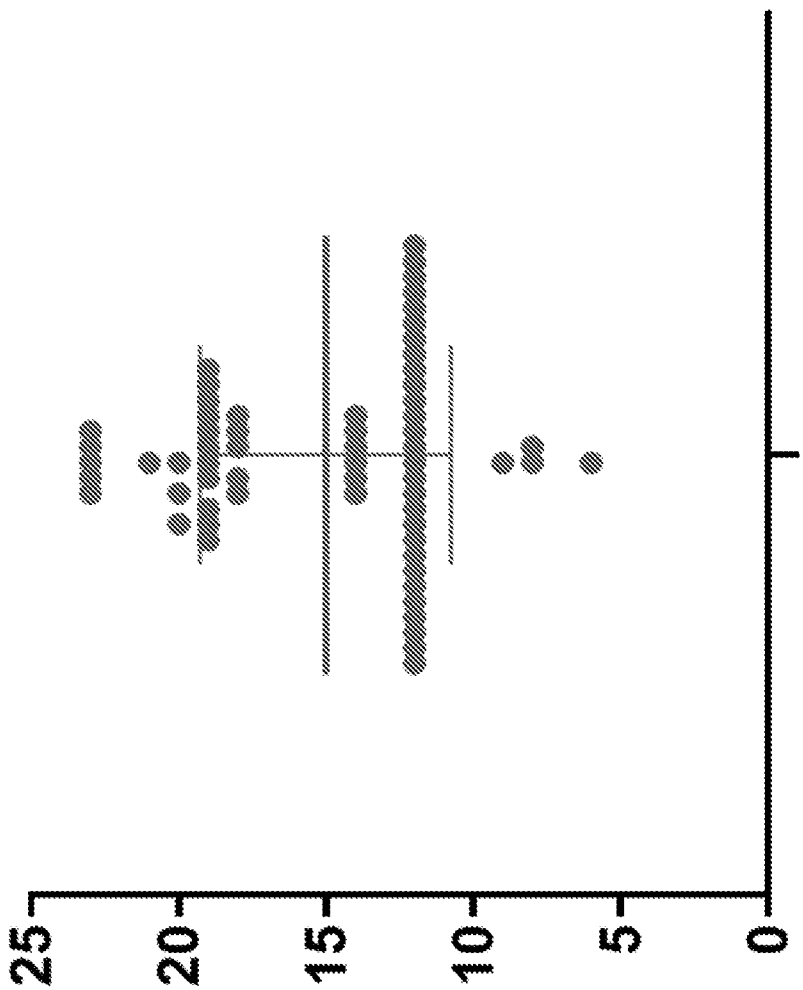


FIG. 22

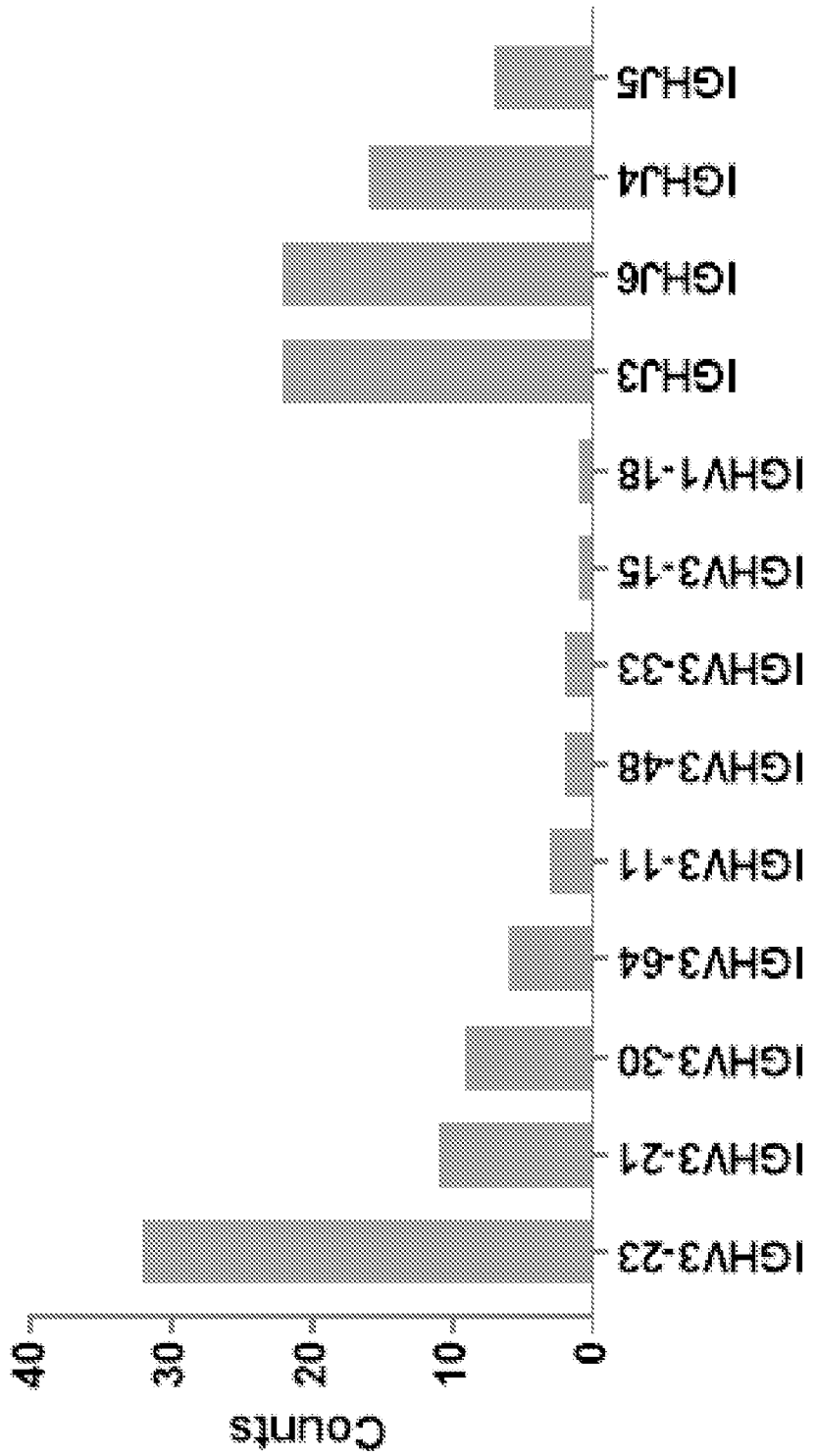

IGHV3-23
IGHV3-21
IGHV3-30
IGHV3-64
IGHV3-11
IGHV3-48
IGHV3-33
IGHV3-15
IGHV1-18
IGHJ3
IGHJ6
IGHJ4
IGHJ5
0
10
20
30
40
Counts

FIG. 23

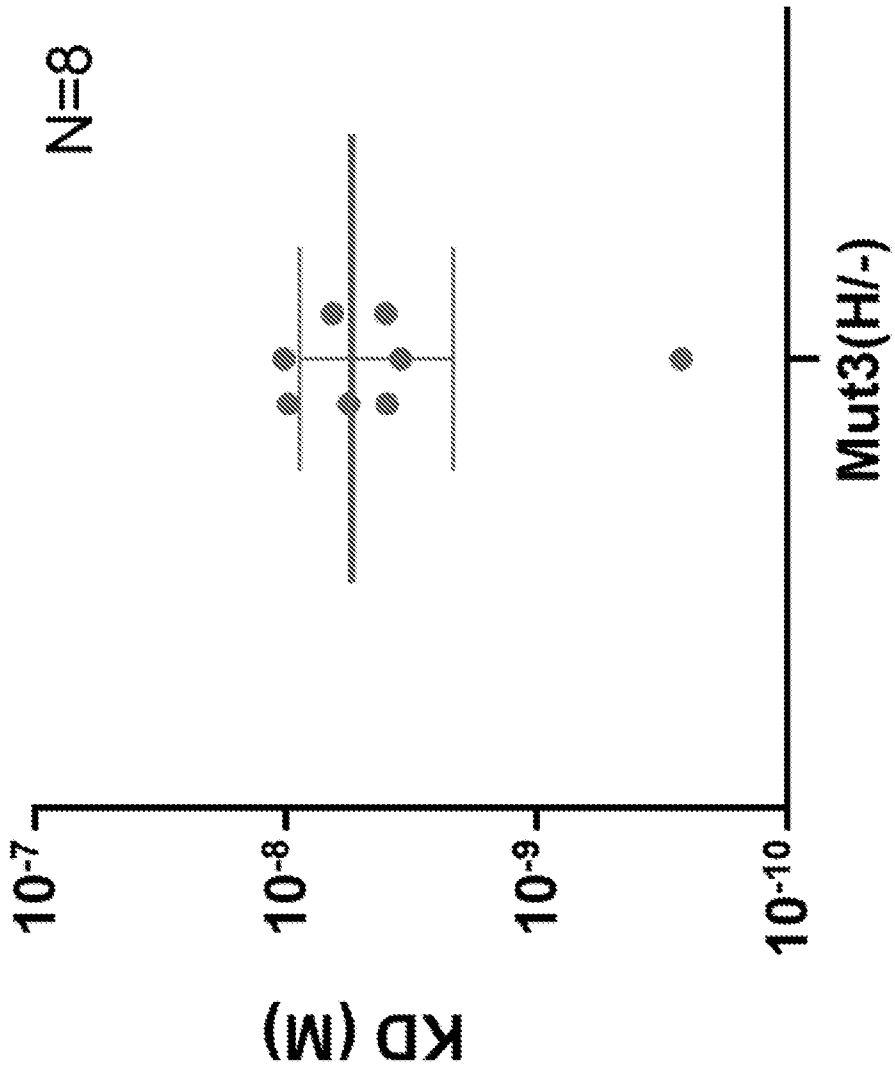


FIG. 24

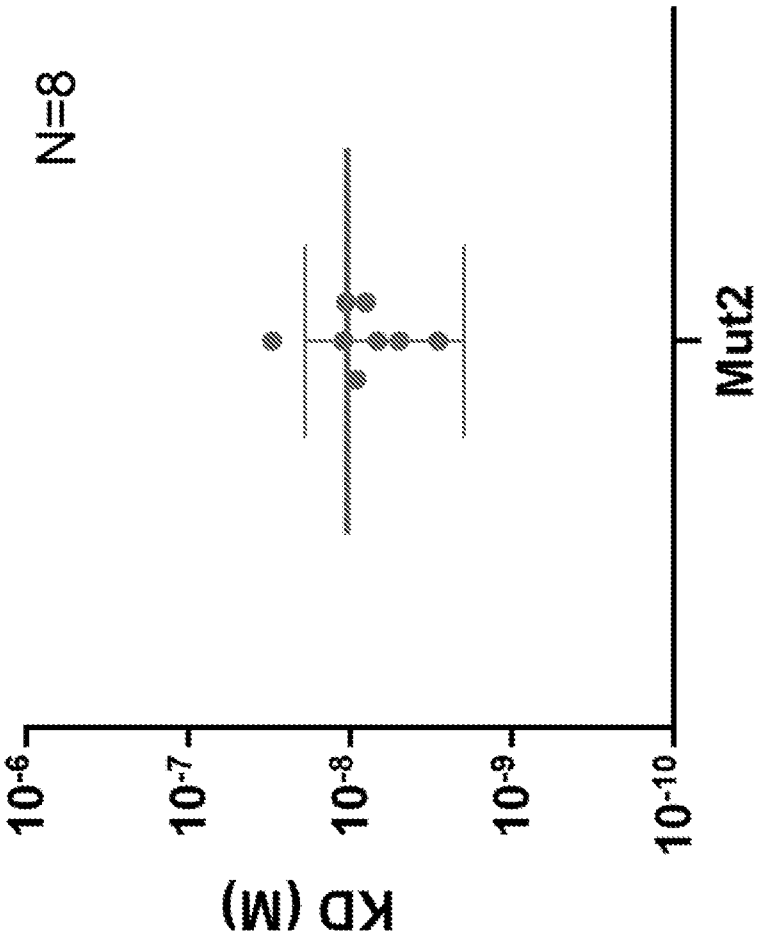


FIG. 25

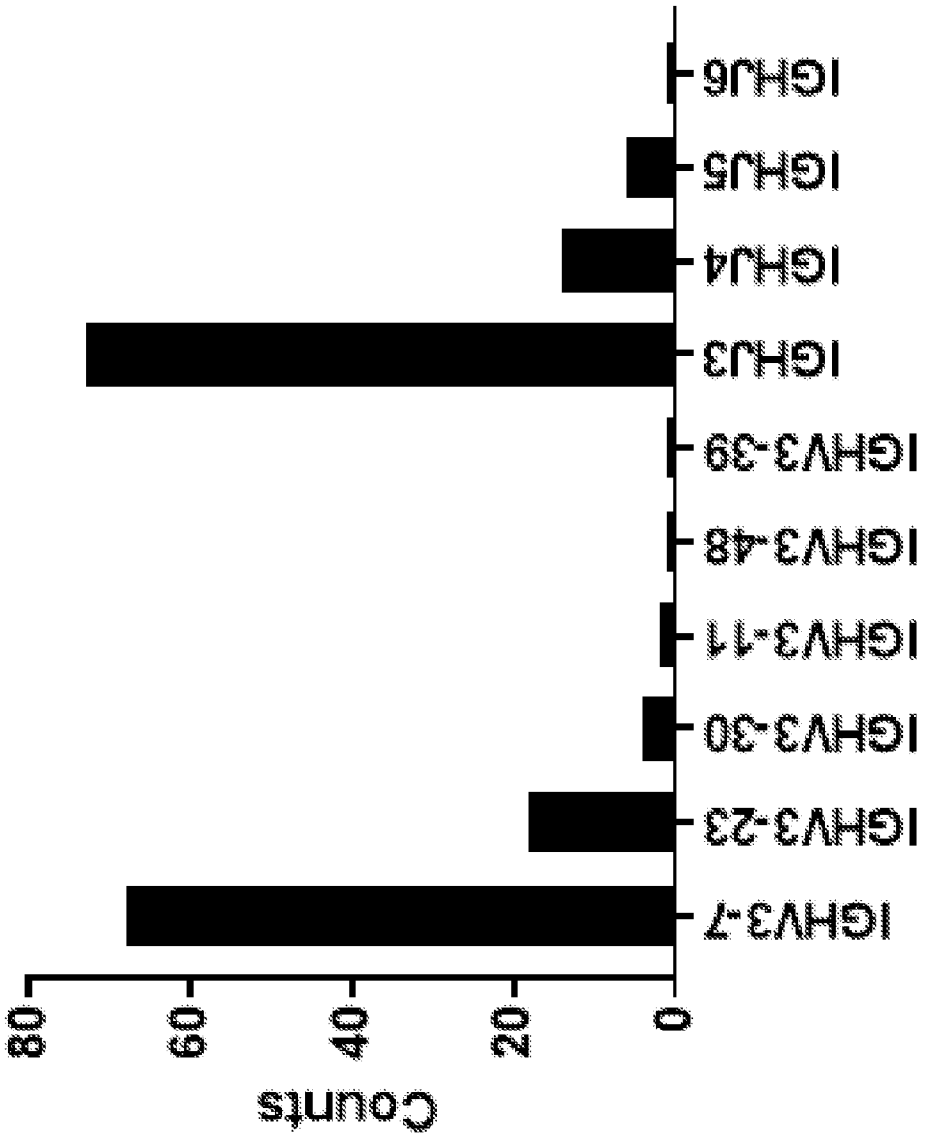

IGHV3-7
IGHV3-23
IGHV3-30
IGHV3-11
IGHV3-48
IGHV3-39
IGHJ3
IGHJ4
IGHJ5
IGHJ6
0
20
40
60
80
Counts

FIG. 26

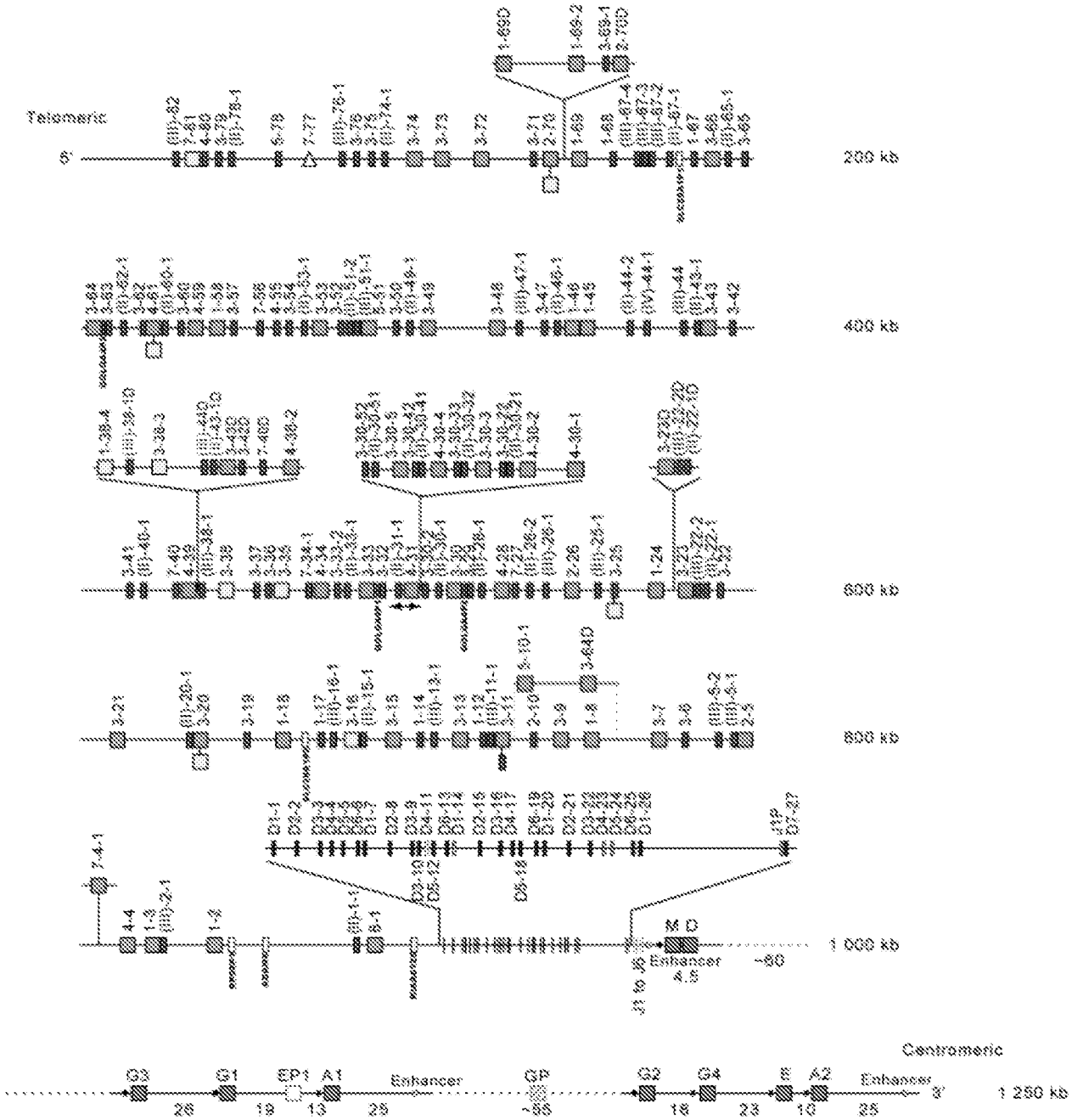

Telomeric
5'
200 kb
400 kb
600 kb
800 kb
1 000 kb
Enhancer
~80
Centromeric
G3
G1
EP1
A1
Enhancer
GP
G2
G4
E
A2
Enhancer
3'
1 250 kb
26
19
13
29
~55
18
23
10
25

FIG. 27

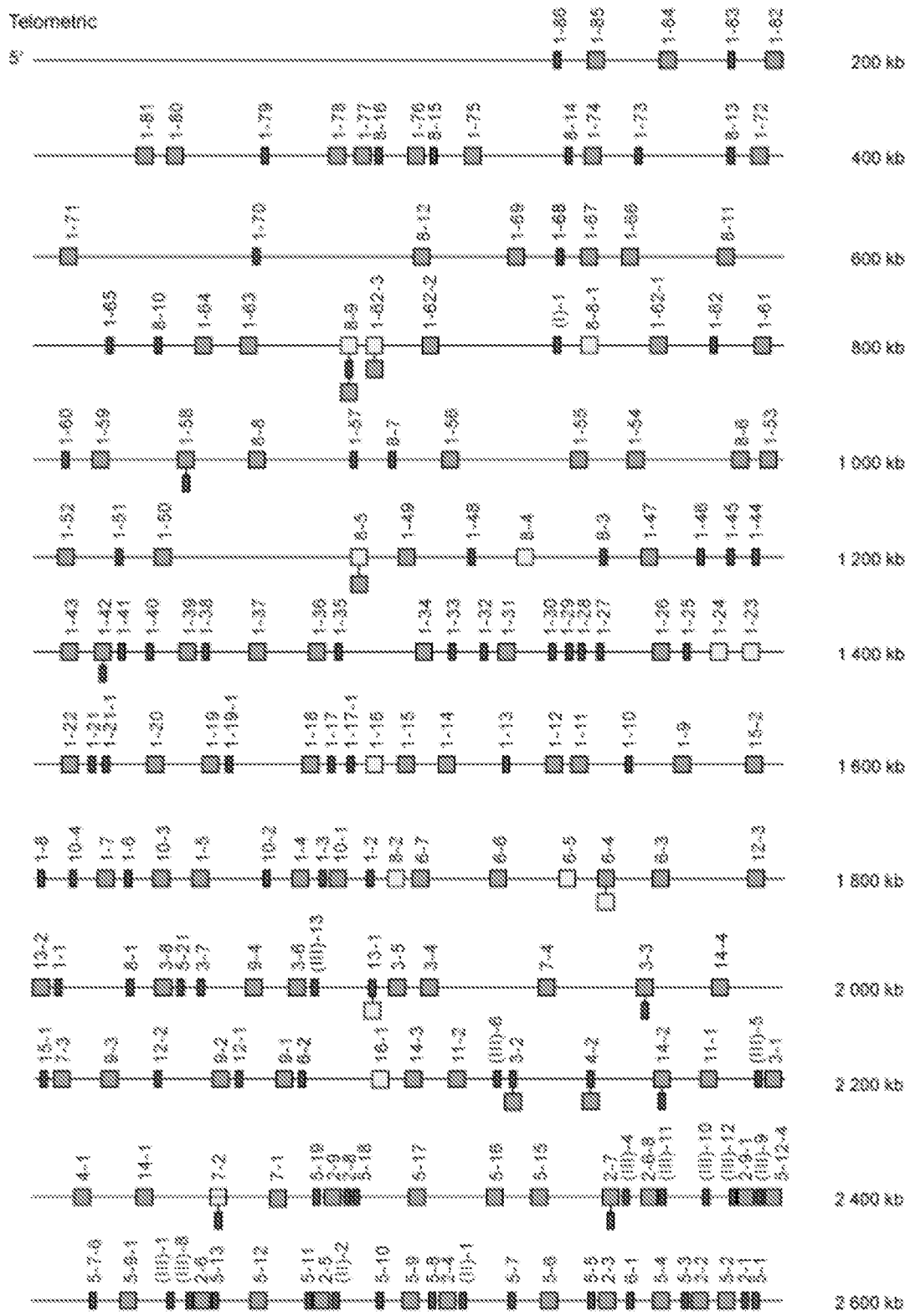

Telomeric
5'
200 kb
400 kb
600 kb
800 kb
1 000 kb
1 200 kb
1 400 kb
1 600 kb
1 800 kb
2 000 kb
2 200 kb
2 400 kb
2 600 kb

FIG. 28

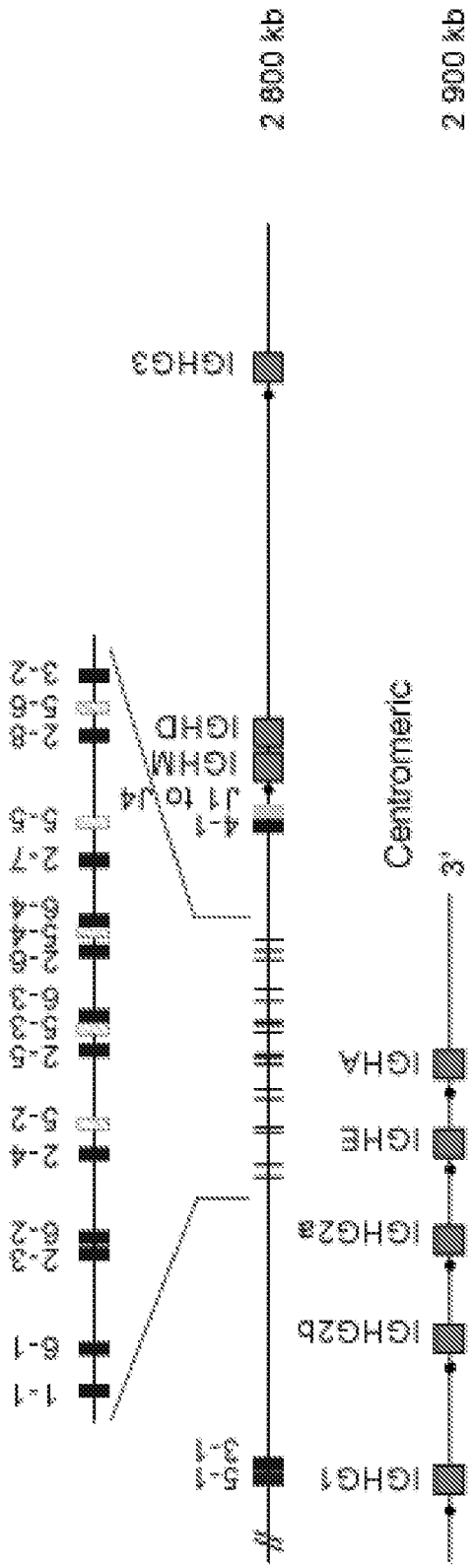


FIG. 28 (Continued)

FIG. 29

IMGT							NCBI		
Gene	Location	Orientation	Locus	Position	Accession number	Positions	Gene positions on chromosome	Accession number	Gene positions in sequences
IGHA2	14q32.33	REV	IGH		J00221	1227..1618	105124322..105124713	NT_026437.11	87054322..87054713
IGHA2	14q32.33	REV	IGH		J00221	684..1004	105124936..105125256	NT_026437.11	87054936..87055256
IGHA2	14q32.33	REV	IGH		J00221	164..469	105125474..105125776	NT_026437.11	87055474..87055776
IGHE	14q32.33	REV	IGH		J00222	2540..2868	105137451..105137779	NT_026437.11	87067451..87067779
IGHE	14q32.33	REV	IGH		J00222	2133..2456	105137863..105138186	NT_026437.11	87067863..87068186
IGHE	14q32.33	REV	IGH		J00222	1726..2046	105138273..105138593	NT_026437.11	87068273..87068593
IGHE	14q32.33	REV	IGH		J00222	1210..1518	105138801..105139109	NT_026437.11	87068801..87069109
IGHG4	14q32.33	REV	IGH		K01316	1481..1800	105161861..105162180	NT_026437.11	87091861..87092180
IGHG4	14q32.33	REV	IGH		K01316	1054..1383	105162278..105162607	NT_026437.11	87092278..87092607
IGHG4	14q32.33	REV	IGH		K01316	900..935	105162726..105162761	NT_026437.11	87092726..87092761
IGHG2	14q32.33	REV	IGH		K01316	216..509	105163154..105163447	NT_026437.11	87093154..87093447
IGHG2	14q32.33	REV	IGH		J00230	1480..1799	105180588..105180907	NT_026437.11	87110588..87110907
IGHG2	14q32.33	REV	IGH		J00230	1056..1382	105181005..105181331	NT_026437.11	87111005..87111331
IGHG2	14q32.33	REV	IGH		J00230	902..937	105181450..105181485	NT_026437.11	87111450..87111485
IGHG2	14q32.33	REV	IGH		J00230	216..509	105181878..105182171	NT_026437.11	87111878..87112171
IGHGP	14q32.33	REV	IGH		X06766	1474..1790	105205635..105205947	NT_026437.11	87135635..87135947
IGHGP	14q32.33	REV	IGH		X06766	1047..1376	105206045..105206374	NT_026437.11	87136045..87136374
IGHGP	14q32.33	REV	IGH		X06766	878..928	105206493..105206543	NT_026437.11	87136493..87136543
IGHGP	14q32.33	REV	IGH		X06766	195..488	105206933..105207226	NT_026437.11	87136933..87137226
IGHA1	14q32.33	REV	IGH		J00220	1244..1635	105244553..105244944	NT_026437.11	87174553..87174944
IGHA1	14q32.33	REV	IGH		J00220	662..1021	105245167..105245526	NT_026437.11	87175167..87175526
IGHA1	14q32.33	REV	IGH		J00220	142..447	105245741..105246046	NT_026437.11	87175741..87176046
IGHEP1	14q32.33	REV	IGH		J00223	849..1177	105259179..105259507	NT_026437.11	87189179..87189507
IGHEP1	14q32.33	REV	IGH		J00223	447..765	105259591..105259909	NT_026437.11	87189591..87189909
IGHG1	14q32.33	REV	IGH		J00228	1481..1800	105278858..105279177	NT_026437.11	87208858..87209177
IGHG1	14q32.33	REV	IGH		J00228	1055..1384	105279275..105279604	NT_026437.11	87209275..87209604
IGHG1	14q32.33	REV	IGH		J00228	892..936	105279723..105279767	NT_026437.11	87209723..87209767
IGHG1	14q32.33	REV	IGH		J00228	210..503	105280159..105280452	NT_026437.11	87210159..87210452
IGHG3	14q32.33	REV	IGH		X03604	2061..2380	105306622..105306941	NT_026437.11	87236622..87236941
IGHG3	14q32.33	REV	IGH		X03604	1634..1963	105307039..105307368	NT_026437.11	87237039..87237368
IGHG3	14q32.33	REV	IGH		X03604	1471..1515	105307487..105307531	NT_026437.11	87237487..87237531
IGHG3	14q32.33	REV	IGH		X03604	1283..1327	105307675..105307719	NT_026437.11	87237675..87237719
IGHG3	14q32.33	REV	IGH		X03604	1095..1139	105307863..105307907	NT_026437.11	87237863..87237907
IGHG3	14q32.33	REV	IGH		X03604	901..951	105308051..105308101	NT_026437.11	87238051..87238101
IGHG3	14q32.33	REV	IGH		X03604	216..509	105308494..105308787	NT_026437.11	87238494..87238787
IGHD	14q32.33	REV	IGH		K02880	25..50	105375785..105375810	NT_026437.11	87305785..87305810
IGHD	14q32.33	REV	IGH		K02879	101..424	105377748..105378071	NT_026437.11	87307748..87308071
IGHD	14q32.33	REV	IGH		K02878	101..424	105378282..105378605	NT_026437.11	87308282..87308605
IGHD	14q32.33	REV	IGH		K02877	101..172	105379342..105379413	NT_026437.11	87309342..87309413
IGHD	14q32.33	REV	IGH		K02876	101..202	105382203..105382304	NT_026437.11	87312203..87312304

FIG. 29 (Continued)

IGHD	14q32.33	REV	IGH		K02875	101..403	105382753..105383055	NT_026437.11	87312753..87313055
IGHM	14q32.33	REV	IGH		X14940	1708..2096	105391496..105391887	NT_026437.11	87321496..87321887
IGHM	14q32.33	REV	IGH		X14940	1212..1529	105392067..105392384	NT_026437.11	87322067..87322384
IGHM	14q32.33	REV	IGH		X14940	631..966	105392630..105392965	NT_026437.11	87322630..87322965
IGHM	14q32.33	REV	IGH		X14940	233..544	105393056..105393367	NT_026437.11	87323056..87323367
IGHJ6	14q32.33	REV	IGH	960175..960275	J00256	2909..3009	105400453..105400553	NT_026437.11	87330453..87330553
IGHJ3P	14q32.33	REV	IGH	959970..960057	J00256	2709..2797	105400671..105400758	NT_026437.11	87330671..87330758
IGHJ5	14q32.33	REV	IGH	959571..959659	J00256	2317..2404	105401084..105401156	NT_026437.11	87331069..87331141
IGHJ4	14q32.33	REV	IGH	959172..959258	J00256	1873..1959	105401470..105401556	NT_026437.11	87331470..87331556
IGHJ3	14q32.33	REV	IGH	958789..958886	J00256	1498..1586	105401865..105401930	NT_026437.11	87331842..87331907
IGHJ2P	14q32.33	REV	IGH	958583..958682	J00256	1284..1380	105402046..105402145	NT_026437.11	87332045..87332145
IGHJ2	14q32.33	REV	IGH	958184..958274	J00256	893..984	105402454..105402544	NT_026437.11	87332454..87332544
IGHJ1	14q32.33	REV	IGH	957977..958066	J00256	686..774	105402662..105402751	NT_026437.11	87332662..87332751
IGHD7-27	14q32.33	REV	IGH	957884..957950	J00256	592..659	105402781..105402845	NT_026437.11	87332778..87332842
IGHJ1P	14q32.33	REV	IGH	957759..957849	J00256	467..557	105402879..105402969	NT_026437.11	87332879..87332969
IGHD1-26	14q32.33	REV	IGH	942598..942673	X97051	72141..72216	105417909..105417984	NT_026437.11	87347909..87347984
IGHD6-25	14q32.33	REV	IGH	942095..942168	X97051	71638..71711	105418414..105418487	NT_026437.11	87348414..87348487
IGHD5-24	14q32.33	REV	IGH	939730..939805	X97051	69272..69347	105420778..105420853	NT_026437.11	87350778..87350853
IGHD4-23	14q32.33	REV	IGH	938765..938839	X97051	68306..68380	105421745..105421819	NT_026437.11	87351745..87351819
IGHD3-22	14q32.33	REV	IGH	937592..937678	X93616	2..88	105422906..105422992	NT_026437.11	87352906..87352992
IGHD2-21	14q32.33	REV	IGH	935075..935158	J00235	1..63	105425447..105425509	NT_026437.11	87355447..87355509
IGHD1-20	14q32.33	REV	IGH	932446..932518	X97051	61987..62059	105428066..105428138	NT_026437.11	87358066..87358138
IGHD6-19	14q32.33	REV	IGH	931934..932010	X97051	61475..61551	105428574..105428650	NT_026437.11	87358574..87358650
IGHD5-18	14q32.33	REV	IGH	930093..930168	X13972	2865..2960	105430417..105430492	NT_026437.11	87360417..87360492
IGHD4-17	14q32.33	REV	IGH	929131..929202	X97051	58671..58742	105431383..105431454	NT_026437.11	87361383..87361454
IGHD3-16	14q32.33	REV	IGH	927984..928076	X93614	2..94	105432511..105432601	NT_026437.11	87362511..87362601
IGHD2-15	14q32.33	REV	IGH	925668..925753	J00234	1..86	105434853..105434918	NT_026437.11	87364853..87364918
IGHD1-14	14q32.33	REV	IGH	923001..923073	X13972	14490..14562	105437513..105437585	NT_026437.11	87367513..87367585
IGHD6-13	14q32.33	REV	IGH	922493..922569	X13972	13983..14059	105438017..105438093	NT_026437.11	87368017..87368093
IGHD5-12	14q32.33	REV	IGH	920984..921062	X13972	12478..12556	105439524..105439602	NT_026437.11	87369524..87369602
IGHD4-11	14q32.33	REV	IGH	920024..920095	X13972	11522..11593	105440491..105440562	NT_026437.11	87370491..87370562
IGHD3-10	14q32.33	REV	IGH	919128..919214	X13972	10631..10717	105441372..105441458	NT_026437.11	87371372..87371458
IGHD3-9	14q32.33	REV	IGH	918944..919030	X13972	10447..10533	105441556..105441642	NT_026437.11	87371556..87371642
IGHD2-8	14q32.33	REV	IGH	916414..916500	X13972	7921..8007	105444086..105444172	NT_026437.11	87374086..87374172
IGHD1-7	14q32.33	REV	IGH	913731..913803	X13972	5238..5310	105446783..105446855	NT_026437.11	87376783..87376855
IGHD6-6	14q32.33	REV	IGH	913227..913300	X13972	4734..4807	105447287..105447359	NT_026437.11	87377286..87377358
IGHD5-5	14q32.33	REV	IGH	911378..911453	X97051	59633..59708	105449133..105449208	NT_026437.11	87379133..87379208
IGHD4-4	14q32.33	REV	IGH	910417..910488	X13972	1924..1995	105450098..105450169	NT_026437.11	87380098..87380169
IGHD3-3	14q32.33	REV	IGH	909265..909351	X13972	776..862	105451235..105451321	NT_026437.11	87381235..87381321
IGHD2-2	14q32.33	REV	IGH	906798..906884	J00232	1..86	105453723..105453788	NT_026437.11	87383723..87383788
IGHD1-1	14q32.33	REV	IGH	904136..904208	X97051	33686..33758	105456378..105456450	NT_026437.11	87386378..87386450
IGHV4-30-2	14q32.33	REV	IGH		L10089	1..297	105476885..105476952	NT_026437.11	87406660..87406727
IGHV4-4	14q32.33	REV	IGH	811002..811472	L10098	1..325	105476906..105476984	NT_026437.11	87406894..87406972
IGHV(II)-1-1	14q32.33	REV	IGH	878294..878475	AB019441	78294..78475	105482111..105482292	NT_026437.11	87412111..87412292
IGHV3-21	14q32.33	REV	IGH	597417..597909	AB019439	197417..197909	106282330..106282428	NT_026437.11	88211936..88212034

FIG. 29 (Continued)

IGHV1-2	14q32.33	REV	IGH	836435..836911	X07448	126..603	105523675..105524151	NT_026437.11	87453675..87454151
IGHV3-41	14q32.33	REV	IGH	390047..390538	M99670	189..680	106119868..106120165	NT_026437.11	88049868..88050165
IGHV1-3	14q32.33	REV	IGH	817860..818336	X62109	21..497	105542250..105542726	NT_026437.11	87472250..87472726
IGHV4-59	14q32.33	REV	IGH	205856..206326	AB019438	1..291	106154301..106154591	NT_026437.11	88084301..88084591
IGHV2-5	14q32.33	REV	IGH	794964..795446	X62111	71..553	105565140..105565622	NT_026437.11	87495140..87495622
IGHV(III)-5-1	14q32.33	REV	IGH	793545..793643	AB019440	193545..193643	105566943..105567041	NT_026437.11	87496943..87497041
IGHV(III)-5-2	14q32.33	REV	IGH	787450..787749	AB019440	187450..187749	105572837..105573136	NT_026437.11	87502837..87503136
IGHV3-6	14q32.33	REV	IGH	777310..777803	M99650	13..506	105582783..105583276	NT_026437.11	87512783..87513276
IGHV3-7	14q32.33	REV	IGH	770688..771182	M99649	184..678	105589404..105589898	NT_026437.11	87519404..87519898
IGHV6-1	14q32.33	REV	IGH	883485..883971	J04097	337..823	105763091..105763163	NT_026437.11	87692677..87692749
IGHV3-9	14q32.33	REV	IGH	736812..737296	M99651	132..616	105623290..105623774	NT_026437.11	87553290..87553774
IGHV2-10	14q32.33	REV	IGH	729561..730040	M99647	71..550	105630549..105631025	NT_026437.11	87560549..87561025
IGHV3-11	14q32.33	REV	IGH	715861..716349	M99652	48..536	105644237..105644725	NT_026437.11	87574237..87574725
IGHV(III)-11-1	14q32.33	REV	IGH	712240..712656	AB019440	112240..112656	105647930..105648346	NT_026437.11	87577930..87578346
IGHV1-12	14q32.33	REV	IGH	710238..710527	X92210	191..628	105649911..105650348	NT_026437.11	87579911..87580348
IGHV3-13	14q32.33	REV	IGH	702954..703445	X92217	2..493	105657141..105657632	NT_026437.11	87587141..87587632
IGHV(III)-13-1	14q32.33	REV	IGH	690249..690586	AB019440	90249..90586	105670000..105670337	NT_026437.11	87600000..87600337
IGHV1-14	14q32.33	REV	IGH	686675..687009	AB019440	86675..87009	105673577..105673911	NT_026437.11	87603577..87603911
IGHV3-15	14q32.33	REV	IGH	678769..679269	X92216	2..502	105681317..105681817	NT_026437.11	87611317..87611817
IGHV(II)-15-1	14q32.33	REV	IGH	669050..669320	AB019440	69050..69320	105691266..105691536	NT_026437.11	87621266..87621536
IGHV3-16	14q32.33	REV	IGH	667196..667688	M99655	30..522	105692898..105693390	NT_026437.11	87622898..87623390
IGHV(III)-16-1	14q32.33	REV	IGH	661801..662105	AB019440	61801..62105	105698481..105698785	NT_026437.11	87628481..87628785
IGHV1-17	14q32.33	REV	IGH	658074..658549	M99640	45..521	105702037..105702512	NT_026437.11	87632037..87632512
IGHV1-18	14q32.33	REV	IGH	647544..648019	M99641	47..522	105712567..105713042	NT_026437.11	87642567..87643042
IGHV3-19	14q32.33	REV	IGH	635792..636211	M99656	211..630	105724375..105724794	NT_026437.11	87654375..87654794
IGHV3-20	14q32.33	REV	IGH	621508..622001	M99657	10..504	105738585..105739078	NT_026437.11	87668585..87669078
IGHV(II)-20-1	14q32.33	REV	IGH	619576..620076	AB019440	19576..20076	105740510..105741010	NT_026437.11	87670510..87671010
IGHV1-58	14q32.33	REV	IGH	210733..211209	M29809	151..627	106149377..106149853	NT_026437.11	88079377..88079853
IGHV3-22	14q32.33	REV	IGH	574720..575220	M99659	85..585	105785366..105785866	NT_026437.11	87715366..87715866
IGHV(II)-22-1	14q32.33	REV	IGH	569320..569588	AB019439	169320..169588	105790998..105791266	NT_026437.11	87720998..87721266
IGHV(III)-22-2	14q32.33	REV	IGH	568065..568262	AB019439	168065..168262	105792324..105792521	NT_026437.11	87722324..87722521
IGHV3-23	14q32.33	REV	IGH	563887..564381	M99660	10..504	105796205..105796699	NT_026437.11	87726205..87726699
IGHV1-24	14q32.33	REV	IGH	555962..556438	M99642	68..544	105804148..105804624	NT_026437.11	87734148..87734624
IGHV3-25	14q32.33	REV	IGH	543468..543960	M99661	78..570	105816626..105817118	NT_026437.11	87746626..87747118
IGHV(III)-25-1	14q32.33	REV	IGH	539140..539381	AB019439	139140..139381	105821205..105821446	NT_026437.11	87751205..87751446
IGHV2-26	14q32.33	REV	IGH	531449..531931	M99648	21..503	105828655..105829137	NT_026437.11	87758655..87759137
IGHV(III)-26-1	14q32.33	REV	IGH	523568..523874	AB019439	123568..123874	105836712..105837018	NT_026437.11	87766712..87767018
IGHV(II)-26-2	14q32.33	REV	IGH	518318..518630	AB019439	118318..118630	105841956..105842268	NT_026437.11	87771956..87772268
IGHV7-27	14q32.33	REV	IGH	515042..515455	M99643	65..526	105845083..105845544	NT_026437.11	87775083..87775544
IGHV4-31	14q32.33	REV	IGH	483897..484373	X05713	153..626	105851517..105851990	NT_026437.11	87781517..87781990
IGHV(II)-28-1	14q32.33	REV	IGH	503603..503855	AB019439	103603..103855	105856731..105856983	NT_026437.11	87786731..87786983
IGHV3-29	14q32.33	REV	IGH	501726..502222	AB019439	101726..102222	105858364..105858860	NT_026437.11	87788364..87788860
IGHV3-30	14q32.33	REV	IGH	498085..498577	M83134	1782..2274	105862009..105862501	NT_026437.11	87792009..87792501
IGHV(II)-30-1	14q32.33	REV	IGH	490666..490939	AB019439	90666..90939	105869647..105869920	NT_026437.11	87799647..87799920
IGHV3-30-2	14q32.33	REV	IGH	488784..489232	AB019439	88784..89232	105871354..105871802	NT_026437.11	87801354..87801802

FIG. 29 (Continued)

IGHV4-28	14q32.33	REV	IGH	508596..509069	X05714	151..624	105549114..105549584	NT_026437.11	87479114..87479584
IGHV(II)-31-1	14q32.33	REV	IGH	478924..479232	AB019439	78924..79232	105881354..105881662	NT_026437.11	87811354..87811662
IGHV3-32	14q32.33	REV	IGH	477013..477509	AB019439	77013..77509	105883077..105883573	NT_026437.11	87813077..87813573
IGHV2-33	14q32.33	REV	IGH	473368..473860	AB019439	73368..73860	105886726..105887218	NT_026437.11	87816726..87817218
IGHV(II)-33-1	14q32.33	REV	IGH	465947..466220	AB019439	65947..66220	105894366..105894639	NT_026437.11	87824366..87824639
IGHV3-33-2	14q32.33	REV	IGH	464064..464512	AB019439	64064..64512	105896074..105896522	NT_026437.11	87826074..87826522
IGHV4-34	14q32.33	REV	IGH	459517..459988	AB019439	1..343	105900598..105900940	NT_026437.11	87830598..87830940
IGHV7-34-1	14q32.33	REV	IGH	455877..456349	AB019439	55877..56349	105904237..105904709	NT_026437.11	87834237..87834709
IGHV3-35	14q32.33	REV	IGH	443767..444259	M99666	140..632	105916327..105916819	NT_026437.11	87846327..87846819
IGHV3-36	14q32.33	REV	IGH	440399..440861	M99667	218..681	105919724..105920187	NT_026437.11	87849724..87850187
IGHV3-37	14q32.33	REV	IGH	436515..437003	M99668	10..489	105923586..105924071	NT_026437.11	87853586..87854071
IGHV3-38	14q32.33	REV	IGH	422685..423175	M99669	10..497	105937413..105937901	NT_026437.11	87867413..87867901
IGHV(III)-38-1	14q32.33	REV	IGH	415315..415645	AB019439	15315..15645	105944941..105945271	NT_026437.11	87874941..87875271
IGHV4-30-4	14q32.33	REV	IGH		Z14238	1..438	105876252..105876689	NT_026437.11	87806252..87806689
IGHV7-40	14q32.33	REV	IGH	408075..408316	M99644	164..497	105952270..105952592	NT_026437.11	87882270..87882592
IGHV(II)-40-1	14q32.33	REV	IGH	392620..392696	AB019438	192620..192696	105967890..105967966	NT_026437.11	87897890..87897966
IGHV3-30-3	14q32.33	REV	IGH		X92283	1..294	105970093..105970379	NT_026437.11	87900093..87900378
IGHV3-42	14q32.33	REV	IGH	369925..370362	M99671	1..461	105762906..105763129	NT_026437.11	87692869..87693089
IGHV3-43	14q32.33	REV	IGH	362897..363393	M99672	170..666	105997193..105997689	NT_026437.11	87927193..87927689
IGHV(II)-43-1	14q32.33	REV	IGH	360618..360826	AB019438	160618..160826	105999760..105999968	NT_026437.11	87929760..87929968
IGHV(III)-44	14q32.33	REV	IGH	355142..355318	M99673	62..511	106005056..106005444	NT_026437.11	87935056..87935444
IGHV(IV)-44-1	14q32.33	REV	IGH	343764..344231	AB019438	143764..144231	106016355..106016822	NT_026437.11	87946355..87946822
IGHV(II)-44-2	14q32.33	REV	IGH	339146..339382	AB019438	136146..139382	106021204..106024440	NT_026437.11	87951204..87954440
IGHV1-45	14q32.33	REV	IGH	326175..326651	X92209	2..478	106033935..106034411	NT_026437.11	87963935..87964411
IGHV1-46	14q32.33	REV	IGH	322057..322533	X92343	153..629	106038053..106038529	NT_026437.11	87968053..87968529
IGHV(II)-46-1	14q32.33	REV	IGH	317614..317771	AB019438	117614..117771	106042815..106042972	NT_026437.11	87972815..87972972
IGHV3-47	14q32.33	REV	IGH	314580..315074	Z18900	1..291	106045553..106045843	NT_026437.11	87975553..87975843
IGHV(III)-47-1	14q32.33	REV	IGH	302103..302406	AB019438	102103..102406	106058180..106058483	NT_026437.11	87988180..87988483
IGHV3-48	14q32.33	REV	IGH	295274..295768	M99675	174..668	106064818..106065312	NT_026437.11	87994818..87995312
IGHV3-49	14q32.33	REV	IGH	276144..276644	M99676	224..724	106083946..106084442	NT_026437.11	88013946..88014442
IGHV(II)-49-1	14q32.33	REV	IGH	268955..269225	AB019438	68955..69225	106091361..106091631	NT_026437.11	88021361..88021631
IGHV3-50	14q32.33	REV	IGH	266998..267453	M99677	10..283	106093316..106093588	NT_026437.11	88023316..88023585
IGHV5-51	14q32.33	REV	IGH	254379..254853	M99686	168..641	106105734..106106207	NT_026437.11	88035734..88036207
IGHV(III)-51-1	14q32.33	REV	IGH	249747..250053	AB019438	49747..50053	106110533..106110839	NT_026437.11	88040533..88040839
IGHV(II)-51-2	14q32.33	REV	IGH	248381..248843	AB019438	48581..48843	106111743..106112005	NT_026437.11	88041743..88042005
IGHV3-52	14q32.33	REV	IGH	246731..247220	M99678	212..701	106113366..106113855	NT_026437.11	88043366..88043855
IGHV3-53	14q32.33	REV	IGH	240421..240910	M99679	38..527	106282134..106282436	NT_026437.11	88212134..88212436
IGHV(II)-53-1	14q32.33	REV	IGH	233605..233874	AB019438	33605..33874	106126712..106126981	NT_026437.11	88056712..88056981
IGHV3-54	14q32.33	REV	IGH	231746..232232	M99680	147..631	106128356..106128840	NT_026437.11	88058356..88058840
IGHV4-55	14q32.33	REV	IGH	226980..227454	M99685	230..704	106133132..106133606	NT_026437.11	88063132..88063606
IGHV7-56	14q32.33	REV	IGH	223335..223808	M29810	136..429	106136958..106137251	NT_026437.11	88066958..88067251
IGHV3-57	14q32.33	REV	IGH	214314..214658	M29815	7..568	106146106..106146421	NT_026437.11	88076062..88076381
IGHV1-8	14q32.33	REV	IGH	750026..750503	M99637	58..535	105610083..105610560	NT_026437.11	87540083..87540560
IGHV4-39	14q32.33	REV	IGH	411486..411963	AB019439	1..328	105948662..105948989	NT_026437.11	87878662..87878989
IGHV3-60	14q32.33	REV	IGH	201883..202375	AB019438	1883..2375	106158211..106158703	NT_026437.11	88088211..88088703

FIG. 29 (Continued)

IGHV(II)-60-1	14q32.33	REV	IGH	195563..195869	AB019437	195563..195831	106164755..106165023	NT_026437.11	88094755..88095023
IGHV4-61	14q32.33	REV	IGH	193980..194456	M29811	151..628	106166130..106166606	NT_026437.11	88096130..88096606
IGHV3-62	14q32.33	REV	IGH	189953..190447	AB019437	189953..190447	106170139..106170633	NT_026437.11	88100139..88100633
IGHV(II)-62-1	14q32.33	REV	IGH	182749..183060	AB019437	182749..183060	106177526..106177837	NT_026437.11	88107526..88107837
IGHV3-63	14q32.33	REV	IGH	180646..181313	M99681	10..506	106179251..106179740	NT_026437.11	88109251..88109740
IGHV3-64	14q32.33	REV	IGH	175347..175841	M99682	78..575	106184745..106185242	NT_026437.11	88114745..88115242
IGHV3-65	14q32.33	REV	IGH	166993..167487	Z27503	1..301	106193140..106193440	NT_026437.11	88123140..88123440
IGHV(III)-2-1	14q32.33	REV	IGH	821563..821903	AB019441	21563..21903	105538683..105539023	NT_026437.11	87468683..87469023
IGHV(II)-65-1	14q32.33	REV	IGH	161481..161753	AB019437	161481..161753	106198833..106199105	NT_026437.11	88128833..88129105
IGHV3-66	14q32.33	REV	IGH	158060..158549	X92218	2..491	106202037..106202526	NT_026437.11	88132037..88132526
IGHV1-67	14q32.33	REV	IGH	152482..152960	AB019437	1..434	106207626..106208059	NT_026437.11	88137626..88138059
IGHV(II)-67-1	14q32.33	REV	IGH	146982..147131	AB019437	146982..147131	106213455..106213604	NT_026437.11	88143455..88143604
IGHV(III)-67-2	14q32.33	REV	IGH	146324..146422	AB019437	146324..146422	106214164..106214262	NT_026437.11	88144164..88144262
IGHV(III)-67-3	14q32.33	REV	IGH	140636..140910	AB019437	140636..140910	106219676..106219950	NT_026437.11	88149676..88149950
IGHV(III)-67-4	14q32.33	REV	IGH	138127..138432	AB019437	138127..138432	106222154..106222459	NT_026437.11	88152154..88152459
IGHV1-68	14q32.33	REV	IGH	129238..129717	AB019437	129238..129717	106230869..106231348	NT_026437.11	88160869..88161348
IGHV1-69	14q32.33	REV	IGH	119174..119651	L22582	233..710	106240935..106241412	NT_026437.11	88170935..88171412
IGHV2-70	14q32.33	REV	IGH	110279..110761	L21969	1..444	106249864..106250307	NT_026437.11	88179864..88180307
IGHV3-71	14q32.33	REV	IGH	105684..106184	AB019437	105684..106184	106254402..106254902	NT_026437.11	88184402..88184902
IGHV3-72	14q32.33	REV	IGH	90150..90650	X92206	87..587	106269936..106270436	NT_026437.11	88199936..88200436
IGHV3-73	14q32.33	REV	IGH	78150..78650	X70197	1..300	105990226..105990526	NT_026437.11	87920226..87920526
IGHV3-74	14q32.33	REV	IGH	70412..70906	L33851	23..476	106289721..106290174	NT_026437.11	88219721..88220174
IGHV(II)-74-1	14q32.33	REV	IGH	59900..60068	AB019437	59900..60068	106300518..106300686	NT_026437.11	88230518..88230686
IGHV3-75	14q32.33	REV	IGH	57164..57672	Z27510	4..308	106302955..106303259	NT_026437.11	88232955..88233259
IGHV3-76	14q32.33	REV	IGH	53003..53492	AB019437	1..292	106307134..106307425	NT_026437.11	88237134..88237425
IGHV(III)-76-1	14q32.33	REV	IGH	49245..49551	AB019437	49245..49551	106311035..106311341	NT_026437.11	88241035..88241341
IGHV5-78	14q32.33	REV	IGH	29780..30228	X92213	592..1027	106330373..106330806	NT_026437.11	88260373..88260806
IGHV(II)-78-1	14q32.33	REV	IGH	15431..15744	AB019437	15431..15744	106344842..106345155	NT_026437.11	88274842..88275155
IGHV3-79	14q32.33	REV	IGH	13237..13688	AB019437	13237..13688	106346898..106347349	NT_026437.11	88276898..88277349
IGHV4-80	14q32.33	REV	IGH	8147..8547	AB019437	8147..8547	106352039..106352439	NT_026437.11	88282039..88282439
IGHV7-81	14q32.33	REV	IGH	6315..6790	AB019437	6315..6790	106353796..106354271	NT_026437.11	88283796..88284271
IGHV(III)-82	14q32.33	REV	IGH	1490..1781	AB019437	1490..1781	106358805..106359096	NT_026437.11	88288805..88289096

FIG. 30

IMGT Gene	Chromosomal localization	Orientation	NCBI Gene ID	Reference GRCm38.p3 C57BL/6J	NCBI Gene positions in sequence	Orientation sequence	MGI ID	MGI symbol
IGHV1-86	12F2 (62.59 cM)	REV	668599	NC_000078.6	116009659..116009954	-1	MGI:3645729	Ighv1-86
IGHV1-85	12F2 (62.59 cM)	REV	668597	NC_000078.6	116000028..116000321	-1	MGI:3645723	Ighv1-85
IGHV1-84	12F2 (62.59 cM)	REV	434609	NC_000078.6	115980702..115981134	-1	MGI:3644235	Ighv1-84
IGHV1-83	12F2 (62.59 cM)	REV	620140	NC_000078.6	115963817..115964109	-1	MGI:3648939	Ighv1-83
IGHV1-82	12F2 (62.59 cM)	REV	100775175	NC_000078.6	115952538..115952831	-1	MGI:4439671	Ighv1-82
IGHV1-81	12F2 (62.59 cM)	REV	668591	NC_000078.6	115920279..115920572	-1	MGI:4439635	Ighv1-81
IGHV1-80	12F2 (62.59 cM)	REV	668589	NC_000078.6	115912344..115912637	-1	MGI:4439738	Ighv1-80
IGHV1-79	12F2 (62.59 cM)	REV	435331	NC_000078.6	115888096..115888389	-1	MGI:3648258	Ighv1-79
IGHV1-78	12F2 (62.59 cM)	REV	213570	NC_000078.6	115868773..115869066	-1	MGI:4439736	Ighv1-78
IGHV1-77	12F2 (62.59 cM)	REV	619994	NC_000078.6	115861867..115862160	-1	MGI:4439670	Ighv1-77
IGHV8-16	12F2 (62.59 cM)	REV	100775177	NC_000078.6	115858008..115858301	-1	MGI:5434409	Ighv8-16
IGHV1-76	12F2 (62.59 cM)	REV	100775174	NC_000078.6	115847881..115848174	-1	MGI:4439737	Ighv1-76
IGHV8-15	12F2 (62.59 cM)	REV	100775176	NC_000078.6	115844016..115844309	-1	MGI:5434408	Ighv8-15
IGHV1-75	12F2 (62.59 cM)	REV	622728	NC_000078.6	115833950..115834243	-1	MGI:4439735	Ighv1-75
IGHV8-14	12F2 (62.59 cM)	REV	668582	NC_000078.6	115808427..115808727	-1	MGI:3648977	Ighv8-14
IGHV1-74	12F2 (62.59 cM)	REV	100775173	NC_000078.6	115802690..115803080	-1	MGI:4439634	Ighv1-74
IGHV1-73	12F2 (62.59 cM)	REV	668577	NC_000078.6	115790593..115790887	-1	MGI:3645131	Ighv1-73
IGHV8-13	12F2 (62.59 cM)	REV	100775172	NC_000078.6	115765335..115765635	-1	MGI:4439734	Ighv8-13
IGHV1-72	12F2 (62.59 cM)	REV	619916	NC_000078.6	115757984..115758277	-1	MGI:4439633	Ighv1-72
IGHV1-71	12F2 (62.59 cM)	REV	619886	NC_000078.6	115742205..115742506	-1	MGI:3643033	Ighv1-71
IGHV1-70	12F2 (62.59 cM)	REV	638634	NC_000078.6	115691889..115692182	-1	MGI:3644620	Ighv1-70
IGHV8-12	12F2 (62.59 cM)	REV	780960	NC_000078.6	115647945..115648245	-1	MGI:3642873	Ighv8-12
IGHV1-69	12F2 (62.59 cM)	REV	619833	NC_000078.6	115623203..115623595	-1	MGI:4439632	Ighv1-69
IGHV1-68	12F2 (62.59 cM)	REV	780959	NC_000078.6	115611340..115611633	-1	MGI:5009923	Ighv1-68

FIG. 30 (Continued)

IGHV1-67	12F2 (62.59 cM)	REV	435328	NC_000078.6	115603940..115604233	-1	MGI:3645228	Ighv1-67
IGHV1-66	12F2 (62.59 cM)	REV	380824	NC_000078.6	115593110..115593403	-1	MGI:4439825	Ighv1-66
IGHV8-11	12F2 (62.59 cM)	REV	636462	NC_000078.6	115567149..115567449	-1	MGI:3644925	Ighv8-11
IGHV1-65	12F2 (62.58 cM)	REV	780958	NC_000078.6	115532055..115532347	-1	MGI:5009922	Ighv1-65
IGHV8-10	12F2 (62.58 cM)	REV	780957	NC_000078.6	115519340..115519640	-1	MGI:5009921	Ighv8-10
IGHV1-64	12F2 (62.58 cM)	REV	380823	NC_000078.6	115507587..115507977	-1	MGI:4439789	Ighv1-64
IGHV1-63	12F2 (62.58 cM)	REV	780956	NC_000078.6	115495625..115496058	-1	MGI:3642755	Ighv1-63
IGHV8-9	12F2 (62.57 cM)	REV	432709	NC_000078.6	115468331..115468632	-1	MGI:3644475	Ighv8-9
IGHV1-62-3	12F2 (62.57 cM)	REV	668549	NC_000078.6	115460999..115461431	-1	MGI:3648544	Ighv1-62-3
IGHV1-62-2	12F2 (62.57 cM)	REV	238448	NC_000078.6	115446417..115446710	-1	MGI:3644968	Ighv1-62-2
IGHV1-62-1	12F2 (62.55 cM)	REV	668544	NC_000078.6	115386707..115386988	-1	MGI:3704125	Ighv1-62-1
IGHV1-62	12F2 (62.55 cM)	REV	668542	NC_000078.6	115371986..115372278	-1	MGI:5009844	Ighv1-62
IGHV1-61	12F2 (62.55 cM)	REV	674018	NC_000078.6	115359140..115359433	-1	MGI:4439824	Ighv1-61
IGHV1-60	12F2 (62.55 cM)	REV	780940	NC_000078.6	115344537..115344830	-1	MGI:5009920	Ighv1-60
IGHV1-59	12F2 (62.54 cM)	REV	432708	NC_000078.6	115335057..115335513	-1	MGI:3644474	Ighv1-59
IGHV1-58	12F2 (62.54 cM)	REV	780939	NC_000078.6	115312166..115312597	-1	MGI:4439557	Ighv1-58
IGHV8-8	12F2 (62.54 cM)	REV	780938	NC_000078.6	115294066..115294362	-1	MGI:3815333	Ighv8-8
IGHV1-57	12F2 (62.53 cM)	REV	780937	NC_000078.6	115268424..115268715	-1	MGI:5009919	Ighv1-57
IGHV8-7	12F2 (62.53 cM)	REV	780936	NC_000078.6	115258139..115258432	-1	MGI:5009918	Ighv8-7
IGHV1-56	12F2 (62.53 cM)	REV	382695	NC_000078.6	115242802..115243095	-1	MGI:4439715	Ighv1-56
IGHV1-55	12F2 (62.52 cM)	REV	780932	NC_000078.6	115208137..115208527	-1	MGI:4439716	Ighv1-55
IGHV1-54	12F2 (62.52 cM)	REV	211331	NC_000078.6	115193675..115193968	-1	MGI:3647133	Ighv1-54
IGHV8-6	12F2 (62.51 cM)	REV	629930	NC_000078.6	115165777..115166077	-1	MGI:3645823	Ighv8-6
IGHV1-53	12F2 (62.51 cM)	REV	780931	NC_000078.6	115158445..115158835	-1	MGI:3576502	Ighv1-53
IGHV1-52	12F2 (62.51 cM)	REV	382696	NC_000078.6	115145487..115145780	-1	MGI:4439752	Ighv1-52
IGHV1-51	12F2 (62.51 cM)	REV	780930	NC_000078.6	115131375..115131668	-1	MGI:5009917	Ighv1-51
IGHV1-50	12F2 (62.5 cM)	REV	780929	NC_000078.6	115119748..115120041	-1	MGI:4439753	Ighv1-50
IGHV8-5	12F2 (62.49 cM)	REV	640506	NC_000078.6	115067560..115067860	-1	MGI:3645478	Ighv8-5
IGHV1-49	12F2 (62.49 cM)	REV	629925	NC_000078.6	115055223..115055516	-1	MGI:4439754	Ighv1-49
IGHV1-48	12F2 (62.49 cM)	REV	780928	NC_000078.6	115038227..115038520	-1	MGI:5009916	Ighv1-48

FIG. 30 (Continued)

IGHV8-4	12F2 (62.48 cM)	REV	629919	NC_000078.6	115024026..115024318	-1	MGI:3646040	Ighv8-4
IGHV8-3	12F2 (62.48 cM)	REV	780927	NC_000078.6	115003418..115003712	-1	MGI:5009915	Ighv8-3
IGHV1-47	12F2 (62.48 cM)	REV	629915	NC_000078.6	114991108..114991401	-1	MGI:4439890	Ighv1-47
IGHV1-46	12F2 (62.48 cM)	REV	668522	NC_000078.6	114977693..114977986	-1	MGI:5009843	Ighv1-46
IGHV1-45	12F2 (62.47 cM)	REV	780894	NC_000078.6	114970055..114970348	-1	MGI:5009913	Ighv1-45
IGHV1-44	12F2 (62.47 cM)	REV	668521	NC_000078.6	114963380..114963674	-1	MGI:5009842	Ighv1-44
IGHV1-43	12F2 (62.47 cM)	REV	629908	NC_000078.6	114945950..114946243	-1	MGI:3704124	Ighv1-43
IGHV1-42	12F2 (62.47 cM)	REV	629906	NC_000078.6	114937154..114937544	-1	MGI:3704123	Ighv1-42
IGHV1-41	12F2 (62.47 cM)	REV	629904	NC_000078.6	114932371..114932664	-1	MGI:3649084	Ighv1-41
IGHV1-40	12F2 (62.47 cM)	REV	780892	NC_000078.6	114924941..114925233	-1	MGI:5009912	Ighv1-40
IGHV1-39	12F2 (62.46 cM)	REV	780891	NC_000078.6	114914599..114914892	-1	MGI:4439888	Ighv1-39
IGHV1-38	12F2 (62.46 cM)	REV	780890	NC_000078.6	114910011..114910303	-1	MGI:5009911	Ighv1-38
IGHV1-37	12F2 (62.46 cM)	REV	668517	NC_000078.6	114896238..114896531	-1	MGI:4439640	Ighv1-37
IGHV1-36	12F2 (62.46 cM)	REV	629897	NC_000078.6	114879888..114880181	-1	MGI:4439639	Ighv1-36
IGHV1-35	12F2 (62.46 cM)	REV	668515	NC_000078.6	114875302..114875597	-1	MGI:3644935	Ighv1-35
IGHV1-34	12F2 (62.45 cM)	REV	628614	NC_000078.6	114851190..114851483	-1	MGI:4439659	Ighv1-34
IGHV1-33	12F2 (62.45 cM)	REV	641247	NC_000078.6	114843693..114844083	-1	MGI:4439617	Ighv1-33
IGHV1-32	12F2 (62.45 cM)	REV	780888	NC_000078.6	114835223..114835516	-1	MGI:5009910	Ighv1-32
IGHV1-31	12F2 (62.45 cM)	REV	629893	NC_000078.6	114829264..114829557	-1	MGI:4439889	Ighv1-31
IGHV1-30	12F2 (62.44 cM)	REV	668511	NC_000078.6	114817099..114817390	-1	MGI:4937173	Ighv1-30
IGHV1-29	12F2 (62.44 cM)	REV	780887	NC_000078.6	114812321..114812605	-1	MGI:5009909	Ighv1-29
IGHV1-28	12F2 (62.44 cM)	REV	668510	NC_000078.6	114809421..114809714	-1	MGI:3644938	Ighv1-28
IGHV1-27	12F2 (62.44 cM)	REV	780886	NC_000078.6	114804616..114804909	-1	MGI:5009908	Ighv1-27
IGHV1-26	12F2 (62.44 cM)	REV	629884	NC_000078.6	114788372..114788665	-1	MGI:4439641	Ighv1-26
IGHV1-25	12F2 (62.44 cM)	REV	640497	NC_000078.6	114781409..114781702	-1	MGI:3648295	Ighv1-25
IGHV1-24	12F2 (62.44 cM)	REV	780885	NC_000078.6	114772928..114773221	-1	MGI:3815052	Ighv1-24
IGHV1-23	12F2 (62.43 cM)	REV	780884	NC_000078.6	114764450..114764743	-1	MGI:3815050	Ighv1-23
IGHV1-22	12F2 (62.43 cM)	REV	780883	NC_000078.6	114746273..114746566	-1	MGI:4439784	Ighv1-22
IGHV1-21	12F2 (62.43 cM)	REV	668499	NC_000078.6	114740383..114740840	-1	MGI:5009841	Ighv1-21
IGHV1-21-1	12F2 (62.43 cM)	REV	631595	NC_000078.6	114736245..114736538	-1	MGI:3645737	Ighv1-21-1

FIG. 30 (Continued)

IGHV1-20	12F2 (62.43 cM)	REV	668497	NC_000078.6	114723772..114724065	-1	MGI:3644607	Ighv1-20
IGHV1-19	12F2 (62.42 cM)	REV	382692	NC_000078.6	114708648..114708941	-1	MGI:4439779	Ighv1-19
IGHV1-19-1	12F2 (62.42 cM)	REV	668493	NC_000078.6	114704147..114704440	-1	MGI:3644603	Ighv1-19-1
IGHV1-18	12F2 (62.42 cM)	REV	629871	NC_000078.6	114682632..114682925	-1	MGI:4439780	Ighv1-18
IGHV1-17	12F2 (62.42 cM)	REV	629869	NC_000078.6	114676749..114677042	-1	MGI:3647700	Ighv1-17
IGHV1-17-1	12F2 (62.42 cM)	REV	780926	NC_000078.6	114671968..114672262	-1	MGI:5009914	Ighv1-17-1
IGHV1-16	12F2 (62.42 cM)	REV	629866	NC_000078.6	114665872..114666165	-1	MGI:3647704	Ighv1-16
IGHV1-15	12F2 (62.41 cM)	REV	629865	NC_000078.6	114657353..114657646	-1	MGI:4439782	Ighv1-15
IGHV1-14	12F2 (62.41 cM)	REV	629863	NC_000078.6	114646572..114646865	-1	MGI:4439781	Ighv1-14
IGHV1-13	12F2 (62.41 cM)	REV	780865	NC_000078.6	114630680..114630973	-1	MGI:5009907	Ighv1-13
IGHV1-12	12F2 (62.41 cM)	REV	629860	NC_000078.6	114615850..114616143	-1	MGI:3646284	Ighv1-12
IGHV1-11	12F2 (62.41 cM)	REV	629859	NC_000078.6	114612244..114612536	-1	MGI:5009835	Ighv1-11
IGHV1-10	12F2 (62.40 cM)	REV	637050	NC_000078.6	114597878..114598175	-1	MGI:5009837	Ighv1-10
IGHV1-9	12F2 (62.40 cM)	REV	668478	NC_000078.6	114583569..114583862	-1	MGI:4439621	Ighv1-9
IGHV15-2	12F2 (62.40 cM)	REV	780864	NC_000078.6	114564578..114564874	-1	MGI:4947963	Ighv15-2
IGHV1-8	12F2 (62.39 cM)	REV	780863	NC_000078.6	114555610..114555902	-1	MGI:5009906	Ighv1-8
IGHV10-4	12F2 (62.39 cM)	REV	780862	NC_000078.6	114547258..114547556	-1	MGI:5009905	Ighv10-4
IGHV1-7	12F2 (62.39 cM)	REV	668474	NC_000078.6	114538495..114538788	-1	MGI:3704122	Ighv1-7
IGHV1-6	12F2 (62.39 cM)	REV	780861	NC_000078.6	114532972..114533264	-1	MGI:5009904	Ighv1-6
IGHV10-3	12F2 (62.39 cM)	REV	380809	NC_000078.6	114523441..114523887	-1	MGI:3648785	Ighv10-3
IGHV1-5	12F2 (62.39 cM)	REV	668469	NC_000078.6	114513330..114513623	-1	MGI:3704121	Ighv1-5
IGHV10-2	12F2 (62.38 cM)	REV	780860	NC_000078.6	114496003..114496300	-1	MGI:5009903	Ighv10-2
IGHV1-4	12F2 (62.38 cM)	REV	432702	NC_000078.6	114487136..114487429	-1	MGI:4439618	Ighv1-4
IGHV1-3	12F2 (62.38 cM)	REV	780859	NC_000078.6	114481615..114481916	-1	MGI:5009902	Ighv1-3
IGHV10-1	12F2 (62.38 cM)	REV	380808	NC_000078.6	114479007..114479451	-1	MGI:4439620	Ighv10-1
IGHV1-2	12F2 (62.38 cM)	REV	238428	NC_000078.6	114469075..114469371	-1	MGI:4439924	Ighv1-2
IGHV8-2	12F2 (62.38 cM)	REV	629849	NC_000078.6	114462290..114462589	-1	MGI:3647474	Ighv8-2
IGHV6-7	12F2 (62.38 cM)	REV	432701	NC_000078.6	114455626..114455925	-1	MGI:3644477	Ighv6-7
IGHV6-6	12F2 (62.37 cM)	REV	238427	NC_000078.6	114434788..114435087	-1	MGI:4439619	Ighv6-6
IGHV6-5	12F2 (62.37 cM)	REV	639510	NC_000078.6	114416596..114416895	-1	MGI:3647134	Ighv6-5

FIG. 30 (Continued)

IGHV6-4	12F2 (62.37 cM)	REV	628398	NC_000078.6	114406474..114406773	-1	MGI:3704120	Ighv6-4
IGHV6-3	12F2 (62.36 cM)	REV	629845	NC_000078.6	114391712..114392010	-1	MGI:4439854	Ighv6-3
IGHV12-3	12F2 (62.36 cM)	REV	435321	NC_000078.6	114366526..114366819	-1	MGI:3646760	Ighv12-3
IGHV13-2	12F2 (62.36 cM)	REV	629842	NC_000078.6	114357761..114358060	-1	MGI:3705861	Ighv13-2
IGHV1-1	12F2 (62.36 cM)	REV	780834	NC_000078.6	114351705..114351995	-1	MGI:5009901	Ighv1-1
IGHV8-1	12F2 (62.35 cM)	REV	780833	NC_000078.6	114332796..114333086	-1	MGI:5009900	Ighv8-1
IGHV3-8	12F2 (62.35 cM)	REV	780831	NC_000078.6	114322376..114322666	-1	MGI:3645298	Ighv3-8
IGHV5-21	12F2 (62.35 cM)	REV	780832	NC_000078.6	114320030..114320324	-1	MGI:5009899	Ighv5-21
IGHV3-7	12F2 (62.35 cM)	REV	780830	NC_000078.6	114314186..114314479	-1	MGI:5009898	Ighv3-7
IGHV9-4	12F2 (62.34 cM)	REV	636260	NC_000078.6	114299961..114300254	-1	MGI:3646379	Ighv9-4
IGHV3-6	12F2 (62.34 cM)	REV	780829	NC_000078.6	114288152..114288447	-1	MGI:4439856	Ighv3-6
IGHV13-1	12F2 (62.33 cM)	REV	780828	NC_000078.6	114267607..114267906	-1	MGI:4439855	Ighv13-1
IGHV3-5	12F2 (62.33 cM)	REV	633457	NC_000078.6	114262652..114262950	-1	MGI:3648045	Ighv3-5
IGHV3-4	12F2 (62.33 cM)	REV	629833	NC_000078.6	114253617..114253915	-1	MGI:3644034	Ighv3-4
IGHV7-4	12F2 (62.32 cM)	REV	632654	NC_000078.6	114222788..114223254	-1	MGI:4439763	Ighv7-4
IGHV3-3	12F2 (62.31 cM)	REV	668438	NC_000078.6	114196439..114196875	-1	MGI:3643949	Ighv3-3
IGHV14-4	12F2 (62.30 cM)	REV	629826	NC_000078.6	114176438..114176867	-1	MGI:4439765	Ighv14-4
IGHV15-1	12F2 (62.30 cM)	REV	780827	NC_000078.6	114158025..114158315	-1	MGI:5009897	Ighv15-1
IGHV7-3	12F2 (62.30 cM)	REV	629822	NC_000078.6	114153180..114153644	-1	MGI:4439766	Ighv7-3
IGHV9-3	12F2 (62.29 cM)	REV	780825	NC_000078.6	114140692..114141122	-1	MGI:3642720	Ighv9-3
IGHV12-2	12F2 (62.29 cM)	REV	780824	NC_000078.6	114127533..114127984	-1	MGI:5009896	Ighv12-2
IGHV9-2	12F2 (62.28 cM)	REV	629818	NC_000078.6	114109001..114109430	-1	MGI:3643816	Ighv9-2
IGHV12-1	12F2 (62.28 cM)	REV	780823	NC_000078.6	114107259..114107563	-1	MGI:5295669	Ighv12-1
IGHV9-1	12F2 (62.28 cM)	REV	195180	NC_000078.6	114093928..114094221	-1	MGI:4439911	Ighv9-1
IGHV6-2	12F2 (62.28 cM)	REV	780822	NC_000078.6	114089387..114089681	-1	MGI:5009895	Ighv6-2
IGHV16-1	12F2 (62.27 cM)	REV	629812	NC_000078.6	114068828..114069126	-1	MGI:3643819	Ighv16-1
IGHV14-3	12F2 (62.27 cM)	REV	238418	NC_000078.6	114059845..114060273	-1	MGI:4439764	Ighv14-3
IGHV11-2	12F2 (62.27 cM)	REV	780818	NC_000078.6	114048241..114048704	-1	MGI:4947968	Ighv11-2
IGHV3-2	12F2 (62.26 cM)	REV	780817	NC_000078.6	114033803..114034097	-1	MGI:4439923	Ighv3-2
IGHV4-2	12F2 (62.26 cM)	REV	668423	NC_000078.6	114013144..114013439	-1	MGI:5009840	Ighv4-2

FIG. 30 (Continued)

IGHV14-2	12F2 (62.25 cM)	REV	668421	NC_000078.6	113994469..113994898	-1	MGI:4439607	Ighv14-2
IGHV11-1	12F2 (62.25 cM)	REV	780804	NC_000078.6	113981879..113982174	-1	MGI:4439535	Ighv11-1
IGHV3-1	12F2 (62.24 cM)	REV	780803	NC_000078.6	113964390..113964683	-1	MGI:4439534	Ighv3-1
IGHV4-1	12F2 (62.24 cM)	REV	780802	NC_000078.6	113948284..113948577	-1	MGI:4439536	Ighv4-1
IGHV14-1	12F2 (62.23 cM)	REV	780801	NC_000078.6	113931953..113932246	-1	MGI:4439920	Ighv14-1
IGHV7-2	12F2 (62.23 cM)	REV	780800	NC_000078.6	113912025..113912483	-1	MGI:4439623	Ighv7-2
IGHV7-1	12F2 (62.22 cM)	REV	780799	NC_000078.6	113896408..113896946	-1	MGI:4439622	Ighv7-1
IGHV5-19	12F2 (62.22 cM)	REV	780798	NC_000078.6	113884761..113885342	-1	MGI:5009894	Ighv5-19
IGHV2-9	12F2 (62.22 cM)	REV	640046	NC_000078.6	113879097..113879388	-1	MGI:4439624	Ighv2-9
IGHV2-8	12F2 (62.22 cM)	REV	780797	NC_000078.6	113877361..113877659	-1	MGI:5009893	Ighv2-8
IGHV5-18	12F2 (62.22 cM)	REV	780796	NC_000078.6	113875772..113876239	-1	MGI:5009892	Ighv5-18
IGHV5-17	12F2 (62.21 cM)	REV	780794	NC_000078.6	113859149..113859442	-1	MGI:4439533	Ighv5-17
IGHV5-16	12F2 (62.21 cM)	REV	633568	NC_000078.6	113838528..113838821	-1	MGI:4439556	Ighv5-16
IGHV5-15	12F2 (62.20 cM)	REV	780792	NC_000078.6	113826648..113826941	-1	MGI:4439812	Ighv5-15
IGHV2-7	12F2 (62.20 cM)	REV	780793	NC_000078.6	113807314..113807747	-1	MGI:4439521	Ighv2-7
IGHV2-6-8	12F2 (62.19 cM)	REV	791090	NC_000078.6	113796138..113796430	-1	MGI:4439811	Ighv2-6-8
IGHV2-9-1	12F2 (62.18 cM)	REV	791089	NC_000078.6	113769857..113770142	-1	MGI:4439519	Ighv2-9-1
IGHV5-12-4	12F2 (62.18 cM)	REV	628098	NC_000078.6	113762251..113762544	-1	MGI:3645977	Ighv5-12-4
IGHV5-9-1	12F2 (62.17 cM)	REV	641178	NC_000078.6	113736113..113736406	-1	MGI:4439810	Ighv5-9-1
IGHV2-6	12F2 (62.17 cM)	REV	630486	NC_000078.6	113716672..113716962	-1	MGI:4439518	Ighv2-6
IGHV5-12	12F2 (62.16 cM)	REV	668395	NC_000078.6	113702126..113702419	-1	MGI:4439516	Ighv5-12
IGHV5-11	12F2 (62.16 cM)	REV	777784	NC_000078.6	113687229..113687505	-1	MGI:5009891	Ighv5-11
IGHV2-5	12F2 (62.16 cM)	REV	638605	NC_000078.6	113685482..113685774	-1	MGI:4439517	Ighv2-5
IGHV5-10	12F2 (62.16 cM)	REV	777783	NC_000078.6	113669673..113669968	-1	MGI:3642589	Ighv5-10
IGHV5-9	12F2 (62.15 cM)	REV	544896	NC_000078.6	113661771..113662228	-1	MGI:4439873	Ighv5-9
IGHV5-8	12F2 (62.15 cM)	REV	777782	NC_000078.6	113654967..113655245	-1	MGI:5009890	Ighv5-8
IGHV2-4	12F2 (62.15 cM)	REV	668389	NC_000078.6	113653291..113653583	-1	MGI:3643263	Ighv2-4
IGHV5-7	12F2 (62.14 cM)	REV	777781	NC_000078.6	113634686..113634981	-1	MGI:3642590	Ighv5-7
IGHV5-6	12F2 (62.14 cM)	REV	777780	NC_000078.6	113625508..113625801	-1	MGI:4439815	Ighv5-6
IGHV5-5	12F2 (62.14 cM)	REV	777779	NC_000078.6	113612965..113613442	-1	MGI:5009889	Ighv5-5

FIG. 30 (Continued)

IGHV2-3	12F2 (62.14 cM)	REV	238412	NC_000078.6	113611184..113611476	-1	MGI:4439872	Ighv2-3
IGHV6-1	12F2 (62.14 cM)	REV	777689	NC_000078.6	113603460..113603928	-1	MGI:5009888	Ighv6-1
IGHV5-4	12F2 (62.13 cM)	REV	777688	NC_000078.6	113597448..113597741	-1	MGI:4439895	Ighv5-4
IGHV5-3	12F2 (62.13 cM)	REV	777687	NC_000078.6	113589967..113590243	-1	MGI:5009887	Ighv5-3
IGHV2-2	12F2 (62.13 cM)	REV	777686	NC_000078.6	113588267..113588700	-1	MGI:4439894	Ighv2-2
IGHV5-2	12F2 (62.13 cM)	REV	777685	NC_000078.6	113578504..113578990	-1	MGI:4439893	Ighv5-2
IGHV2-1	12F2 (62.13 cM)	REV	777684	NC_000078.6	113574247..113574540	-1	MGI:4936981	Ighv2-1
IGHV5-1	12F2 (62.13 cM)	REV	777683	NC_000078.6	113572929..113573222	-1	MGI:5009886	Ighv5-1
IGHD5-1	12F2 (62.11 cM)	REV	777682	NC_000078.6	113526800..113526809	-1	MGI:4937117	Ighd5-1
IGHD3-1	12F2 (62.11 cM)	REV	777681	NC_000078.6	113525313..113525329	-1	MGI:4439891	Ighd3-1
IGHD1-1	12F2 (62.10 cM)	REV	777657	NC_000078.6	113482170..113482192	-1	MGI:4439871	Ighd1-1
IGHD6-1	12F2 (62.10 cM)	REV	777680	NC_000078.6	113480143..113480171	-1	MGI:4937254	Ighd6-1
IGHD2-3	12F2 (62.10 cM)	REV	777679	NC_000078.6	113475400..113475416	-1	MGI:4439708	Ighd2-3
IGHD6-2	12F2 (62.10 cM)	REV	777678	NC_000078.6	113474875..113474903	-1	MGI:4937120	Ighd6-2
IGHD2-4	12F2 (62.10 cM)	REV	777677	NC_000078.6	113470694..113470710	-1	MGI:4439709	Ighd2-4
IGHD5-2	12F2 (62.10 cM)	REV	777676	NC_000078.6	113469426..113469435	-1	MGI:4936898	Ighd5-2
IGHD2-5	12F2 (62.10 cM)	REV	777675	NC_000078.6	113466027..113466043	-1	MGI:4439705	Ighd2-5
IGHD5-3	12F2 (62.10 cM)	REV	777674	NC_000078.6	113464761..113464770	-1	MGI:4937297	Ighd5-3
IGHD5-7	12F2 (62.10 cM)	REV	777673	NC_000078.6	113464524..113464552	-1	MGI:4936973	Ighd5-7
IGHD2-6	12F2 (62.10 cM)	REV	777672	NC_000078.6	113461369..113461385	-1	MGI:4439865	Ighd2-6
IGHD5-4	12F2 (62.10 cM)	REV	777670	NC_000078.6	113460101..113460110	-1	MGI:4937058	Ighd5-4
IGHD5-8	12F2 (62.10 cM)	REV	777669	NC_000078.6	113459864..113459892	-1	MGI:4937171	Ighd5-8
IGHD2-7	12F2 (62.10 cM)	REV	777667	NC_000078.6	113456720..113456736	-1	MGI:4439866	Ighd2-7
IGHD5-5	12F2 (62.10 cM)	REV	777665	NC_000078.6	113454942..113454951	-1	MGI:4937334	Ighd5-5
IGHD2-8	12F2 (62.10 cM)	REV	777664	NC_000078.6	113450851..113450867	-1	MGI:4439706	Ighd2-8
IGHD5-6	12F2 (62.10 cM)	REV	777663	NC_000078.6	113449588..113449597	-1	MGI:4937234	Ighd5-6
IGHD3-2	12F2 (62.10 cM)	REV	777662	NC_000078.6	113448214..113448229	-1	MGI:4439707	Ighd3-2
IGHD4-1	12F2 (62.10 cM)	REV	777661	NC_000078.6	113430528..113430538	-1	MGI:4439801	Ighd4-1
IGHJ1	12F2 (62.10 cM)	REV	777648	NC_000078.6	113429781..113429833	-1	MGI:4439802	Ighj1
IGHJ2	12F2 (62.10 cM)	REV	777654	NC_000078.6	113429468..113429515	-1	MGI:4439803	Ighj2

FIG. 30 (Continued)

IGHJ3	12F2 (62.10 cM)	REV	777655	NC_000078.6	113429085..113429132	-1	MGI:4439767	Ighj3
IGHJ4	12F2 (62.10 cM)	REV	777656	NC_000078.6	113428514..113428567	-1	MGI:4439658	Ighj4
IGHM	12F2 (62.10 cM)	REV	16019	NC_000078.6	113418826..113422730	-1	MGI:96448	Ighm
IGHD	12F2 (62.10 cM)	REV	380797	NC_000078.6	113407535..113416324	-1	MGI:96447	Ighd
IGHG3	12F2 (62.10 cM)	REV	380795	NC_000078.6	113357442..113361232	-1	MGI:2144790	Ighg3
IGHG1	12F2 (62.10 cM)	REV	16017	NC_000078.6	113326544..113330523	-1	MGI:96446	Ighg1
IGHG2B	12F2 (62.10 cM)	REV	16016	NC_000078.6	113304314..113307933	-1	MGI:96445	Ighg2b
IGHG2C	12F2 (62.10 cM)	REV	404711	NC_000078.6	113287389..113288932	-1	MGI:2686979	Ighg2c
IGHE	12F2 (62.10 cM)	REV	380792	NC_000078.6	113269263..113273248	-1	MGI:2685746	Ighe
IGHA	12F2 (62.09 cM)	REV	238447	NC_000078.6	113256204..113260236	-1	MGI:96444	Igha

FIG. 31

SEQ ID NO.1 Cy1ΔCH1 knock-in sequence 5193bp (artificial sequence)

gtgagaggacgtatagggaggaggaggttcactagaggtgaggctcaaggcattagcctgcctaaaaccaaccaggctggacagccatcaccaggaaaatggatctcagc
ccagaaagatcgaaagttgttcttctccctctggagattctatgtccttacactcattggttggattcccacacatcttgacaaacagagac
aattyaytatcaccagccaaaagtcatacccaaaaacagcctggcatgacctcacaccaqactcaaaccctaccctacctcttctatcctggtgctcttcctcatctccagac
cccagtaacacatagctttctctccacagtgcccagggatctgtggtctgcatatgtacaggtcaagtcagtagaggccttcacctgacccagatgcaac
aagtggccatgttagaggtggtcccaggtattgacctatttcctctttctcctctaygtcccagaagtatcatctgtcttcatcttcccccaaagcccaa
ggatgtgctcaccattactctgactgtgttgtggtagacatcagcaaggatgatcccgagttccagttcaggtttgtagatgatgtggagg
tgcacacagctcagacgaaaccccgggagcagatcaacagtttccgttcagtcagtgaacttcccatcatgcacaaggactggctccaatggcaggagttc
aaagtgcagggtcaacagagtgcagctttcccctgcccccatcgagaaaaccatctccaaaaaccaaagggtgagagctgcagtgatgaaggctgcaataggtcagtc
catagacagagcttggcataacagaycccctgccttgtcccatgtgctaaccaactctctttaccccacccagcgaccgaaggcctccacaggtgtacacc
attccacctcccaaggagcagcagaatggccaaggataaagtcagtctgcattgtacacaaacttcttcccctgaagacattactgtgygytgcagtggaatgggca
gccagcygagaactacaagaacacctcagcccatcatggacacagatggctcttactttcgtctacagcaagctcaatgtgcaggaaggcagtgaaata
ctttcacctgctctgtgttacatgaygcctgcacaaccacctactgagaagagcctctcccactctcctggtaaatgatcccagtgtcctgaattgcctggtc
ctacaggactctgacacctcctccaccccctcctgtgtaaataaagcacccagcactgcctcggaccctgcaataatgtcctggtgatttctgagatgagtc
tagctaggtcatgggaatgagggggtctccatgggtttgaggcctgagttgtyactaagggaaaaaccctaggcctacactgccaccaccctgaattttgcctg
acatgaaaagaatttacctctccttgtaaaygtgagcctiatccctaggcagttccctactaccagacctcctcttagcttgcacttgtttcttggcacagaatgtgtc
taaccccccaaagcaaggaagacacaccctcctacctcctcactctgcctttccctgtgctaaagcatctcactgagtgcgctgaatagatgcatgtggcc
acagtcttgcagacagacccttgccatctctccactcagctttccagaggctaagtctagtccctatgctatgctcctatgctcatc
agactcccaagtggaggatgaacatggaccctttaaaaccaacctgcgcagcaacaccctgccaataaggcccgtatgtgaaaatgtgccacacatctacacatgcac
aggcacacacacacacacatgcatgggcacacacatacagagagagaggaatcacagaaactcccatgagcatcctatacagtactcaaagataaaaaggtaccag
gtctaccacatgatcatccggcaattacaagtgggccaactgatacaggtgtataggcttccttactttcctgtgcatcact
tctgtccctgaacctcagactgtgccgctgactcatacagctgtataggcttcctttactttcctgtgtgtgaagtacact
catcacacagaagaggccctgtgattactctggccctctgtcitiggtcatcagagaatagacagaagaagatcaggcaaactacacagacatcccacaatcatcaca
ggccctgactctgctcccagtctcaaaactgaaggctggagcacacagaataagctcctgcacaggccagctatgtgtctgactgagcccag
ggacaaaatggcagcactttgggaactgagctgtccaaygaaagaggaatgggcctcaggtcctgcaggtgtcctgcagctgcagct
ttctcctgggcctccatacagcctccctgccacacagggaatgggcctaccgccctctctgtccaggggctgcaactggacg
agacctgtgctgagggcccaggacgggggctctggacgacatcaccatcttcatcagcctcttcctgctcagacgcgtctgctcaca
ctcttcaagqtcagccatactgtcccacagtgtctacaaatgtcctcatactcttcccatactgtccccatgtgctcccatgctcaataga
ccacagtcttacatgctatgtaatgctgtctaccttctgggaactgaggttctgggtccaaygaaygaggaatgggcctcagtcaggtcccattccaacccat
gcagccctttgttccccgctatgctgtccccatgctattgtcigtaittttcatgctctttccctagttgtccacattgttgtccaccacattg
tcccccactctgtacacagcctccacactgtactacccgataatgttcccctgttgtcccaacatcctcctcctgcaccatttgtcaactgtcccctgaattccca
tgttgttcccacactgttagtgttaatgtgctctgtcccaggtgtactcttgttccgtgctgtcttgtactaacccactctat
caccacactgtgcctatagcactgtcccacattgtcctacatcagtgtcccaatgatctacccacactatttccact
tcatgcccctacaatttcccctatttcctctggtcaccatgccatccttcccactcctgcacagctggagggactcccgggatgagtccttgcccagatg
agctacctatctagaggagtcttcaggtgggaaggggaatgcagtctgattcaccctgtctcacaggtaaagttgatcttcctcctcggtggtggaag
ctgaagcagacactggttcctgatacaaggaacatgattgggcaagcacctagggccacctcctgtaatggcattttcccaggcccgaaggaccctgtccaataatgc
caagcagcacaactgagatcacactgtctgctcatctcgcctttcctccccgagacccgagactcagctactctcaaatttttcctctctgaaaggaccatgtggacattac
attgctccaggccacagccaccaggacctaaaaacaccatcacagcagcaccaaagacactggatagacccacaaaggcaatagtttcctcaacagtatatcccaaaact
gttgggacaaacgagcaattcactgaagaagtgacaagttcccacaatgtcagtgtccaggctgagaaaggggacaaaaagtggtaccagccctgtccacaccacctcta

FIG. 31 (Continued)

attcacaggaatacgtgatagaagaggcaggttgtagatccgaaagatgagacagattttatcaactccagaaagagctgggtccaactgaattattctagcgacct
tggcattgtcatgacctgccatgaccttcctcctttaacacttcgataaaccctgggatatggaaaatgcctgtgtttctcagggtttgggaaagaaccatccatgtt
gggattcttgtgtagatcctccttctggtcacagatgcaatacactggattttcaggcaaaggagcaaattcacagacaactctggccctacagccctaagacctag
acaccaccatctccttggaattatcaaatttaacacccggcacacaacaaagaaggactgggactttgaggcctttgtgtagccctagaggggcagaggccactga
gcagggattgggtgatcacaaggacctcctggagagggacctgaggagcaggttccaattgggccaaagaaagaagaacaacaatagagatgaaggatgctggaaag
agccatggtacagcagtcttgtccttcagacatgactcttacagcccaggactcttacagtagctagctggagcagaagtccaagggattaccatgccctagggcca
caggctactggagggtggagtgagtctactacacaggtccaatgcctgtttctccattgcttctcagccaatgagaaatcagagtctccacctccaagaaaaaggaa
ggtggaaatgaaaggtgagcacctgccttcccgtgactggcagaaagatctccacggactcaaggctttgacttcaaacttcccgtagatccctgatgtctttaagg
actctgtctgctttgttggttttgtttgtttgttttgtactttgtcaataaaacatttttcaaatattctacaattgctgtgctctttcctatgcaaactggccctgg
cacttcaaaacatggcacagttaagttgaccagtgggccatgcagagcatactacccctcctggtcctgtgtcctaccagatatcctctaagtgtcccttcactgct
gggctccagtttctgctgccctgaaccacacaggtcctcagtctgtcctccctatggg

SEQ ID NO: 8: Linkage sequence between Sμ and Sγ1
ttcctgtggctagaaggcagctccatagccctgctgcagtttcgtcctgtataccaggttcacctactaccatatctagccctgcctgcctaaactgagatggg

SEQ ID NO: 9: Linkage sequence between Sμ and Cγ1ΔCH1'
aggcaagaagacagattcttacccctccatttctcttttatcccctctctggtcctgaggacgtatagggaggaggggttcactagaggtgaggctcaagccat

SEQ ID NO: 10: Linkage sequence
ctgacaggaggcaagaagacagattcttaccccctccatttctcttttatcccctctctggtcctaaccaaaaccctcccagcaggggtgccccaggcccaggcatggccc
agagggagcagc

SEQ ID NO: 13: CμΔCH1 knock-in sequence 4959 bp
aaccaaaccctcccagcaggggtgcccaggcccaggcatggcccagagggagcagcggggtggggcttaggccaagctgagctcacaccttgacctttcattccagc
tgtcgcagagatgaaccccaatgtaaatgtgttcgtcccaccacgggatggcttctctggccctgcaccacgcaagtctaaactcatctgcgaggccacgaacttca
ctccaaaaccgatcacagtatcctggctaaaggatgggaagctcgtggaatctggcttcaccacagatccggtgaccatcgagaacaaaggatccacaccccaaaacc
tacaaggtcataagcacacttaccatctctgaaatcgactggctgaacctgaatgtgtacacctgccgtgtggatcacaggggtctcaccttcttgaagaacgtgtc
ctccacatgtgctgccagtgagtggcctgggctaagcccaatgcctagccctcccagattagggaagtcctcctacaattatggccaatgccacccagacatggtca
tttgctccttgaactttggctccccagagtggccaaggacaagaatgagcaataggcagtagaggggtgagaatcagctggaaggaccagcatcttcccttaagtag
gtttgggggatggagactaagctttttccaacttcacaactagatatgtcataacctgacacagtgttctcttgactgcaggtcctccacagacatcctaaccttc
caccatccccccctcctttgccgacatcttcctcagcaagtccgctaacctgacctgtctggtctcaaacctggcaacctatgaaaccctgaatatctcctgggctt
ctcaaagtggtgaaccactggaaaccaaaattaaaatcatggaaagccatcccaatggcaccttcagtgctaagggtgtggctagtgtttgtgtggaagactggaat
aacaggaaggaatttgtgtgtactgtgactcacagggatctgccttcaccacagaagaaattcatctcaaaacccaatggtaggtatccccccttcccttcccctcc
aattgcaggacccttcctgtacctcatagggagggcaggtcctcttccaccctatcctcactactgtcttcatttacagaggtgcacaaacatccacctgctgtgta
cctgctgccaccagctcgtgagcaactgaacctgagggagtcagccacagtcacctgcctggtgaagggcttctctcctgcagacatcagtgtgcagtggcttcaga
gagggcaactcttgccccaagagaagtatgtgaccagtgccccgatgccagagcctgggggccccaggcttctactttacccacagcatcctgactgtgacagaggag
gaatggaactccggagagacctataccctgtgttgtaggccacgaggccctgccacacctggtgaccgagaggaccgtggacaagtccactggtaaacccacactgta
caatgtctccctgatcatgtctgacacaggcggcaccctgctattgaccatgctagcgctcaaccaggcaggccctgggtgtccagttgctctgtgtatgcaaactaa
ccatgtcagagtgagatgttgcatttttataaaaattagaaataaaaaaaatccattcaaacgtcactggttttgattatacaatgctcatgcctgctgagacagttg
tgttttgcttgctctgcacacaccctgcatacttgcctccaccctggcccttcctctaccttgccagtttcctccttgtgtgtgaactcagtcaggcttacaacaga
cagagtatgaacatgcgattcctccagctacttctagatatatggctgaaagcttgcctaacctggtgcaggcagcattcaggcacatatatagacacacatgcatt

FIG. 31 (Continued)

tatacatagatatataggtacacatgtgtagacacatacatgaatgtgtattcatggacacagacaaaggtacacatatacacatgaytcatgcgcacacac
atgcatggacacttacaaacgccttcagagacacaaataggcatacagacacacaaccactcacagaaacagataccaatatgcatggcctgtgtacacagaaacagact
atagggcaaatatacacaaaataaactatatagatacatacacacatgtacagaaaacatcttcacatgtgtacacttaacacatgtgaacaggtatagca
cacagatacacctggacctcgaccaggcctgtaatctccaaggctcagagagcctcacactaggctgggtcactgtatactcctcaggagcccacactcatga
ttgggagagataacccaggtacaaagtatgcctatctgtcttcaacacctatgggcagaagatacticccactaaccacccatgacagaaagttagccttggctgtc
tccattaatagaacactcagaagaccaatgtgaaattgcctaacccactcacaccccacttgatcttcccagttcaaaatgcagaaaacataatgcagttgtccaaaa
gatgcccaccacacacacacacacacacacacacacacacacacacacacacatcaagggagcctcgtgtaaggagtcaccaccaccaat
aacactgccctctttgggcctatatcctggacattcttcatatccattttgggccttaggcttttagatatcccaaggggctcatctttacagggatcagagat
cccaataaatgccctgtgtcccacagcctcccctcaggtatctgtctgtttatctcttggtaccaagaccaacgcattgctggcagggtagaggcaaggcaacggga
actctgatcaaagaaagtcatgagatgccttcaggaaggaggacaacctctggtatcccctgttcttattgctaaagcccaagagacagggagacc
tgctctaaaattctcagtcctaaacagcaccggatgggcaccacctgctcagggaaagtcccagaqcacaccaaatatcatttttgccacagttcctgagtcctgcca
ggtccatacattgcatctgtcttgctctgctgccccagggctcctggaacaaaggctccaaattagtgtgtcctacagcttggcctgttctgtgctcctcgtct
agcttgagctattagggaccagtcaatactcgctaaagaatcccagggccacccccaaccctttatgcaaatgctcagtcacacccaagaactttggcttgac
cctccctctgtgtcccctcatagagggggaagtgaatgctgaggaggcaaagcctttgagaaccctgtgcaccactgcatcctcttcctcctgag
cctcttctacagcaccaccgtcacccttgtccaaggtagtgtggggctgagggacacagggcctgggacaggggagtcaccagtcctcactgcctcctacctcctac
tccctacaagtgggacagcaattcacactgtcctcgtgcaccctgtcaggtgaaatgactctcagcatggaaggcaccaagaqaccaagaatcctcccacagggacacta
cctctgggcctgggatacctgactgtatgactagtaaacctcatcctcacgctctttcctgcgctgcttcctgagctgtcttttctctgagactga
ccaaagactttttgtcaacctgtaccaaatccigaagcaatgtctgggcccacagagacagtcgcaccctcaaacagcaaagtacacaacacagccaatggatagcaci
actttatgtgaagggaattttgtttttgtctttcaaacctttcctgcggtgttgacagcccaagtatactctgaatagagcttaggaatggaacagtgcagt
ctgattggtacttaaggggagaaaaaaaggagcaagaaaaaagcatgtccttgcctatcccaaagctttgtggtatgtggtgccttgggagttggctg
gctctgtcccagcccaccatcttgggctggtgttgcttcaaggctagtaatttaggtcttatcccaaagctttgtggtatgtggtgccttgggagttggctg
agatttgaagatgtttcgtacctctcccacaaacatgacaagccctagggtttagtcaataactcaaattctctgtctatgacaactgcgtgtatgacatatgaaga
atgggataaagatgctatagtcacticgaggacagataaaggcgctgctcacagagcaagaatggaaaggagcaggtcatgggcttcagttcagagaaggaa
gcacagcaaatacaaggtgaggtcatgttactcccctgttcatctgtggacaggctagagaagcccactgacatgggaagtctgggctaactgactacattg
agaggaggaccaggaagcagggcatgagagagcatgaccatgggaggtgatctgtgtagatgccaggtgcctaaagagagtctttaaataaaatgtgtgggtag
cacctggtgatggtcccctcctcctgagttgggattctagtgggtagctgcaggatacagaggcagaaaccatctaatgggaaaccactgtagagatgaatcag
tcaccaaattccagatatattgtgtagtcattccaccatg

SEQ ID NO: 41 C6ΔCH1 knock-in sequence 8505bp

tcaatccctggaccatgcagaaggggccttgtgggacttctcataccacactcagctaatcccccaaagaattgactgaagcaagacatccccccaactttgatcccctc
ccctccttgctcttcccccattcattccctcccctttccagctgcttcatcaacacaacaacttgacatccacacctcagaaacccctgtactaacagtgtcaaaggat
gcagctgggtgagcaaaggcactagtgctaccctttccagctgcttcatcaacacaacaacttgacatccacacctcagaaacccctgtactaacagtgtcaaaggat
ggaatactctgatctcccaccatcacctttttgtccatctcttaactccagtcccagtgagtcactccaactcccaagcaaagaa
tcacticcacagaaggccaccaaaagctattaccaccaaaaaggacatagaaggttaagaagtccaatgtgccaacctcatctcctccttgcctcctgtgaa
acctccatgactcttttctctgagctgtcaatgtacatcccaggtiagaacaggttactaaaaatgacttctgtactactgcaagccaggccctatatcccaggccacaac
aaatacccaccacctaacccctgagaagccatcatcctttttgagctaatgcctaaacagaggaagaggacaaggacctactagcccatagggacctgagctctgtgtatc
tgatccaaactgtgaggtcctcatggattctgtaggccatgtgaccagacaaaacaacaaagacattttctgaggtgcatagaggcctagagaccatta
caccttggtgaagactttattgtgtaccccacatgttgtcctcctaactatatctctctcctcctctcctctctctctctctctctctctctctctctctct
ctctctctctctctcacacacacacacacacacacacacacacacacacacac

FIG. 31 (Continued)

FIG. 31 (Continued)

tcttagttctgtcgactagcacctcctgacttgtctctatgatgccattcttcccctcctacccatccatcaccctgttgccaaggcactctctgagacatcttgtt
tctacatagatacagaagaacatttgattaggcatgtaggccatggtgaaagaaaaacaccaagtaaacttttaacatcacaattttatttaaaactactagagagg
actcagcaatgctcaaagagactaaattcccaggtgaaaggagatgggatgttgtattctaggcacagacagacagagcagtgggacatatgtaatcaccaagaaat
cagccaaatttttaataaacaccttagctgccgagagggatgagaacacagaagaatcccaaattctgagccaatcaccactcacttaacctcccacccaggacact
gctaagcaagaaacaccttcgagcacagtgtgggcagaatatccatcttgtaccagctccaaggcctgctgaatataagggttatcccaaagcacctagagccagaa
gtcaactatgataccccagctccagcacagctcagttcccagttccaaaaaacaggtcttccatgcctggtggataagatgatatgcacttgtatctcaccccccaga
ctctacaatccatgtatatacaatgcctagcatgcaatcagtaattacaagatacaaatgattgataaccaagacagagacagtacatgggattagacctgaggtac
gggaaaggttgggaaagcattcatagtgtcccttgctacaagaagtcccagagctttgcagggtgtgagcagattctaacctcccccattggcaagtcttctcaaag
ctgcttctgctactagagaaccctctttcccagccacctgccactatttatcacttatctgagttttccccatgattctaggatgctaccacctcctgcctgagtca
gacggtccttccaggagacctgatggtcctgcccttgcctgagacctttctaggctgaatggtcatcatgtccactgtctggttagcttggttgcctctgtgttaag
ttgcttcatactacatgtagcaaggaaagttacctttccacttctcaggacactgtaaagaagcacttctgtaaattatagtgagaaacgatcaataataaatgaaa
tgaacaaacataaacctctcttatctctatgggacctgccctaaggcctcttgtggatgcctgaaaccttgagtgttcttaagccatgtgtcacttcagtcattaac
ctgtggctgaggtggatcctgactgacaaacagacaggcatgtagacagcttgaagcctggacaaagagaaggttgacgttccaggcaaagcaagagcctatcacat
tgttccaaagtgtgtgatttagaatcagaagatttcttgtttaaaatatactcaaaaagttataaaatatcagccaggactagtcataaagcttaaatctgagagag
agagagagagagagagagagagagagaggtgttcatgtatatacatatgtgtgtgtgtatatatatgtatgtatatgtatgtatatgtgtatattggtgtttggg
tgtcagattccctggaattaaagttataaacagttatgagctgcaaagtaggctttccaatgtcttgttctccaaccataggtaattgaaactggaaactacaggta
aaggaggctgccaagtgtttcttattatgctaagattcagtgttctacaatattcacatgatttcactttgatatggtgaaatgcctttaggccaacgcatccacag
acttcacagtgacccccagctgagcttacatataatgtcaaacagcacttgttgaatagatacccaggcagccctgtgacccatggtgttttccttaattcttcccat
ctgccttggactttcacaccttgagttgagatgtgctttctgtagaaatgtatttccctaggtctgtacatcagccaccacccttttctgatcttggaccatacttca
gtctccagctgggatattcatatgcatgcttagatatggcataaccatcgttaacgtcttctgttaccttaggattaagttaccgtagctaggatgctttccaatc
tctccctccatttgcaaaggctcagggttccttctgactctggagtagtactctttcccaatgagtgcaccatgaccactattcttccttagaatgttactccccct
tctctgccttttttcatggcacatctttctgagaaaacctacctggtctagctgctctttctgggtcttcttaagtagtaacctaactcctgcaccatccgtgtggca
gtgacatagatccacatcaggtttggattcagatattccaatcccatgcagaccaatctatgtgtgagattccaggatccagcaggctggggcatgggttgaggggt
aggattctttggggcaaagtggttagtaggcacttgggactggaagtgataagtattagcgggaggagggcatatgatcaaaacattgctgctggatgtgaaagctg
taccttcaggttcagtgttcaggagaccaggttaataagaaatgatctttctgtagggtggagaccttctcatgagcactagttcttccctaggcatagtcaacacc
atccaaccactcgtgtatcatggatgagcaaagtgacagctacatggacttagaggaggagaacggcctgtggcccacaatgtgcaccttcgtggccctcttcctgct
cacactgctctacagtggcttcgtcaccttcatcaaggtaggcctcctaccacctccttctgaaccctttcctgggcacatgtaagataccttccaggggcattttg
aaacccatcggtacttctggagtggctagggaatgcacaagtaattgggaggccatggtcactgttcagcttctgggccctatgtagctccacggactttccctgtt
tctcagacaacagcctctgtatgacttcacggctctctggcaggtgaagtag

hTFR1 (SEQ ID NO: 70):

MMDQARSAFSNLFGGEPLSYTRFSLARQVDGDNSHVEMKLAVDEEENADNNTKANVTKPKRCSGSICYGTIAVIVFFLIGFMIGYLGYCKGVEPKTECERLAGTESP
VREEPGEDFPAARRLYWDDLKRKLSEKLDSTDFTGTIKLLNENSYVPREAGSQKDENLALYVENQFREFKLSKVWRDQHFVKIQVKDSAQNSVIIVDKNGRLVYLVE
NPGGYVAYSKAATVTGKLVHANFGTKKDFEDLYTPVNGSIVIVRAGKITFAEKVANAESLNAIGVLIYMDQTKFPIVNAELSFFGHAHLGTGDPYTPGFPSFNHTQF
PPSRSSGLPNIPVQTISRAAAEKLFGNMEGDCPSDWKTDSTCRMVTSESKNVKLTVSNVLKEIKILNIFGVIKGFVEPDHYVVVGAQRDAWGPGAAKSGVGTALLLK
LAQMFSDMVLKDGFQPSRSIIFASWSAGDFGSVGATEWLEGYLSSLHLKAFTYINLDKAVLGTSNFKVSASPLLYTLIEKTMQNVKHPVTGQFLYQDSNWASKVEKL
TLDNAAFPFLAYSGIPAVSFCFCEDTDYPYLGTTMDTYKELIERIPELNKVARAAAEVAGQFVIKLTHDVELNLDYERYNSQLLSFVRDLNQYRADIKEMGLSLQWL
YSARGDFFRATSRLTTDFGNAEKTDRFVMKKLNDRVMRVEYHFLSPYVSPKESPFRHVFWGSGSHTLPALLENLKLRKQNNGAFNETLFRNQLALATWTIQGAANAL
SGDVWDIDNEF

FasTFR1 VH (SEQ ID NO: 71):

FIG. 31 (Continued)

MERVQPLEENVGNAARPRFERNKLLLVASVIQGLGLLLCFTYICLHFSALGGGGSCRGVEPKTECERLAGTESPAREEPEEDFPAAPRLYWDDLKRKLSEKLDTTDF
TSTIKLLNENLYVPREAGSQKDENLALYIENQFREFKLSKVWRDQHFVKIQVKDSAQNSVIIVDKNGGLVYLVENPGGYVAYSKAATVTGKLVHANFGTKKDFEDLD
SPVNGSIVIVRAGKITFAEKVANAESLNAIGVLIYMDQTKFPIVKADLSFFGHAHLGTGDPYTPGFPSFNHTQFPPSQSSGLPNIPVQTISRAAAEKLFGNMEGDCP
SDWKTDSTCRMVTSENKSVKLTVSNVLKETKILNIFGVIKGFVEPDHYVVVGAQRDAWGPGAAKSSVGTALLLKLAQMFSDMVLKDGFQPSRSIIFASWSAGDFGSV
GATEWLEGYLSSLHLKAFTYINLDKAVLGTSNFKVSASPLLYTLIEKTMQDVKHPVTGRSLYQDSNWASKVEKLTLDNAAFPFLAYSGIPAVSFCFCEDTDYPYLGT
TMDTYKELVERIPELNKVARAAAEVAGQFVIKLTHDTELNLDYERYNSQLLLFLRDLNQYRADVKEMGLSLQWLYSARGDFFRATSRLTTDFRNAEKRDKFVMKKLN
DRVMRVEYYFLSPYVSPKESPFRHVFWGSGSHTLSALLESLKLRRQNNSAFNETLFRNQLALATWTIQGAANALSGDVWDIDNEF

JR141 VH (SEQ ID NO: 72): EVQLVQSGAEVKKPGESLKISCKGSGYSFTNYWLGWVRQMPGKGLEWMGDIYPGGDYPTYSEKFKVQVTISADKSISTAYL
QWSSLKASDTAMYYCARSGNYDEVAYWGQGTLVTVSS

JR141 VL (SEQ ID NO: 73): DIVMTQTPLSLSVTPGQPASISCRSSQSLVHSNGNTYLHWYLQKPGQSPQLLIYKVSNRFSGVPDRFSGSGSGTDFTLKIS
RVEAEDVGVYYCSQSTHVPWTFGQGTKVEIK

Chimeric TFR1 protein (SEQ ID NO: 74)
MMDQARSAFSNLFGGEPLSYTRFSLARQVDGDNSHVEMKLAADEENADNNMKASVRKPKRFNGRLCFAAIALVIFFLIGFMSGYLGYCKGVEPKTECERLAGTESP
VREEPGEDFPAARRLYWDDLKRKLSEKLDSTDFTGTIKLLNENSYVPREAGSQKDENLALYVENQFREFKLSKVWRDQHFVKIQVKDSAQNSVIIVDKNGRLVYLVE
NPGGYVAYSKAATVTGKLVHANFGTKKDFEDLYTPVNGSIVIVRAGKITFAEKVANAESLNAIGVLIYMDQTKFPIVNAELSFFGHAHLGTGDPYTPGFPSFNHTQF
PPSRSSGLPNIPVQTISRAAAEKLFGNMEGDCPSDWKTDSTCRMVTSESKNVKLTVSNVLKEIKILNIFGVIKGFVEPDHYVVVGAQRDAWGPGAAKSGVGTALLLK
LAQMFSDMVLKDGFQPSRSIIFASWSAGDFGSVGATEWLEGYLSSLHLKAFTYINLDKAVLGTSNFKVSASPLLYTLIEKTMQNVKHPVTGQFLYQDSNWASKVEKL
TLDNAAFPFLAYSGIPAVSFCFCEDTDYPYLGTTMDTYKELIERIPELNKVARAAAEVAGQFVIKLTHDVELNLDYERYNSQLLSFVRDLNQYRADIKEMGLSLQWL
YSARGDFFRATSRLTTDFGNAEKTDRFVMKKLNDRVMRVEYHFLSPYVSPKESPFRHVFWGSGSHTLPALLENLKLRKQNNGAFNETLFRNQLALATWTIQGAANAL
SGDVWDIDNEF

Human IgG1 Fc region (SEQ ID NO: 75)
DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKA
LPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHN
HYTQKSLSLSPGK

Human IgG1 Fc region with N297A mutation (SEQ ID NO: 76)
DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYASTYRVVSVLTVLHQDWLNGKEYKCKVSNKA
LPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHN
HYTQKSLSLSPGK

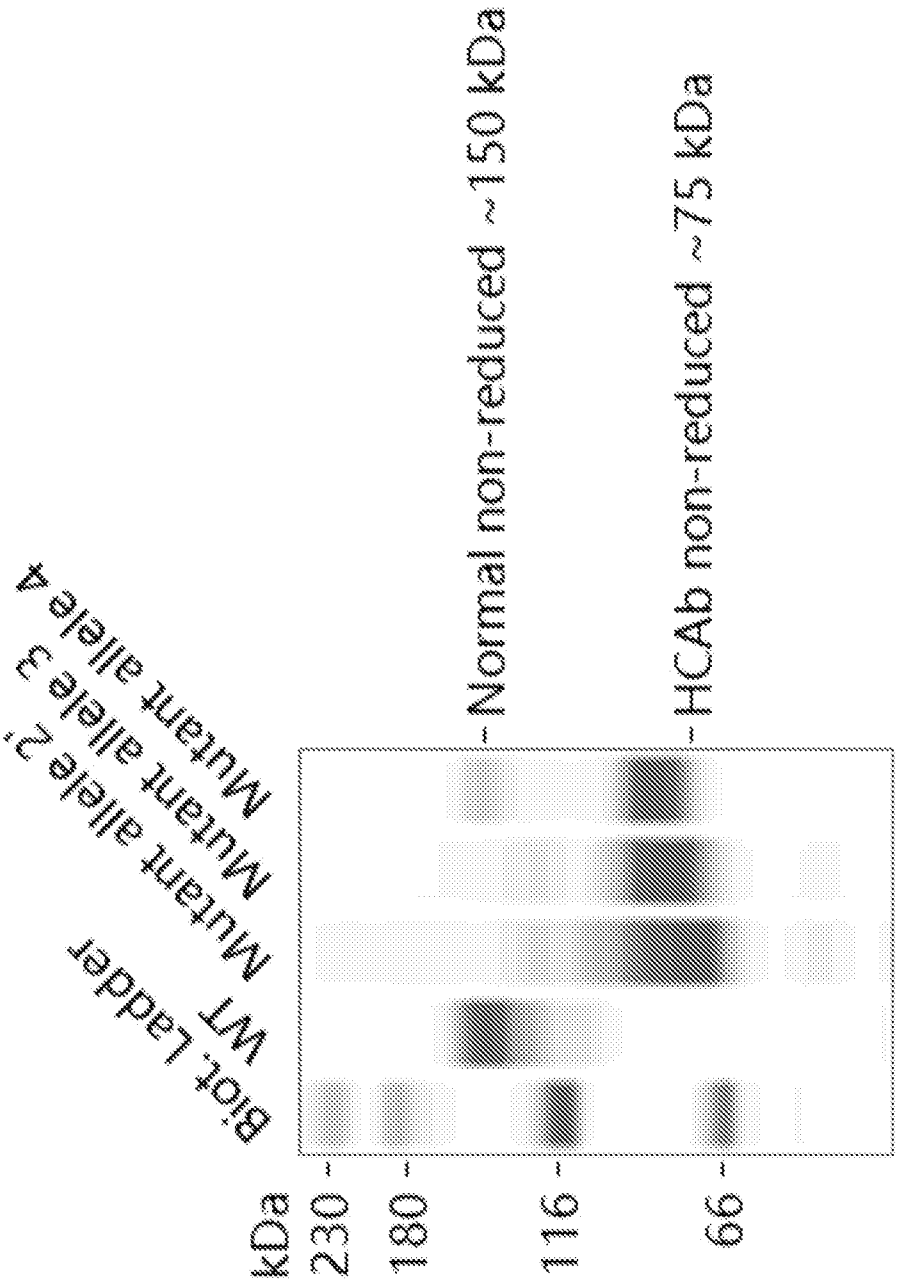

kDa
230
180
116
66
Biotᵢ Ladder
WT
Mutant allele 2
Mutant allele 3
Mutant allele 4
Normal non-reduced ~150 kDa
HCAb non-reduced ~75 kDa
HCAb: heavy-chain antibody

FIG. 32

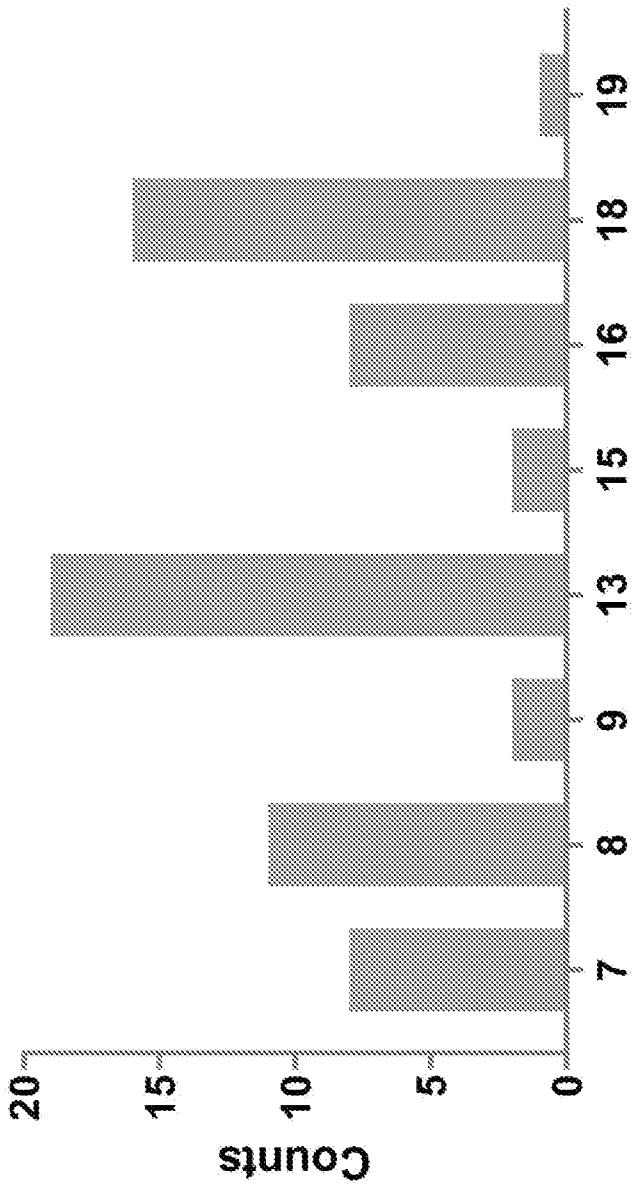


FIG. 33

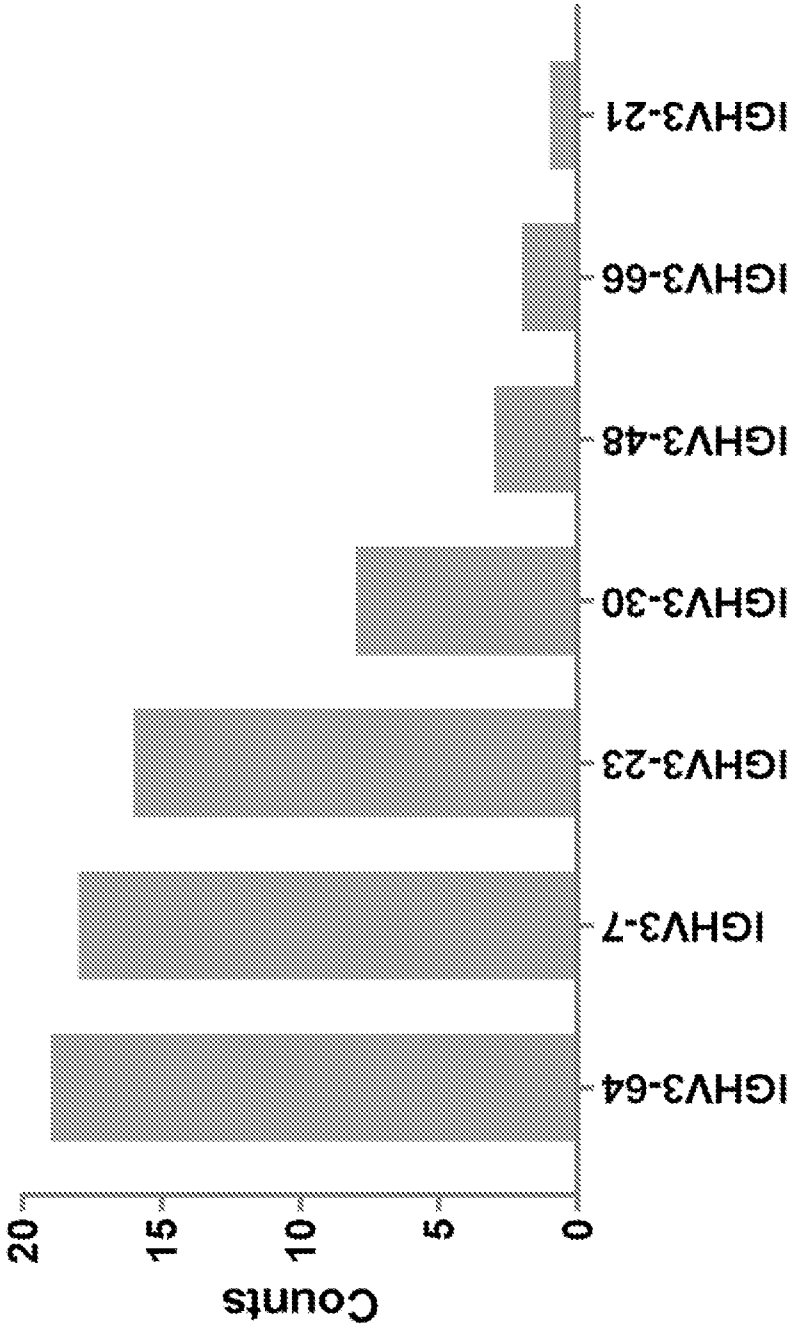

IGHV3-64
IGHV3-7
IGHV3-23
IGHV3-30
IGHV3-48
IGHV3-66
IGHV3-21
0
5
10
15
20
Counts

FIG. 34

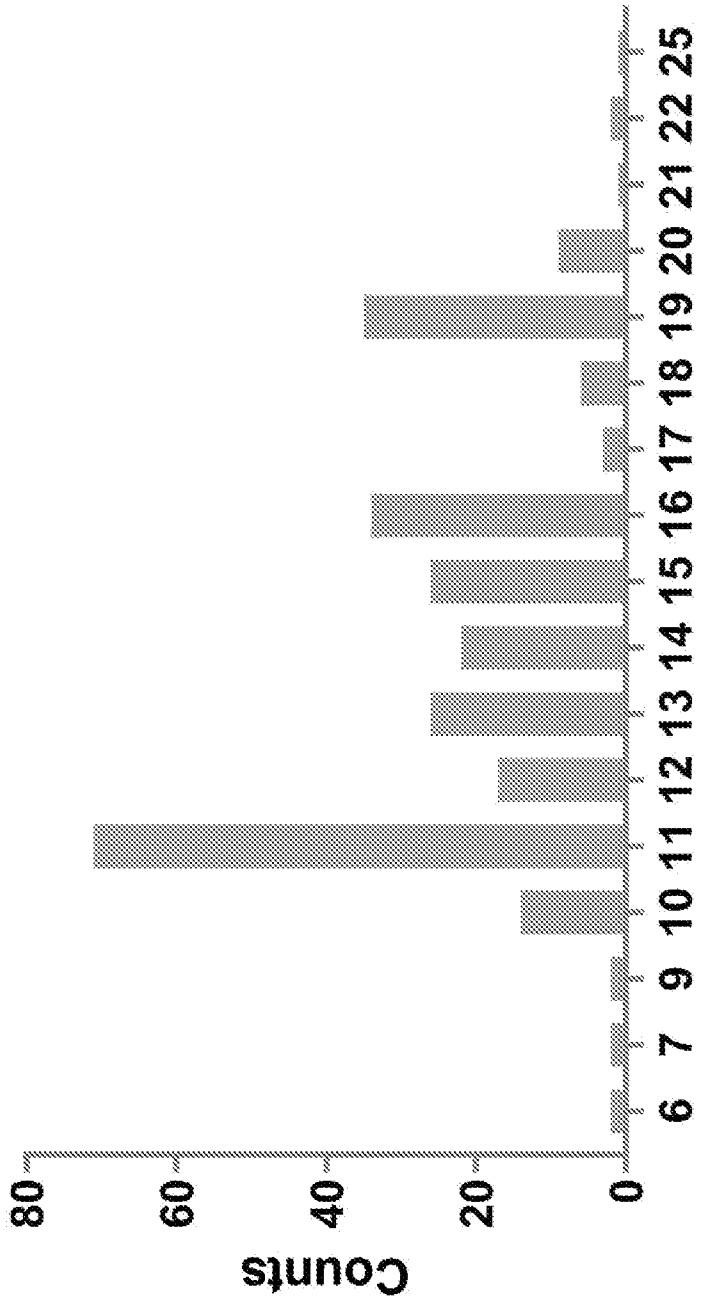


FIG. 35

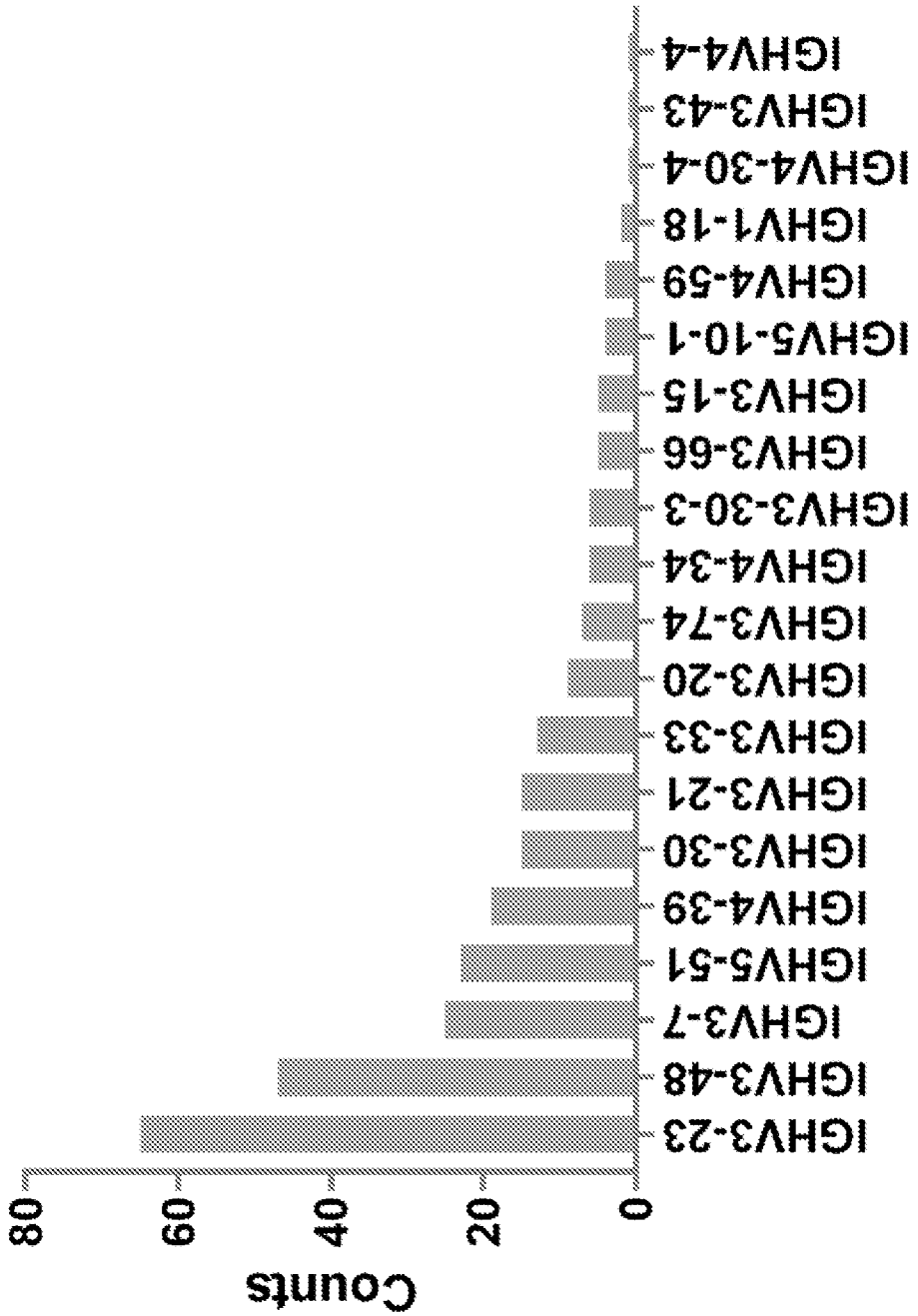

Counts
0
20
40
60
80
IGHV3-23
IGHV3-48
IGHV3-7
IGHV5-51
IGHV4-39
IGHV3-30
IGHV3-21
IGHV3-33
IGHV3-20
IGHV3-74
IGHV4-34
IGHV3-30-3
IGHV3-66
IGHV3-15
IGHV5-10-1
IGHV4-59
IGHV1-18
IGHV4-30-4
IGHV3-43
IGHV4-4

FIG. 36

Kabat CDR for anti-TFR1 antibodies

Antibody	VHH CDR1	SEQ ID NO	VHH CDR2	SEQ ID NO	VHH CDR3	SEQ ID NO
23B8	NAWMN	42	HIKSKTDGGTTDYAAPVKG	43	DRAYCGGDCSHYYHDLDV	44
24A1	SYAMS	45	EISGSGSSTNYADSVKG	46	DDSSGWAFNI	47
24C9	SYGMY	48	FIRFDESKKYYADSVKG	49	DGEGGYCTNGVCHGLDV	50
24G5	SFAMS	51	EISDSGGNTYYADSVKG	52	DDSSGWAFNI	53

FIG. 37

IMGT CDR for anti-TFR1 antibodies

Antibody	VHH CDR1	SEQ ID NO	VHH CDR2	SEQ ID NO	VHH CDR3	SEQ ID NO
23B8	GFPFSNAW	54	IKSKTDGGTT	55	TTDRAYCGGDCSHYYY HDLDV	56
24A1	GFTFSSYA	57	ISGSGSST	58	AKDDSSGWAFNI	59
24C9	GFTLSSYG	60	IRFDESKK	61	ARDGEGGYCTNGVCH GLDV	62
24G5	GFTFSSFA	63	ISDSGGNT	64	AKDDSSGWAFNI	65

FIG. 38

Antibody	Amino acid sequence of VHH
23B8	QVTLKESGGGLVKPGGSLRLSCAASGFPFSNAWMNWVRQAPGKGLEWVGHIKSKTDGGTTDYAAPVKGRFTIS RDDSKNMVFLQMNSLKTEDTAVYYCTTDRAYCGGDCSHYYYHDLDVWGQGTTVTVSS **(SEQ ID NO: 66)**
24A1	EVQLVESGGGLVQPGGSLRLTCAASGFTFSSYAMSWVRQAPGKGLERVSEISGSGSSTNYADSVKGRFTISRDNSK NTLYLQMNSLRAEDTAVYYCAKDDSSGWAFNIWDQGTVVTVSS **(SEQ ID NO: 67)**
24C9	EVQLVQSGGGVVQPGRSLRLSCAASGFTLSSYGMYWVRQAPGKGREWVAFIRFDESKKYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDGEGGYCTNGVCHGLDVWGQGTTVTVSS **(SEQ ID NO: 68)**
24G5	QLQLQESGGGLVQPGGSLRLTCAASGFTFSSFAMSWVRQAPGKGLERVSEISDSGGNTYYADSVKGRFTISRDNS KNTLYLQMNSLRAEDTAVYYCAKDDSSGWAFNIWDQGTVVTVSS **(SEQ ID NO: 69)**

FIG. 39

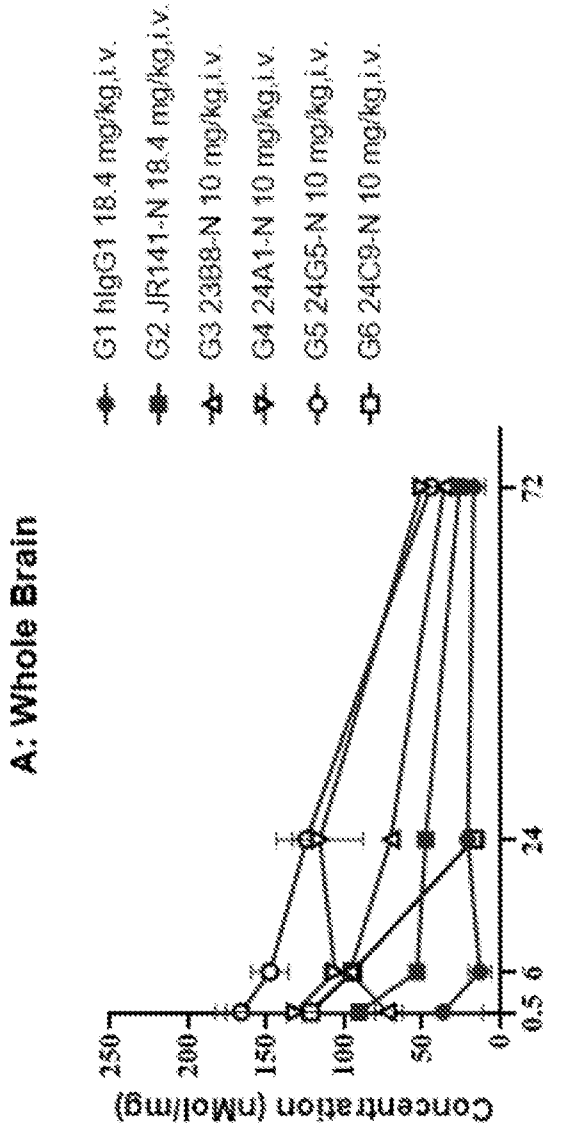

A: Whole Brain
G1 hIgG1 18.4 mg/kg,i.v.
G2 JR141-N 18.4 mg/kg,i.v.
G3 2388-N 10 mg/kg,i.v.
G4 24A1-N 10 mg/kg,i.v.
G5 24G5-N 10 mg/kg,i.v.
G6 24C9-N 10 mg/kg,i.v.
Concentration (nMol/mg)
250
200
150
100
50
0
0.5
6
24
72

FIG. 40A

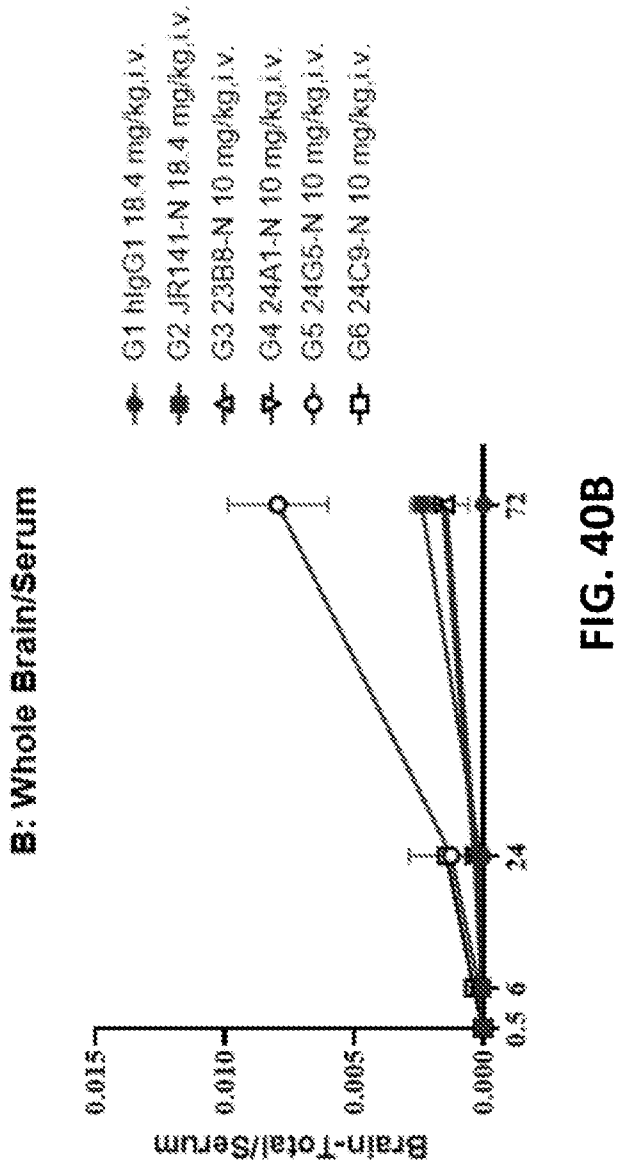

B: Whole Brain/Serum
Brain-Total/Serum
0.015
0.010
0.005
0.000
0.5
6
24
72
G1 hIgG1 18.4 mg/kg,i.v.
G2 JR141-N 18.4 mg/kg,i.v.
G3 23B8-N 10 mg/kg,i.v.
G4 24A1-N 10 mg/kg,i.v.
G5 24G5-N 10 mg/kg,i.v.
G6 24C9-N 10 mg/kg,i.v.

FIG. 40B

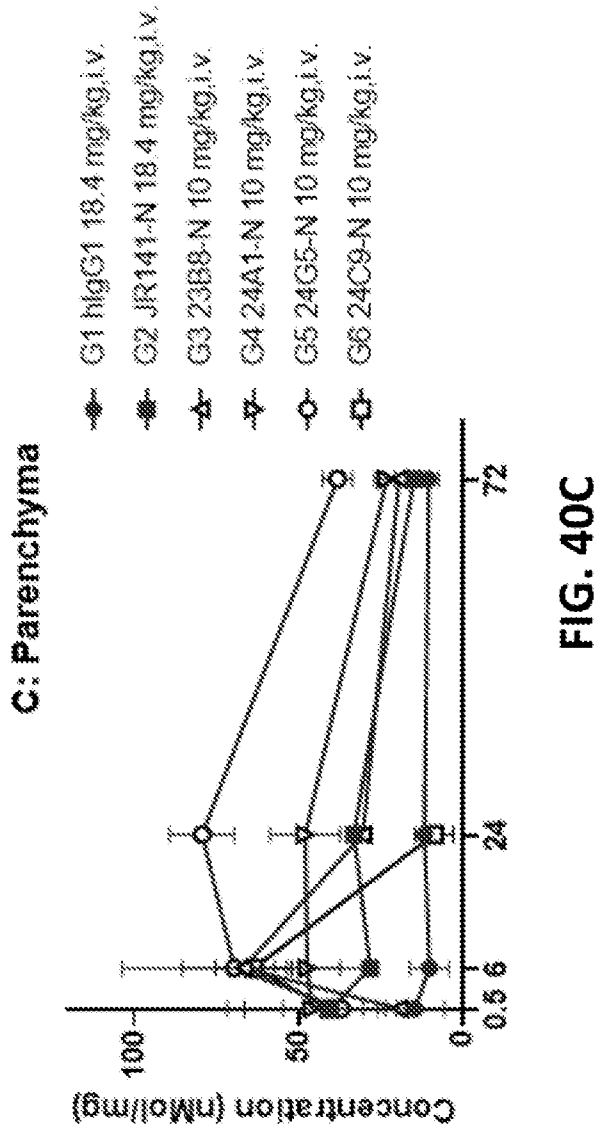

C: Parenchyma
G1 hIgG1 18.4 mg/kg,i.v.
G2 JR141-N 18.4 mg/kg,i.v.
G3 23B8-N 10 mg/kg,i.v.
G4 24A1-N 10 mg/kg,i.v.
G5 24G5-N 10 mg/kg,i.v.
G6 24C9-N 10 mg/kg,i.v.
Concentration (nMol/mg)
100
50
0
0.5
6
24
72

FIG. 40C

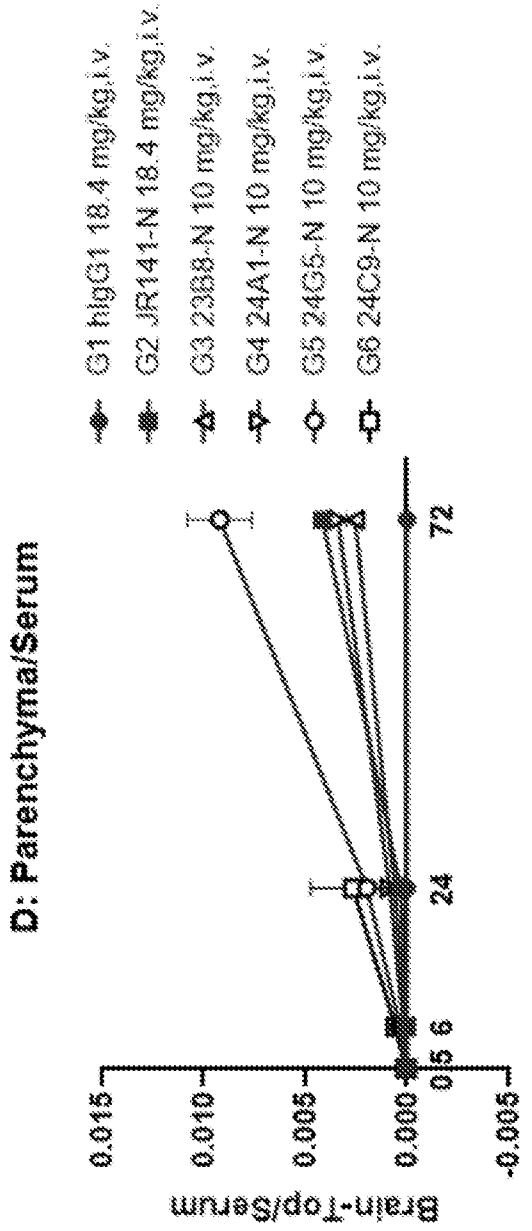

D: Parenchyma/Serum
G1 hIgG1 18.4 mg/kg,i.v.
G2 JR141-N 18.4 mg/kg,i.v.
G3 23B8-N 10 mg/kg,i.v.
G4 24A1-N 10 mg/kg,i.v.
G5 24G5-N 10 mg/kg,i.v.
G6 24C9-N 10 mg/kg,i.v.
Brain-Top/Serum
0.015
0.010
0.005
0.000
-0.005
0.5
6
24
72

FIG. 40D

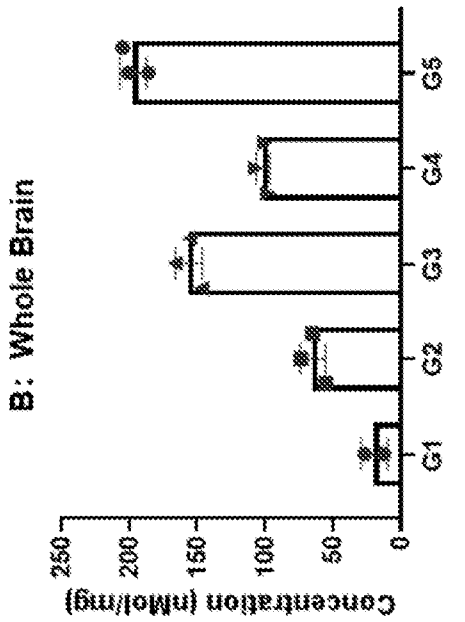


FIG. 41B

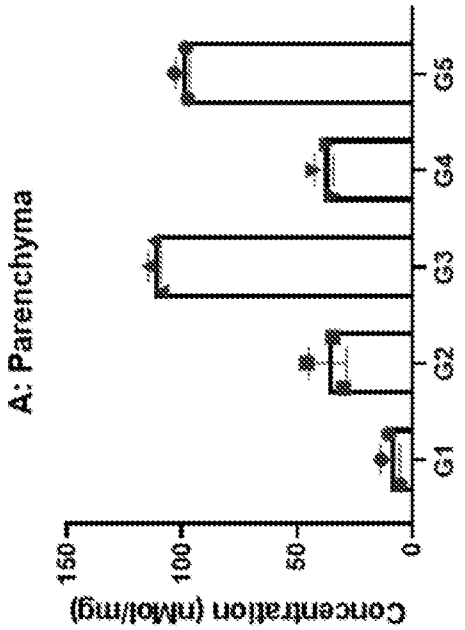


FIG. 41A

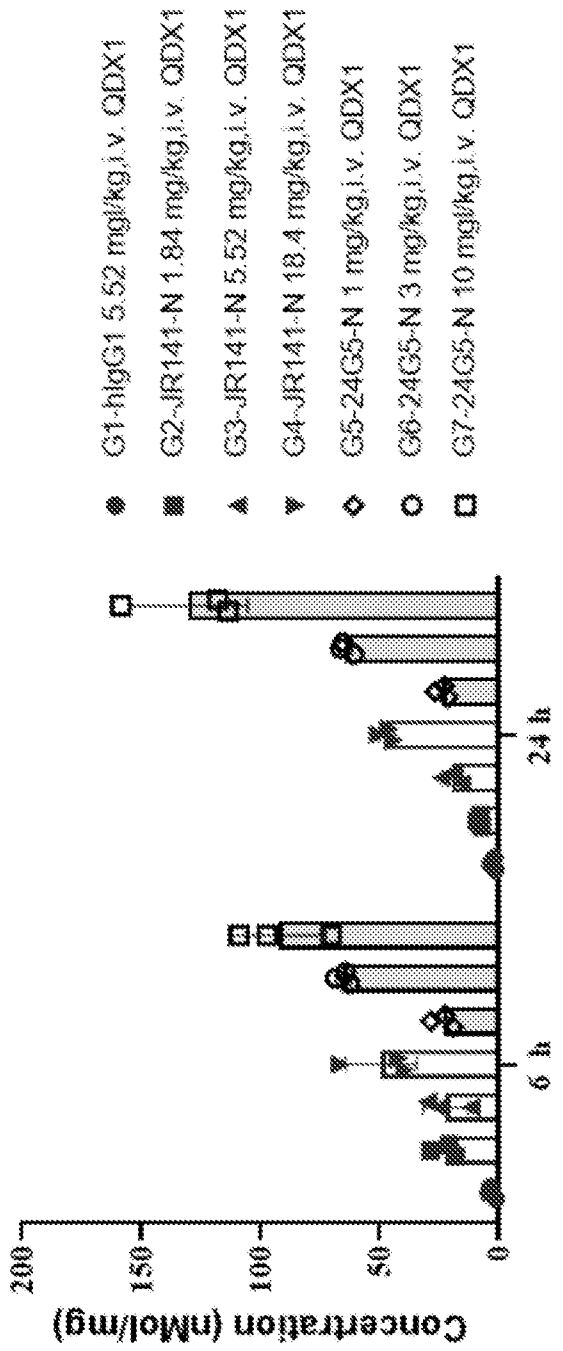

G1-hIgG1 5.52 mg/kg,i.v. QDX1
G2-JR141-N 1.84 mg/kg,i.v. QDX1
G3-JR141-N 5.52 mg/kg,i.v. QDX1
G4-JR141-N 18.4 mg/kg,i.v. QDX1
G5-24G5-N 1 mg/kg,i.v. QDX1
G6-24G5-N 3 mg/kg,i.v. QDX1
G7-24G5-N 10 mg/kg,i.v. QDX1
Concentration (nMol/mg)
200
150
100
50
0
6 h
24 h

FIG. 42

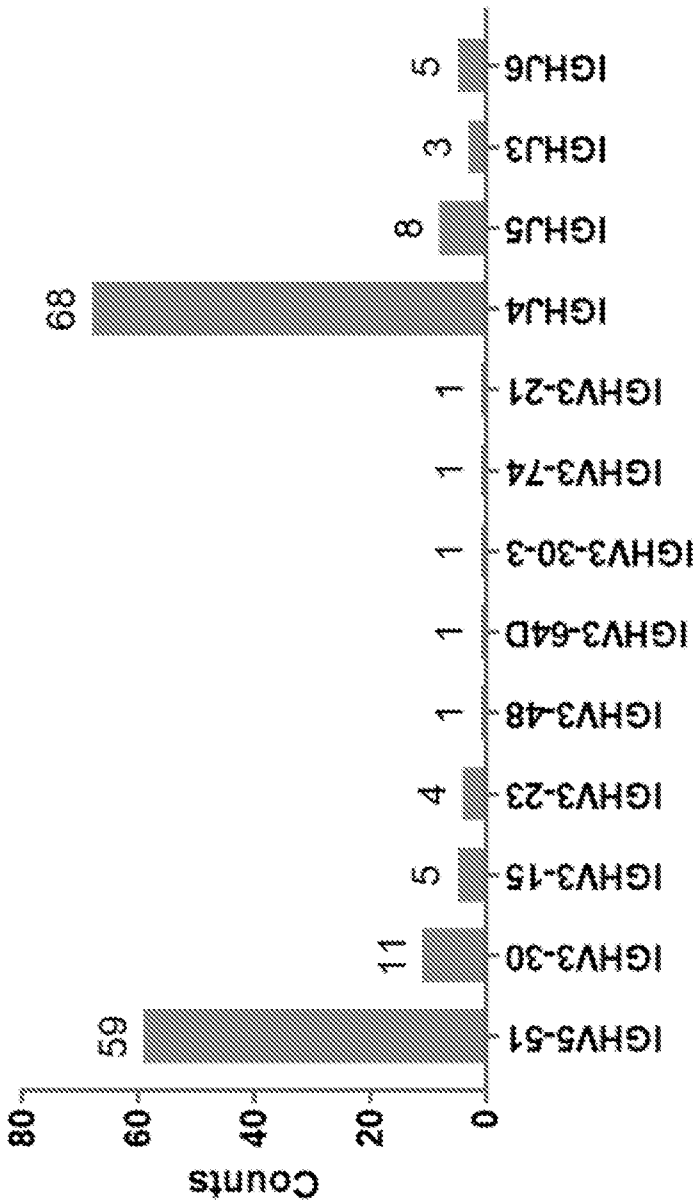

IGHV5-51
IGHV3-30
IGHV3-15
IGHV3-23
IGHV3-48
IGHV3-64D
IGHV3-30-3
IGHV3-74
IGHV3-21
IGHJ4
IGHJ5
IGHJ3
IGHJ6
59
11
5
4
1
1
1
1
1
68
8
3
5
0
20
40
60
80
Counts

FIG. 43


Counts
0
20
40
60
IGHV4-39
IGHV3-66
IGHV3-48
IGHV3-23
IGHV3-49
IGHV4-34
IGHV3-30
IGHV3-7
IGHV3-21
IGHV4-4
IGHV3-15
IGHV3-53
IGHV3-33
IGHV4-59
IGHV3-74
IGHV3-43
IGHJ4
IGHJ3
IGHJ6
IGHJ5
26
16
12
9
6
5
4
4
4
3
2
1
1
1
1
1
55
17
14
10

FIG. 44

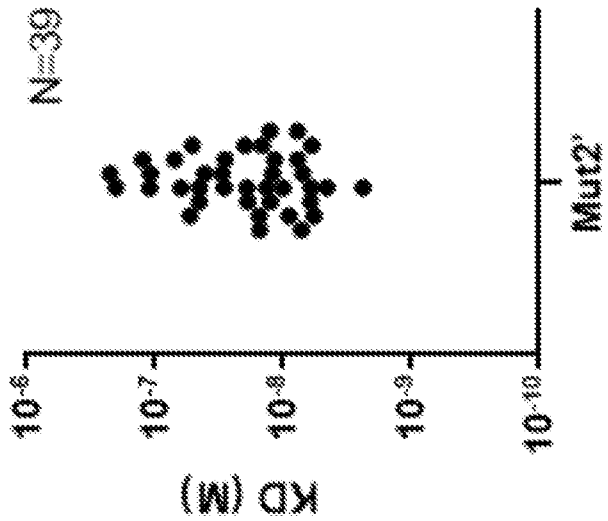


FIG. 45

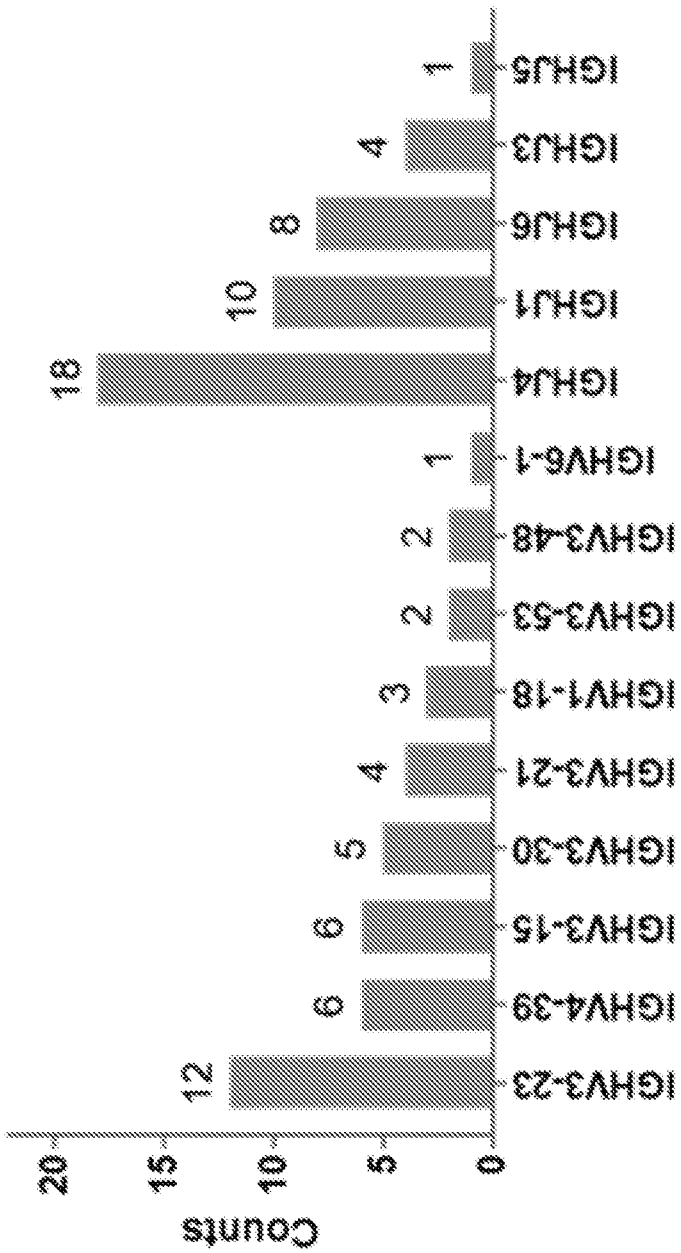

Counts
0
5
10
15
20
IGHV3-23
12
IGHV4-39
6
IGHV3-15
6
IGHV3-30
5
IGHV3-21
4
IGHV1-18
3
IGHV3-53
2
IGHV3-48
2
IGHV6-1
1
IGHJ4
18
IGHJ1
10
IGHJ6
8
IGHJ3
4
IGHJ5
1

FIG. 46

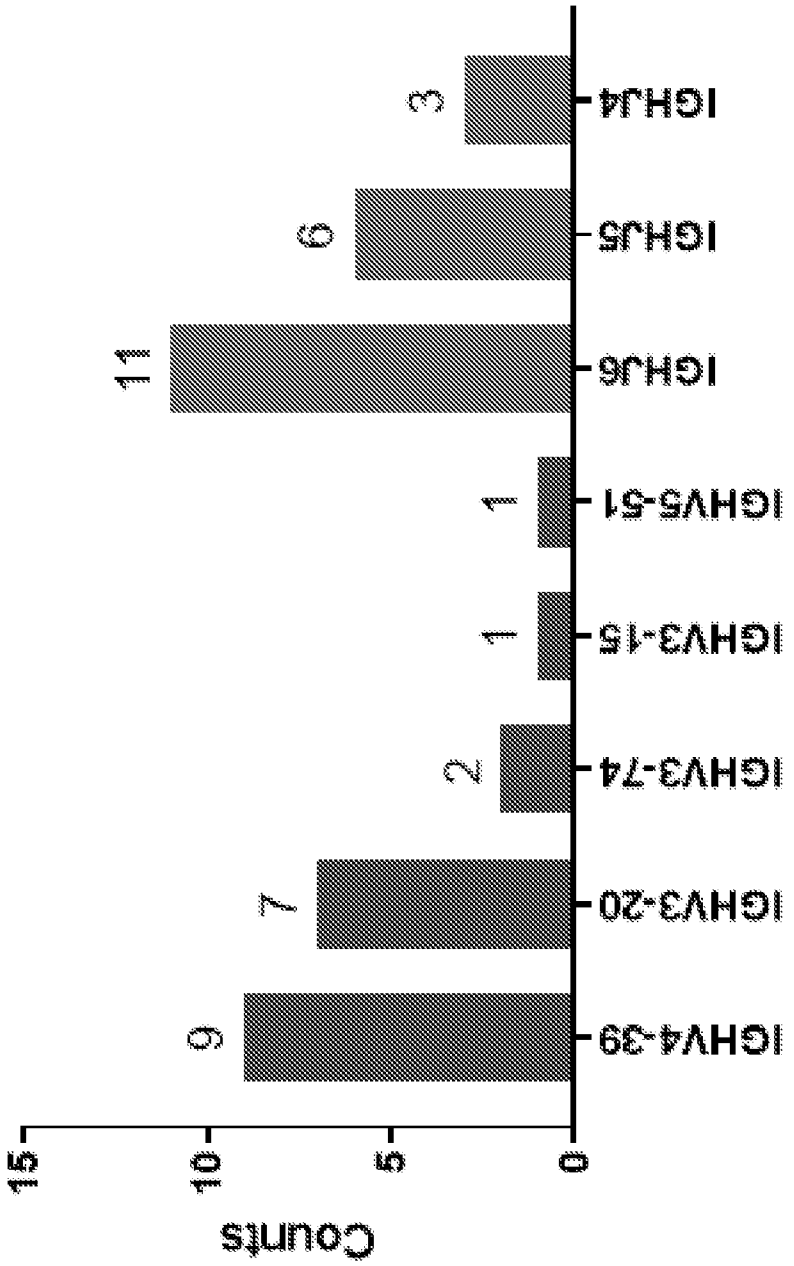

Counts
15
10
5
0
9
7
2
1
1
11
6
3
IGHV4-39
IGHV3-20
IGHV3-74
IGHV3-15
IGHV5-51
IGHJ6
IGHJ5
IGHJ4

FIG. 47

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2023/118958

A. CLASSIFICATION OF SUBJECT MATTER

C12N15/85(2006.01)i; A01K67/027(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: C12N, A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT, VEN: heavy chain locus, IgG, CH1, antibody

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104131035 A (REGENERON PHARMACEUTICALS, INC.) 05 November 2014 (2014-11-05) see claims 1-31,paragraphs 56-138	1-4,44-53,61-85
X	CN 109906030 A (AKEAGEN LLC) 18 June 2019 (2019-06-18) see claims 1-44,paragraphs 32-103	1-4,44-53,61-85
A	CN 102711449 A (Regeneron Pharmaceuticals, Inc.) 03 October 2012 (2012-10-03) the whole document	1-85
A	US 2013247236 A1 (REGN) REGENERON PHARM INC;(REGN) REGENERON PHARMA;19 September 2013 (2013-09-19) the whole document	1-85
A	US 2019292263 A1 (REGN) REGENERON PHARM INC 26 September 2019 (2019-09-26) the whole document	1-85

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
08 November 2023	**05 January 2024**

Name and mailing address of the ISA/CN	Authorized officer
CHINA NATIONAL INTELLECTUAL PROPERTY ADMINISTRATION **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088, China**	**WANG,XiaBing** Telephone No. **(+86) 010-62085146**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2023/118958

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 104131035 A	05 November 2014	None	
CN 109906030 A	18 June 2019	US 2020344983 A1	05 November 2020
		US 2019261611 A1	29 August 2019
		US 10660316 B2	26 May 2020
		WO 2018144097 A1	09 August 2018
CN 102711449 A	03 October 2012	WO 2011072204 A1	16 June 2011
		MY 172472 A	26 November 2019
		SI 2954779 T1	28 June 2019
		CA 2782936 A1	16 June 2011
		CA 2782936 C	18 June 2019
		US 2015197556 A1	16 July 2015
		JP 2015133991 A	27 July 2015
		JP 6293702 B2	14 March 2018
		IL 237819 A	30 March 2017
		HK 1177100 A1	16 August 2013
		US 2015196015 A1	16 July 2015
		MY 173363 A	21 January 2020
		LT 2954779 T	27 May 2019
		EP 3504964 A1	03 July 2019
		KR 20140024970 A	03 March 2014
		KR 101553244 B1	15 September 2015
		US 2015197555 A1	16 July 2015
		LT 2509409 T	12 December 2016
		IL 220150 A0	31 July 2012
		KR 20130055556 A	28 May 2013
		KR 101430514 B1	19 August 2014
		RS 58709 B1	28 June 2019
		TR 201906488 T4	21 May 2019
		HUE 043556 T2	30 September 2019
		US 2015197553 A1	16 July 2015
		US 11234419 B2	01 February 2022
		CY 1118163 T1	28 June 2017
		US 2011145937 A1	16 June 2011
		US 8754287 B2	17 June 2014
		HRP 20190820 T1	20 September 2019
		AU 2010328046 A1	12 July 2012
		AU 2010328046 B2	31 July 2014
		AU 2010328046 C1	18 January 2018
		EP 2509409 A1	17 October 2012
		EP 2509409 B1	27 July 2016
		SI 2509409 T1	30 December 2016
		PL 2954779 T3	31 July 2019
		ES 2724975 T3	18 September 2019
		JP 2013513388 A	22 April 2013
		JP 5909449 B2	26 April 2016
		IL 237818 A	31 August 2016
		ES 2595376 T3	29 December 2016
		US 2015197557 A1	16 July 2015
		EP 2954779 A1	16 December 2015
		EP 2954779 B1	06 February 2019

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2023/118958

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
		US	2014289876	A1	25 September 2014
		US	9888675	B2	13 February 2018
		HK	1220865	A1	19 May 2017
		HUE	030208	T2	28 April 2017
		SG	10201408181	PA	29 January 2015
		DK	2954779	T3	13 May 2019
		US	2015197554	A1	16 July 2015
		PL	2509409	T3	28 February 2017
		DK	2509409	T3	14 November 2016
		PT	2509409	T	29 September 2016
		PT	2954779	T	29 May 2019
		US	2022201993	A1	30 June 2022
		IL	257822	A	30 April 2018
		IL	257822	B	30 January 2020
		JP	2017140058	A	17 August 2017
		JP	6615830	B2	04 December 2019
		CY	1121608	T1	29 May 2020
		RU	2012128856	A	20 January 2014
		RU	2603102	C2	20 November 2016
		SG	181477	A1	30 July 2012
		JP	2019141095	A	29 August 2019
		JP	2015077147	A	23 April 2015
		JP	6271450	B2	31 January 2018
US 2013247236	A1 19 September 2013	NZ	730271	A	30 September 2022
		SI	2825037	T1	30 August 2019
		RS	59076	B1	30 September 2019
		KR	20210018510	A	17 February 2021
		KR	102345232	B1	30 December 2021
		US	2022217955	A1	14 July 2022
		NZ	629077	A	24 December 2015
		IL	264362	A	28 February 2019
		IL	264362	B	31 May 2020
		JP	2015510769	A	13 April 2015
		JP	6228589	B2	08 November 2017
		RU	2014141307	A	10 May 2016
		RU	2664473	C2	17 August 2018
		IL	274268	A	30 June 2020
		IL	274268	B	01 December 2021
		EP	2825037	A1	21 January 2015
		EP	2825037	B1	01 May 2019
		ES	2737990	T3	17 January 2020
		IL	244098	A0	21 April 2016
		IL	244098	B	30 April 2018
		SG	11201405171	WA	30 October 2014
		KR	20140135737	A	26 November 2014
		KR	102213535	B1	08 February 2021
		MX	2021000484	A	08 June 2022
		US	9801362	B2	31 October 2017
		HUE	045537	T2	30 December 2019
		WO	2013138712	A1	19 September 2013

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2023/118958

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
		JP 2017131251 A	03 August 2017
		JP 6574811 B2	11 September 2019
		LT 2825037 T	12 August 2019
		HRP 20191333 T1	18 October 2019
		JP 2021007409 A	28 January 2021
		SG 10202107388 VA	30 August 2021
		MX 2014011048 A	10 April 2015
		MX 355944 B	07 May 2018
		PT 2825037 T	07 August 2019
		EP 3539374 A1	18 September 2019
		PL 2825037 T3	31 October 2019
		PL 2825037 T4	31 March 2020
		DK 2825037 T3	29 July 2019
		IL 234237 B	28 February 2019
		CA 2865645 A1	19 September 2013
		RU 2018128915 A	18 February 2019
		BR 112014022855 A2	18 July 2017
		NZ 712508 A	29 June 2018
		SG 10201700360 VA	30 March 2017
		US 2017367308 A1	28 December 2017
		US 11224207 B2	18 January 2022
		US 2014082760 A1	20 March 2014
		US 9648856 B2	16 May 2017
		KR 20220002718 A	06 January 2022
		KR 102459666 B1	27 October 2022
		HK 1200270 A1	07 August 2015
		JP 2018164462 A	25 October 2018
		JP 7011549 B2	10 February 2022
US 2019292263 A1	26 September 2019	EP 3772927 A1	17 February 2021
		IL 277353 A	30 November 2020
		AU 2019242551 A1	22 October 2020
		JP 2021519062 A	10 August 2021
		RU 2020131383 A	26 April 2022
		KR 20200135397 A	02 December 2020
		WO 2019190922 A1	03 October 2019
		US 2021051929 A1	25 February 2021
		BR 112020019026 A2	29 December 2020
		TW 202003564 A	16 January 2020
		MX 2020009989 A	14 October 2020
		CA 3093995 A1	03 October 2019
		SG 11202008799 PA	29 October 2020
		JP 2023089284 A	27 June 2023