MICE THAT MAKE BINDING PROTEINS COMPRISING VL DOMAINS

FIELD OF INVENTION

[0001] Immunoglobulin-like binding proteins comprising an immunoglobulin heavy chain constant region fused with an immunoglobulin light chain variable domain are provided, as well as binding proteins having an immunoglobulin light chain variable domain fused to a light chain constant domain and an immunoglobulin light chain variable domain fused to a heavy chain constant domain. Cells expressing such binding proteins, mice that make them, and related methods and compositions are also provided.

BACKGROUND

[0002] Antibodies typically comprise a tetrameric structure having two identical heavy chains that each comprise a heavy chain constant region (C_H) fused with a heavy chain variable domain (V_H) associated with a light chain constant region (C_L) fused with a light chain variable domain (V_L). For a typical human IgG, an antibody molecule is approximately about 150 kDa to about 170 kDa in size (*e.g.*, for IgG3, which comprises a longer hinge region), depending on the subclass of IgG (*e.g.*, IgG1, IgG3, IgG4) and (varying) length of the variable region.

[0003] In a typical antibody, V_H and V_L domains associate to form a binding site that binds a target antigen. Characteristics of the antibody with respect to affinity and specificity therefore can depend in large part on characteristics of the V_H and V_L domains. In typical antibodies in humans and in mice, V_H domains couple with either λ or κ V_L domains. It is also known, however, that V_L domains can be made that specifically bind a target antigen in the absence of a cognate V_H domain (*e.g.*, a V_H domain that naturally expresses in the context of an antibody and is associated with the particular V_L domain), and that V_H domains can be isolated that specifically bind a target antigen in the absence of a cognate V_L domain. Thus, useful diversity in immunoglobulin-based binding proteins is generally conferred by recombination leading to a particular V_H or V_L (and somatic hypermutation, to the extent that it occurs), as well as by combination of a cognate V_H/V_L pair. It would be useful to develop compositions and methods to exploit other sources of diversity.

[0004] There is a need in the art for binding proteins based on immunoglobulin structures, including immunoglobulin variable regions such as light chain variable regions, and including binding proteins that exhibit enhanced diversity over traditional antibodies. There is also a need for further methods and animals for making useful binding proteins, including binding proteins that comprise diverse light chain immunoglobulin variable region sequences. Also in need are useful formats for immunoglobulin-based binding proteins that provide an enhanced diversity of binding proteins from which to choose, and enhanced

diversity of immunoglobulin variable domains, including compositions and methods for generating somatically mutated and clonally selected immunoglobulin variable domains for use, *e.g.*, in making human therapeutics.

SUMMARY

[0005] In one aspect, binding proteins are described that comprise immunoglobulin variable domains that are derived from light chain (*i.e.*, kappa (κ) and/or lambda (λ)) immunoglobulin variable domains, but not from full-length heavy chain immunoglobulin variable domains. Methods and compositions for making binding proteins, including genetically modified mice, are also provided.

[0006] In one aspect, nucleic acids constructs, cells, embryos, mice, and methods are provided for making proteins that comprise one or more κ and/or λ light chain variable region immunoglobulin sequences and an immunoglobulin heavy chain constant region sequence, including proteins that comprise a human λ or κ light chain variable domain and a human or mouse heavy chain constant region sequence.

[0007] In one aspect, a mouse is provided, comprising an immunoglobulin heavy chain locus comprising a replacement of one or more immunoglobulin heavy chain variable region (V_H) gene segments, heavy chain diversity (D_H) gene segments, and heavy chain joining (J_H) gene segments at an endogenous mouse immunoglobulin heavy chain locus with one or more light chain variable region (V_L) gene segments and one or more light chain joining region (J_L) gene segments.

[0008] In one aspect, a mouse is provided, comprising an immunoglobulin heavy chain locus that comprises a replacement of all or substantially all V_H, D_H, and J_H gene segments with one or more V_L gene segments and one or more J_L gene segments to form a V_L gene segment sequence at an endogenous heavy chain locus (VL_H locus), wherein the VL_H locus is capable of recombining with an endogenous mouse C_H gene to form a rearranged gene that is derived from a V_L gene segment, a J_L gene segment, and an endogenous mouse C_H gene.

[0009] In one embodiment, the V_L segments are human V_L. In one embodiment, the J_L segments are human J_L. In a specific embodiment, the V_L and J_L segments are human V_L and human J_L segments.

[0010] In one embodiment, all or substantially all V_H, D_H, and J_H gene segments are replaced with at least six human Vκ gene segments and at least one Jκ gene segment. In one embodiment, all or substantially all V_H, D_H, and J_H gene segments are replaced with at least 16 human Vκ gene segments (human Vκ) and at least one Jκ gene segment. In one embodiment, all or substantially all V_H, D_H, and J_H gene segments are replaced with at least

30 human Vκ gene segments and at least one Jκ gene segment. In one embodiment, all or substantially all V_H, D_H, and J_H gene segments are replaced with at least 40 human Vκ gene segments and at least one Jκ gene segment. In one embodiment, the at least one Jκ gene segment comprises two, three, four, or five human Jκ gene segments.

[0011] In one embodiment, the V_L segments are human Vκ segments. In one embodiment, the human Vκ segments comprise 4-1, 5-2, 7-3, 2-4, 1-5, and 1-6. In one embodiment, the κ V_L comprise 3-7, 1-8, 1-9, 2-10, 3-11, 1-12, 1-13, 2-14, 3-15, 1-16. In one embodiment, the human Vκ segments comprise 1-17, 2-18, 2-19, 3-20, 6-21, 1-22, 1-23, 2-24, 3-25, 2-26, 1-27, 2-28, 2-29, and 2-30. In one embodiment, the human Vκ segments comprise 3-31, 1-32, 1-33, 3-34, 1-35, 2-36, 1-37, 2-38, 1-39, and 2-40.

[0012] In one embodiment, the V_L segments are human Vκ segments and comprise 4-1, 5-2, 7-3, 2-4, 1-5, 1-6, 3-7, 1-8, 1-9, 2-10, 3-11, 1-12, 1-13, 2-14, 3-15, and 1-16. In one embodiment, the Vκ segments further comprise 1-17, 2-18, 2-19, 3-20, 6-21, 1-22, 1-23, 2-24, 3-25, 2-26, 1-27, 2-28, 2-29, and 2-30. In one embodiment, the Vκ segments further comprise 3-31, 1-32, 1-33, 3-34, 1-35, 2-36, 1-37, 2-38, 1-39, and 2-40.

[0013] In one embodiment, the V_L segments are human Vλ segments and comprise a fragment of cluster A of the human λ light chain locus. In a specific embodiment, the fragment of cluster A of the human λ light chain locus extends from hVλ3-27 through hVλ3-1.

[0014] In one embodiment, the V_L segments comprise a fragment of cluster B of the human λ light chain locus. In a specific embodiment, the fragment of cluster B of the human λ light chain locus extends from hVλ5-52 through hVλ1-40.

[0015] In one embodiment, the V_L segments comprise a human λ light chain variable region sequence that comprises a genomic fragment of cluster A and a genomic fragment of cluster B. In a one embodiment, the human λ light chain variable region sequence comprises at least one gene segment of cluster A and at least one gene segment of cluster B.

[0016] In one embodiment, the V_L segments comprise at least one gene segment of cluster B and at least one gene segment of cluster C.

[0017] In one embodiment, the V_L segments comprise hVλ 3-1, 4-3, 2-8, 3-9, 3-10, 2-11, and 3-12. In a specific embodiment, the V_L segments comprise a contiguous sequence of the human λ light chain locus that spans from Vλ3-12 to Vλ3-1. In one embodiment, the contiguous sequence comprises at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 hVλs. In a specific embodiment, the hVλs include 3-1, 4-3, 2-8, 3-9, 3-10, 2-11, and 3-12. In a specific

embodiment, the hVλs comprises a contiguous sequence of the human λ locus that spans from Vλ3-12 to Vλ3-1.

[0018] In one embodiment, the hVλs comprises 13 to 28 or more hVλs. In a specific embodiment, the hVλs include 2-14, 3-16, 2-18, 3-19, 3-21, 3-22, 2-23, 3-25, and 3-27. In a specific embodiment, the hVλs comprise a contiguous sequence of the human λ locus that spans from Vλ3-27 to Vλ3-1.

[0019] In one embodiment, the V_L segments comprise 29 to 40 hVλs. In a specific embodiment, the V_L segments comprise a contiguous sequence of the human λ locus that spans from Vλ3-29 to Vλ3-1, and a contiguous sequence of the human λ locus that spans from Vλ5-52 to Vλ1-40. In a specific embodiment, all or substantially all sequence between hVλ1-40 and hVλ3-29 in the genetically modified mouse consists essentially of a human λ sequence of approximately 959 bp found in nature (*e.g.,* in the human population) downstream of the hVλ1-40 gene segment (downstream of the 3' untranslated portion), a restriction enzyme site (*e.g.,* PI-SceI), followed by a human λ sequence of approximately 3,431 bp upstream of the hVλ3-29 gene segment found in nature.

[0020] In one embodiment, the Jκ is human and is selected from the group consisting of Jκ1, Jκ2, Jκ3, Jκ4, Jκ5, and a combination thereof. In a specific embodiment, the Jκ comprises Jκ1 through Jκ5.

[0021] In one embodiment, the V_L segments are human Vλ segments, and the Jκ gene segment comprises an RSS having a 12-mer spacer, wherein the RSS is juxtaposed at the upstream end of the Jκ gene segment. In one embodiment, the V_L gene segments are human Vλ and the VL_H locus comprises two or more Jκ gene segments, each comprising an RSS having a 12-mer spacer wherein the RSS is juxtaposed at the upstream end of each Jκ gene segment.

[0022] In a specific embodiment, the V_L segments comprise contiguous human κ gene segments spanning the human κ locus from Vκ4-1 through Vκ2-40, and the J_L segments comprise contiguous gene segments spanning the human κ locus from Jκ1 through Jκ5.

[0023] In one embodiment, where the V_L segments are Vλ segments and no D_H segment is present between the V_L segments and J segments, the V_L segments are flanked downstream (*i.e.*, juxtaposed on the downstream side) with 23-mer RSS, and Jκ segments if present or Jλ segments if present are flanked upstream (*i.e.,* juxtaposed on the upstream side) with 12-mer RSS.

[0024] In one embodiment, where the V gene segments are Vκ gene segments and no D_H gene segment is present between the V gene segments and J gene segments, the Vκ gene segments are each juxtaposed on the downstream side with a 12-mer RSS, and Jκ

segments if present or Jλ segments if present are each juxtaposed on the upstream side with a 23-mer RSS.

[0025] In one embodiment, the mouse comprises a rearranged gene that is derived from a V_L gene segment, a J_L gene segment, and an endogenous mouse C_H gene. In one embodiment, the rearranged gene is somatically mutated. In one embodiment, the rearranged gene comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more N additions. In one embodiment, the N additions and/or the somatic mutations observed in the rearranged gene derived from the V_L segment and the J_L segment are 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 4.5-fold, or at least 5-fold more than the number of N additions and/or somatic mutations observed in a rearranged light chain variable domain (derived from the same V_L gene segment and the same J_L gene segment) that is rearranged at an endogenous light chain locus. In one embodiment, the rearranged gene is in a B cell that specifically binds an antigen of interest, wherein the B cell binds the antigen of interest with a K_D in the low nanomolar range or lower (*e.g.*, a K_D of 10 nanomolar or lower). In a specific embodiment, the V_L segment, the J_L segment, or both, are human gene segments. In a specific embodiment, the V_L and J_L segments are human κ gene segments. In one embodiment, the mouse C_H gene is selected from IgM, IgD, IgG, IgA and IgE. In a specific embodiment, the mouse IgG is selected from IgG1, IgG2A, IgG2B, IgG2C and IgG3. In another specific embodiment, the mouse IgG is IgG1.

[0026] In one embodiment, the mouse comprises a B cell, wherein the B cell makes from a locus on a chromosome of the B cell a binding protein consisting essentially of four polypeptide chains, wherein the four polypeptide chains consist essentially of (a) two identical polypeptides that comprise an endogenous mouse C_H region fused with a V_L; and, (b) two identical polypeptides that comprise an endogenous mouse C_L region fused with a V_L region that is cognate with respect to the V_L region that is fused with the mouse C_H region, and, in one embodiment, is a human *(e.g.*, a human κ) V_L region. In one embodiment, the V_L region fused to the endogenous mouse C_H region is a human V_L region. In a specific embodiment, the human V_L region fused with the mouse C_H region is a Vκ region. In a specific embodiment, the human V_L region fused with the mouse C_H region is identical to a V region encoded by a rearranged human germline light chain nucleotide sequence. In a specific embodiment, the human V_L region fused to the mouse C_H region comprises two, three, four, five, six, or more somatic hypermutations. In one embodiment, the human V_L region fused to the mouse C_H region is encoded by a rearranged gene that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0027] In one embodiment, at least 50% of all IgG molecules made by the mouse comprise a polypeptide that comprises an IgG isotype C_H region and a V_L region, wherein

the length of said polypeptide is no longer than 535, 530, 525, 520, or 515 amino acids. In one embodiment, at least 75% of all IgG molecules comprise the polypeptide recited in this paragraph. In one embodiment, at least 80%, 85%, 90%, or 95% of all IgG molecules comprise the polypeptide recited in this paragraph. In a specific embodiment, all IgG molecules made by the mouse comprise a polypeptide that is no longer than the length of the polypeptide recited in this paragraph.

[0028] In one embodiment, the mouse makes a binding protein comprising a first polypeptide that comprises an endogenous mouse C_H region fused with a variable domain encoded by a rearranged human V gene segment and a J gene segment but not a D_H gene segment, and a second polypeptide that comprises an endogenous mouse C_L region fused with a V domain encoded by a rearranged human V gene segment and a J gene segment but not a D_H gene segment, and the binding protein specifically binds an antigen with an affinity in the micromolar, nanomolar, or picomolar range. In one embodiment, the J segment is a human J segment (*e.g.*, a human κ gene segment). In one embodiment, the human V segment is a human Vκ segment. In one embodiment, the variable domain that is fused with the endogenous mouse C_H region comprises a greater number of somatic hypermutations than the variable region that is fused with the endogenous mouse C_L region; in a specific embodiment, the variable region fused with the endogenous mouse C_H region comprises about 1.5, 2-, 3-, 4-, or 5-fold or more somatic hypermutations than the V region fused to the endogenous mouse C_L region; in a specific embodiment, the V region fused with the mouse C_H region comprises at least 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 or more somatic hypermutations than the V region fused with the mouse C_L region. In one embodiment, the V region fused with the mouse C_H region is encoded by a rearranged gene that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0029] In one embodiment, the mouse expresses a binding protein comprising a first light chain variable domain (V_L1) fused with an immunoglobulin heavy chain constant region sequence and a second light chain variable domain (V_L2) fused with an immunoglobulin light chain constant region, wherein V_L1 comprises a number of somatic hypermutations that is about 1.5- to about 5-fold higher or more than the number of somatic hypermutations present in V_L2. In one embodiment, the number of somatic hypermutations in V_L1 is about 2- to about 4-fold higher than in V_L2. In one embodiment, the number of somatic hypermutations in V_L1 is about 2- to about 3-fold higher than in V_L2. In one embodiment, V_L1 is encoded by a sequence that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0030] In one aspect, a genetically modified mouse is provided that expresses an immunoglobulin that consists essentially of the following polypeptides: a first two identical

polypeptides that each consists essentially of a C_H region fused with a variable domain that is derived from gene segments that consist essentially of a V_L gene segment and a J_L gene segment, and a second two identical polypeptides that each consists essentially of a C_L region fused with a variable domain that is derived from gene segments that consist essentially of a V_L segment and a J_L segment.

[0031] In a specific embodiment, the two identical polypeptides that have the C_H region have a mouse C_H region.

[0032] In a specific embodiment, the two identical polypeptides that have the C_L region have a mouse C_L region.

[0033] In one embodiment, the variable domain fused with the C_L region is a variable domain that is cognate with the variable domain fused to the C_H region.

[0034] In one embodiment, the variable domain that is fused with the endogenous mouse C_H region comprises a greater number of somatic hypermutations than the variable domain that is fused with the endogenous mouse C_L region; in a specific embodiment, the variable domain fused with the endogenous mouse C_H region comprises about 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 4.5-fold, or 5-fold or more somatic hypermutations than the variable domain fused to the endogenous mouse C_L region. In one embodiment, the variable domain fused with the endogenous mouse C_L region is encoded by a gene that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more N additions.

[0035] In one embodiment, one or more of the V segments and the J segments are human gene segments. In a specific embodiment, both the V segments and the J segments are human κ gene segments. In another specific embodiment, both of the V segments and the J segments are human λ gene segments. In one embodiment, the V segments and the J segments are independently selected from human κ and human λ gene segments. In a specific embodiment, the V segments are Vκ segments and the J segments are Jλ segments. In another specific embodiment, the V segments are Vλ segments and the J segments are Jκ segments.

[0036] In one embodiment, one or more of the variable domains fused with the C_L region and the variable domains fused with the C_H region are human variable domains. In a specific embodiment, the human variable domains are human Vκ domains. In another specific embodiment, the human variable domains are Vλ domains. In one embodiment, the human domains are independently selected from human Vκ and human Vλ domains. In a specific embodiment, the human variable domain fused with the C_L region is a human Vλ domain and the human variable domain fused with the C_H region is a human Vκ domain. In another embodiment, the human variable domain fused with the C_L region is a human Vκ domain and the human variable domain fused with the C_H is a human Vλ domain.

[0037] In one embodiment, the V_L gene segment of the first two identical polypeptides is selected from a human Vλ segment and a human Vκ segment. In one embodiment, the V_L segment of the second two identical polypeptides is selected from a human Vλ segment and a human Vκ segment. In a specific embodiment, the V_L segment of the first two identical polypeptides is a human Vκ segment and the V_L segment of the second two identical polypeptides is selected from a human Vκ segment and a human Vλ segment. In a specific embodiment, the V_L segment of the first two identical polypeptides is a human Vλ segment and the V_L segment of the second two identical polypeptides is selected from a human Vλ segment and a human Vκ segment. In a specific embodiment, the human V_L segment of the first two identical polypeptides is a human Vκ segment, and the human V_L segment of the second two identical polypeptides is a human Vκ segment.

[0038] In one embodiment, the IgG of the mouse comprises a binding protein made in response to an antigen, wherein the binding protein comprises a polypeptide that consists essentially of a variable domain and a C_H region, wherein the variable domain is encoded by a nucleotide sequence that consists essentially of a rearranged V_L segment and a rearranged J segment, and wherein the binding protein specifically binds an epitope of the antigen with a K_D in the micromolar, nanomolar, or picomolar range.

[0039] In one aspect, a mouse is provided, wherein all or substantially all of the IgG made by the mouse in response to an antigen comprises a heavy chain that comprises a variable domain, wherein the variable domain is encoded by a rearranged gene derived from gene segments that consist essentially of a V gene segment and a J gene segment. In one embodiment, the rearranged gene comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0040] In one embodiment, the V segment is a V segment of a light chain. In one embodiment, the light chain is selected from a κ light chain and a λ light chain. In a specific embodiment, the light chain is a κ light chain. In a specific embodiment, the V segment is a human V segment. In a specific embodiment, the V segment is a human Vκ segment and the J segment is a human Jκ segment.

[0041] In one embodiment, the J segment is a J segment of a light chain. In one embodiment, the light chain is selected from a κ light chain and a λ light chain. In a specific embodiment, the light chain is a κ light chain. In a specific embodiment, the J segment is a human J segment. In another embodiment, the J segment is a J segment of a heavy chain (*i.e.*, a J_H). In a specific embodiment, the heavy chain is of mouse origin. In another specific embodiment, the heavy chain is of human origin.

[0042] In one embodiment, the variable domain of the heavy chain that is made from no more than a V segment and a J segment is a somatically mutated variable domain.

[0043] In one embodiment, the variable domain of the heavy chain that is made from no more than a V segment and a J segment is fused to a mouse C_H region.

[0044] In a specific embodiment, all or substantially all of the IgG made by the mouse in response to an antigen comprises a variable domain that is derived from no more than one human V segment and no more than one human J segment, and the variable domain is fused to a mouse IgG constant region, and the IgG further comprises a light chain that comprises a human V_L domain fused with a mouse C_L region. In a specific embodiment, the V_L domain fused with the mouse C_L region is derived from a human $V\kappa$ segment and a human $J\kappa$ segment. In a specific embodiment, the V_L domain fused with the mouse C_L region is derived from a human $V\lambda$ segment and a human $J\lambda$ segment.

[0045] In one aspect, a mouse is provided that makes an IgG comprising a first CDR3 on a polypeptide comprising a C_H region and a second CDR3 on a polypeptide comprising a C_L region, wherein both the first CDR3 and the second CDR3 are each independently derived from no more than two gene segments, wherein the two gene segments consist essentially of a V_L gene segment and a J_L gene segment. In one embodiment, the CDR3 on the polypeptide comprising the C_H region comprises a sequence that is derived from a CDR3 nucleotide sequence that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0046] In one embodiment, the V_L segment and the J_L segment are human gene segments. In one embodiment, the V_L segment and the J_L segment are κ gene segments. In one embodiment, the V_L segment and the J_L segment are λ gene segments.

[0047] In one aspect, a mouse is provided that makes an IgG comprising a first CDR3 on a first polypeptide comprising a C_H region and a second CDR3 on a second polypeptide comprising a C_L region, wherein both the first CDR3 and the second CDR3 each comprise a sequence of amino acids wherein more than 75% of the amino acids are derived from a V gene segment. In one embodiment, the CDR3 on the polypeptide comprising the C_H region comprises a sequence that is derived from a CDR3 nucleotide sequence that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0048] In one embodiment, more than 80%, more than 90%, or more than 95% of the amino acids of the first CDR3, and more than 80%, more than 90%, or more than 95% of the amino acids of the second CDR3, are derived from a light chain V segment.

[0049] In one embodiment, no more than two amino acids of the first CDR3 are derived from a gene segment other than a light chain V segment. In one embodiment, no more than two amino acids of the second CDR3 are derived from a gene segment other than a light chain V segment. In a specific embodiment, no more than two amino acids of the first CDR3 and no more than two amino acids of the second CDR3 are derived from a gene segment other than a light chain V segment. In one embodiment, no CDR3 of the IgG comprises an

amino acid sequence derived from a D gene segment. In one embodiment, the CDR3 of the first polypeptide does not comprise a sequence derived from a D segment.

[0050] In one embodiment, the V segment is a human V gene segment. In a specific embodiment, the V segment is a human $V\kappa$ gene segment.

[0051] In one embodiment, the first and/or the second CDR3 have at least one, two, three, four, five, or six somatic hypermutations. In one embodiment, the first CDR3 is encoded by a nucleic acid sequence that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0052] In one embodiment, the first CDR3 consists essentially of amino acids derived from a human light chain V gene segment and a human light chain J gene segment, and the second CDR3 consists essentially of amino acids derived from a human light chain V gene segment and a human light chain J gene segment. In one embodiment, the first CDR3 is derived from a nucleic acid sequence that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions. In one embodiment, the first CDR3 is derived from no more than two gene segments, wherein the no more than two gene segments are a human $V\kappa$ gene segment and a human $J\kappa$ gene segment; and the second CDR3 is derived from no more than two gene segments, wherein the no more than two gene segments are a human $V\kappa$ gene segment and a J gene segment selected from a human $J\kappa$ segment, a human $J\lambda$ segment, and a human J_H segment. In one embodiment, the first CDR3 is derived from no more than two gene segments, wherein the no more than two gene segments are a human $V\lambda$ segment and a J segment selected from a human $J\kappa$ segment, a human $J\lambda$ segment, and a human J_H segment.

[0053] In one aspect, a mouse is provided that makes an IgG that does not contain an amino acid sequence derived from a D_H gene segment, wherein the IgG comprises a first polypeptide having a first V_L domain fused with a mouse C_L region and a second polypeptide having a second V_L domain fused with a mouse C_H region, wherein the first V_L domain and the second V_L domain are not identical. In one embodiment, the first and second V_L domains are derived from different V segments. In another embodiment, the first and second V_L domains are derived from different J segments. In one embodiment, the first and second V_L domains are derived from identical V and J segments, wherein the second V_L domain comprises a higher number of somatic hypermutations as compared to the first V_L domain.

[0054] In one embodiment, the first and the second V_L domains are independently selected from human and mouse V_L domains. In one embodiment, the first and second V_L domains are independently selected from $V\kappa$ and $V\lambda$ domains. In a specific embodiment, the first V_L domain is selected from a $V\kappa$ domain and a $V\lambda$ domain, and the second V_L

domain is a $V\kappa$ domain. In another specific embodiment, the $V\kappa$ domain is a human $V\kappa$ domain.

[0055] In one aspect, a mouse is provided, wherein all or substantially all of the IgG made by the mouse consists essentially of a light chain having a first human V_L domain fused with a mouse C_L domain, and a heavy chain having a second human V_L domain fused with a mouse C_H domain.

[0056] In one embodiment, the human V_L domain fused with the mouse C_H domain is a human $V\kappa$ domain.

[0057] In one embodiment, the first and the second human V_L domains are not identical.

[0058] In one aspect, a mouse is provided, wherein at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or about 100% of the immunoglobulin G made by the mouse consists essentially of a dimer of (a) a first polypeptide that consists essentially of an immunoglobulin V_L domain and an immunoglobulin C_L region; and, (b) a second polypeptide of no more than 535 amino acids in length, wherein the second polypeptide consists essentially of a C_H region and a V domain that lacks a sequence derived from a D_H gene segment.

[0059] In one embodiment, the second polypeptide is about 435-535 amino acids in length. In a specific embodiment, the second polypeptide is about 435-530 amino acids in length. In a specific embodiment, the second polypeptide is about 435-525 amino acids in length. In a specific embodiment, the second polypeptide is about 435-520 amino acids in length. In a specific embodiment, the second polypeptide is about 435-515 amino acids in length.

[0060] In one embodiment, in about 90% or more of the IgG made by the mouse the second polypeptide is no more than about 535 amino acids in length.

[0061] In one embodiment, in about 50% or more of the IgG made by the mouse the second polypeptide is no more than about 535 amino acids in length. In one embodiment, in about 50% or more of the immunoglobulin G made by the mouse the second polypeptide is no more than about 530, 525, 520, 515, 510, 505, 500, 495, 490, 485, 480, 475, 470, 465, 460, 455, or 450 amino acids in length. In one embodiment, about 60%, 70%, 80%, 90%, or 95 % or more of the IgG made by the mouse is of the recited length. In a specific embodiment, all or substantially all of the IgG made by the mouse is of the recited length.

[0062] In one embodiment, the V domain of the second polypeptide is a V_L domain. In a specific embodiment, the V domain of the second polypeptide is selected from a $V\kappa$ and a $V\lambda$ domain. In a specific embodiment, the V domain of the second polypeptide is a human $V\kappa$ or $V\lambda$ domain.

[0063] In one aspect, a mouse is provided that expresses from a nucleotide sequence in its germline a polypeptide that comprises a light chain variable sequence (*e.g.*, a V and/or J sequence), a D_H sequence, and a heavy chain constant region.

[0064] In one embodiment, the mouse expresses the polypeptide from an endogenous mouse heavy chain locus that comprises a replacement of all or substantially all functional endogenous mouse heavy chain variable locus gene segments with a plurality of human gene segments at the endogenous mouse heavy chain locus.

[0065] In one embodiment, the polypeptide comprises a V_L sequence derived from a Vλ or a Vκ gene segment, the polypeptide comprises a CDR3 derived from a D_H gene segment, and the polypeptide comprises a sequence derived from a J_H or Jλ or Jκ gene segment.

[0066] In one embodiment, the mouse comprises an endogenous mouse heavy chain immunoglobulin locus comprising a replacement of all functional V_H gene segments with one or more human light chain Vλ gene segments wherein the one or more human Vλ segments each have juxtaposed on the downstream side a 23-mer spaced recombination signal sequence (RSS), wherein the Vλ segments are operably linked to a human or mouse D_H segment that has juxtaposed upstream and downstream a 12-mer spaced RSS; the D_H gene segment is operably linked with a J segment juxtaposed upstream with a 23-mer spaced RSS that is suitable for recombining with the 12-mer spaced RSS juxtaposing the D_H gene segment; wherein the V, D_H, and J segments are operably linked to a nucleic acid sequence encoding a heavy chain constant region.

[0067] In one embodiment, the mouse comprises an endogenous mouse heavy chain immunoglobulin locus comprising a replacement of all functional V_H gene segments with one or more human Vκ gene segments each juxtaposed on the downstream side with a 12-mer spaced recombination signal sequence (RSS), wherein the V segments are operably linked to a human or mouse D_H segment that is juxtaposed both upstream and downstream with a 23-mer spaced RSS; the D_H segment is operably linked with a J segment juxtaposed on the upstream side with a 12-mer spaced RSS that is suitable for recombining with the 23-mer spaced RSS juxtaposing the D_H segment; wherein the V, D_H, and gene segments are operably linked to a nucleic acid sequence encoding a heavy chain constant region.

[0068] In one embodiment, the heavy chain constant region is an endogenous mouse heavy chain constant region. In one embodiment, the nucleic acid sequence encodes a sequence selected from a C_H1, a hinge, a C_H2, a C_H3, and a combination thereof. In one embodiment, one or more of the C_H1, hinge, C_H2, and C_H3 are human.

[0069] In one embodiment, the mouse comprises an endogenous mouse heavy chain immunoglobulin locus comprising a replacement of all functional V_H gene segments with a plurality of human Vλ or Vκ gene segments each juxtaposed downstream with 23-mer

spaced RSS, a plurality of human D_H segments juxtaposed both upstream and downstream by a 12-mer spaced RSS, a plurality of human J segments (J_H or Jλ or Jκ) juxtaposed both upstream and downstream with a 23-mer spaced RSS, wherein the locus comprises an endogenous mouse constant region sequence selected from C_H1, hinge, C_H2, C_H3, and a combination thereof. In a specific embodiment, the mouse comprises all or substantially all functional human Vλ or Vκ segments, all or substantially all functional human D_H segments, and all or substantially all J_H or Jλ or Jκ segments.

[0070] In one embodiment, the mouse expresses an antigen-binding protein comprising (a) a polypeptide that comprises a human light chain sequence linked to a heavy chain constant sequence comprising a mouse sequence; and (b) a polypeptide that comprises a human light chain variable region linked to a human or mouse light chain constant sequence. In a specific embodiment, the light chain sequence is a human light chain sequence, and upon exposure to a protease that is capable of cleaving an antibody into an Fc and a Fab, a fully human Fab is formed that comprises at least four light chain CDRs, wherein the at least four light chain CDRs are selected from λ sequences, κ sequences, and a combination thereof. In one embodiment, the Fab comprises at least five light chain CDRs. In one embodiment, the Fab comprises six light chain CDRs. In one embodiment, at least one CDR of the Fab comprises a sequence derived from a Vλ segment or a Vκ segment, and the at least one CDR further comprises a sequence derived from a D segment. In one embodiment, the at least one CDR is a CDR3 and the CDR is derived from a human Vκ segment, a human D segment, and a human Jκ segment.

[0071] In one embodiment, the polypeptide of comprises a variable region derived from a human Vλ or Vκ gene segment, a human D_H gene segment, and a human J_H or Jλ or Jκ gene segment. In a specific embodiment, the heavy chain constant sequence is derived from a human C_H1 and a mouse C_H2 and a mouse C_H3 sequence.

[0072] In one aspect, a mouse is provided that comprises in its germline an unrearranged human Vκ or Vλ gene segment operably linked to a human J gene segment and a heavy chain constant region sequence, wherein the mouse expresses a V_L binding protein that comprises a human Vκ domain fused with a heavy chain constant region, and wherein the mice exhibit a population of splenic B cells that express V_L binding proteins in $CD19^+$ B cells, including transitional B cells ($CD19^+IgM^{hi}IgD^{int}$), and mature B cells ($CD19^+IgM^{int}IgD^{hi}$).

[0073] In one aspect, a mouse is provided that comprises in its germline an unrearranged human Vκ or Vλ gene segment operably linked to a human J gene segment and a heavy chain constant region sequence, wherein the mouse expresses on a B cell an immunoglobulin that comprises a light chain variable domain fused with a heavy chain

constant region, wherein the lymphocyte population in bone marrow of the mice exhibit a pro/pre B cell population that is about the same in number as in a pro/pre B cell population of a wild-type mouse (lymphocytes in bone marrow).

[0074] In one embodiment, the mice comprise at least 6 unrearranged hVκ gene segments and one or more unrearranged hJκ gene segments, and the mice comprise a lymphocyte-gated and IgM^+ spleen cell population expressing a V_L binding protein, wherein the population is at least 75% as large as a lymphocyte-gated and IgM^+ spleen cell population of a wild-type mouse.

[0075] In one embodiment, the mice exhibit a mature B cell-gated ($CD19^+$) splenocyte population of IgD^+ cells and IgM^+ cells that total about 90%; in one embodiment, the mature B cell-gated ($CD19^+$) splenocyte population of IgD^+ cells and IgM^+ cells of the modified mouse is about the same (*e.g.*, within 10%, or within 5%) as the total of IgD^+ cells and IgM^+ cells of a wild-type mouse that are mature B cell-gated ($CD19^+$) splenocytes.

[0076] In one aspect, a mouse is provided that expresses an immunoglobulin protein from a modified endogenous heavy chain locus in its germline, wherein the modified endogenous heavy chain locus lacks a functional mouse heavy chain V gene segment and the locus comprises unrearranged light chain V gene segments and unrearranged J gene segments, wherein the unrearranged light chain V gene segments and unrearranged J gene segments are operably linked with a heavy chain constant region sequence; wherein the immunoglobulin protein consists essentially of a first polypeptide and a second polypeptide, wherein the first polypeptide comprises an immunoglobulin light chain sequence and an immunoglobulin heavy chain constant sequence, and the second polypeptide comprises an immunoglobulin light chain variable domain and a light chain constant region.

[0077] In one aspect, a mouse is provided that expresses an immunoglobulin protein, wherein the immunoglobulin protein lacks a heavy chain immunoglobulin variable domain, and the immunoglobulin protein comprises a first variable domain derived from a light chain gene, and a second variable domain derived from a light chain gene, wherein the first variable domain and the second variable domain are cognate with respect to one another, wherein the first and the second light chain variable domains are not identical, and wherein the first and the second light chain variable domains associate and when associated specifically bind an antigen of interest.

[0078] In one aspect, a mouse is provided that makes from unrearranged gene segments in its germline an immunoglobulin protein comprising variable regions that are wholly derived from gene segments that consist essentially of unrearranged human gene segments, wherein the immunoglobulin protein comprises an immunoglobulin light chain

constant sequence and an immunoglobulin heavy chain constant sequence selected from the group consisting of a C_H1, a hinge, a C_H2, a C_H3, and a combination thereof.

[0079] In one aspect, a mouse is provided that makes from unrearranged gene segments in its germline an immunoglobulin protein comprising variable regions, wherein all CDR3s of all variable regions are generated entirely from light chain V and J gene segments, and optionally one or more somatic hypermutations, *e.g.*, one or more N additions.

[0080] In one aspect, a mouse is provided that makes a somatically mutated immunoglobulin protein derived from unrearranged human immunoglobulin light chain variable region gene segments in the germline of the mouse, wherein the immunoglobulin protein lacks a CDR that comprises a sequence derived from a D gene segment, wherein the immunoglobulin protein comprises a first CDR3 on a light chain variable domain fused with a light chain constant region, comprises a second CDR3 on a light chain variable domain fused with a heavy chain constant region, and wherein the second CDR3 is derived from a rearranged light chain variable region sequence that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[0081] In one aspect, a mouse as described herein is provided, wherein the mouse comprises a functionally silenced light chain locus selected from a λ locus, a κ locus, and a combination thereof. In one embodiment, the mouse comprises a deletion of a λ and/or a κ locus, in whole or in part, such that the λ and/or κ locus is nonfunctional.

[0082] In one aspect, a mouse embryo is provided, comprising a cell that comprises a modified immunoglobulin locus as described herein. In one embodiment, the mouse is a chimera and at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95% of the cells of the embryo comprise a modified immunoglobulin locus as described herein. In one embodiment, at least 96%. 97%, 98%, 99%, or 99.8% of the cells of the embryo comprise a modified immunoglobulin locus as described herein. In one embodiment, the embryo comprises a host cell and a cell derived from a donor ES cell, wherein the cell derived from the donor ES cell comprises a modified immunoglobulin locus as described herein. In one embodiment, the embryo is a 2-, 4-, 8, 16-, 32, or 64-cell stage host embryo, or a blastocyst, and further comprises a donor ES cell comprising a modified immunoglobulin locus as described herein.

[0083] In one aspect, a mouse or a cell made using a nucleic acid construct as described herein is provided.

[0084] In one aspect, a mouse made using a cell as described herein is provided. In one embodiment, the cell is a mouse ES cell.

[0085] In one aspect, use of a mouse as described herein to make a nucleic acid sequence encoding a first human light chain immunoglobulin variable sequence (V_L1) that is cognate with a second human light chain immunoglobulin variable sequence (V_L2), wherein the V_L1 fused with a human immunoglobulin light chain constant region (polypeptide 1) expresses with V_L2 fused with a human immunoglobulin heavy chain constant region (polypeptide 2), as a dimer of polypeptide1/polypeptide 2, to form a V_L1-V_L2 antibody.

[0086] In one aspect, use of a mouse as described herein to make a nucleic acid sequence encoding a human immunoglobulin light chain variable sequence that is fused with a human immunoglobulin heavy chain sequence, wherein the nucleic acid sequence encodes a human V_L-C_H polypeptide, wherein the human V_L-C_H polypeptide expresses as a dimer, and wherein the dimer expresses in the absence of an immunoglobulin light chain (*e.g.*, in the absence of a human λ or human κ light chain). In one embodiment, the V_L-C_H dimer specifically binds an antigen of interest in the absence of a λ light chain and in the absence of a κ light chain.

[0087] In one aspect, use of a mouse as described herein to make a nucleic acid sequence encoding all or a portion of an immunoglobulin variable domain. In one embodiment, the immunoglobulin variable domain is a human Vλ or human Vκ domain.

[0088] In one aspect, use of a mouse as described herein to make a fully human Fab (comprising a first human V_L fused with a human light chain constant region, and a second human V_L fused with a human heavy chain constant region sequence) or a fully human $F(ab)_2$ is provided.

[0089] In one aspect, use of a mouse as described herein to make an immortalized cell line is provided. In one embodiment, the immortalized cell line comprises a nucleic acid sequence encoding a human Vλ or Vκ domain operably linked to a nucleic acid sequence that comprises a mouse constant region nucleic acid sequence.

[0090] In one aspect, use of a mouse as described herein to make a hybridoma or quadroma is provided.

[0091] In one aspect, a cell is provided, comprising a modified immunoglobulin locus as described herein. In one embodiment, the cell is selected from a totipotent cell, a pluripotent cell, an induced pluripotent stem cell (iPS), and an ES cell. In a specific embodiment, the cell is a mouse cell, *e.g.*, a mouse ES cell. In one embodiment, the cell is homozygous for the modified immunoglobulin locus.

[0092] In one aspect, a cell is provided, comprising a nucleic acid sequence encoding a first polypeptide that comprises a first somatically mutated human Vκ or Vλ sequence fused to a human heavy chain constant region sequence.

[0093] In one embodiment, the cell further comprises a second polypeptide chain that comprises a second somatically mutated human Vκ or Vλ sequence fused to a human light chain constant region sequence.

[0094] In one embodiment, the human Vκ or Vλ sequence of the first polypeptide is cognate with the human Vκ or Vλ sequence of the second polypeptide.

[0095] In one embodiment, the Vκ or Vλ of the first polypeptide and the human Vκ or Vλ of the second polypeptide when associated specifically bind an antigen of interest. In a specific embodiment, the first polypeptide comprises a variable domain consisting essentially of a human Vκ, and the second polypeptide comprises a variable domain consisting of a human Vκ that is cognate with the human Vκ of the first polypeptide, and the human constant region sequence is an IgG sequence.

[0096] In one embodiment, the cell is selected from a CHO cell, a COS cell, a 293 cell, a HeLa cell, and a human retinal cell expressing a viral nucleic acid sequence (*e.g.*, a PERC.6™ cell.

[0097] In one aspect, a somatic mouse cell is provided, comprising a chromosome that comprises a genetic modification as described herein.

[0098] In one aspect, a mouse germ cell is provided, comprising a nucleic acid sequence that comprises a genetic modification as described herein.

[0099] In one aspect, a pluripotent, induced pluripotent, or totipotent cell derived from a mouse as described herein is provided. In a specific embodiment, the cell is a mouse embryonic stem (ES) cell.

[00100] In one aspect, use of a cell as described herein for the manufacture of a mouse, a cell, or a therapeutic protein (*e.g.*, an antibody or other antigen-binding protein) is provided.

[00101] In one aspect, a nucleic acid construct is provided that comprises a human D_H gene segment juxtaposed upstream and downstream with a 23-mer spaced RSS. In a specific embodiment, the nucleic acid construct comprises a homology arm that is homologous to a human genomic sequence comprising human Vκ gene segments. In one embodiment, the targeting construct comprises all or substantially all human D_H gene segments each juxtaposed upstream and downstream with a 23-mer spaced RSS.

[00102] In one aspect, a nucleic acid construct is provided that comprises a human Jκ gene segment juxtaposed upstream with a 12-mer spaced RSS. In a specific embodiment, the nucleic acid construct comprises a first homology arm that contains homology to a human genomic D_H gene sequence that is juxtaposed upstream and downstream with a 23-mer spaced RSS. In one embodiment, the nucleic acid construct comprises a second homology arm that contains homology to a human genomic J gene sequence or that

contains homology to a mouse heavy chain constant region sequence or that contains homology to a J-C intergenic sequence upstream of a mouse constant region heavy chain sequence.

[00103] In one aspect, a nucleic acid construct is provided that comprises a human Vλ segment juxtaposed downstream with a 23-mer spaced RSS, a human D_H segment juxtaposed upstream and downstream with a 12-mer spaced RSS, and a human J segment selected from a Jκ segment juxtaposed upstream with a 23-mer spaced RSS, a human Jλ segment juxtaposed upstream with a 23-mer spaced RSS, and a human J_H segment juxtaposed upstream with a 23-mer spaced RSS. In one embodiment, the construct comprises a homology arm that contains homology to a mouse constant region sequence, a J-C intergenic mouse sequence, and/or a human Vλ sequence.

[00104] In one embodiment, the nucleic acid construct comprises a human λ light chain variable region sequence that comprises a fragment of cluster A of the human λ light chain locus. In a specific embodiment, the fragment of cluster A of the human λ light chain locus extends from hVλ3-27 through hVλ3-1.

[00105] In one embodiment, the nucleic acid construct comprises a human λ light chain variable region sequence that comprises a fragment of cluster B of the human λ light chain locus. In a specific embodiment, the fragment of cluster B of the human λ light chain locus extends from hVλ5-52 through hVλ1-40.

[00106] In one embodiment, nucleic acid construct comprises a human λ light chain variable region sequence that comprises a genomic fragment of cluster A and a genomic fragment of cluster B. In a one embodiment, the human λ light chain variable region sequence comprises at least one gene segment of cluster A and at least one gene segment of cluster B.

[00107] In one embodiment, the human λ light chain variable region sequence comprises at least one gene segment of cluster B and at least one gene segment of cluster C.

[00108] In one aspect, a nucleic acid construct is provided, comprising a human D_H segment juxtaposed upstream and downstream with a 23-mer spaced RSS normally found in nature flanking either a Jκ, a J_H, a Vλ, or a V_H segment. In one embodiment, the nucleic acid construct comprises a first homology arm homologous to a human V-J intergenic region or homologous to a human genomic sequence comprising a human V gene segment. In one embodiment, the nucleic acid construct comprises a second homology arm homologous to a human or mouse heavy chain constant region sequence. In a specific embodiment, the human or mouse heavy chain constant region sequence is selected from a C_H1, hinge, C_H2, C_H3, and a combination thereof. In one embodiment, the nucleic acid construct comprises a human J gene segment flanked upstream with a 12-mer RSS. In one embodiment, the

nucleic acid construct comprises a second homology arm that contains homology to a J gene segment flanked upstream with a 12-mer RSS. In one embodiment, the J gene segment is selected from a human Jκ, a human Jλ, and a human J_H.

[00109] In one aspect, a nucleic acid construct is provided that comprises a human D_H segment juxtaposed upstream and downstream with a 23-mer spaced RSS, and a site-specific recombinase recognition sequence, *e.g.*, a sequence recognized by a site-specific recombinase such as a Cre, a Flp, or a Dre protein.

[00110] In one aspect, a nucleic acid construct is provided that comprises a human Vλ or a human Vκ segment, a D_H segment juxtaposed upstream and downstream with a 12-mer or a 23-mer spaced RSS, and a human J segment with a 12-mer or a 23-mer spaced RSS, wherein the 12-mer or 23-mer spaced RSS is positioned immediately 5' to the human J segment (*i.e.,* with respect to the direction of transcription). In one embodiment, the construct comprises a human Vλ juxtaposed with a 3' 23-mer spaced RSS, a human D_H segment juxtaposed upstream and downstream with a 12-mer spaced RSS, and a human Jκ segment juxtaposed with a 5' 23-mer spaced RSS. In one embodiment, the construct comprises a human Vκ juxtaposed with a 3' 12-mer spaced RSS, a human D_H segment juxtaposed upstream and downstream with a 23-mer spaced RSS, and a human Jλ segment juxtaposed with a 5' 12-mer spaced RSS.

[00111] In one aspect, a targeting vector is provided, comprising (a) a first targeting arm and a second targeting arm, wherein the first and second targeting arms are independently selected from human and mouse targeting arms, wherein the targeting arms direct the vector to an endogenous or modified immunoglobulin V region gene locus; and, (b) a contiguous sequence of human V_L gene segments or a contiguous sequence of human V_L gene segments and at least one human Jκ gene segment, wherein the contiguous sequence is selected from the group consisting of (i) hVκ4-1 through hVκ1-6 and Jκ1, (ii) hVκ4-1 through hVκ1-6 and Jκ1 through Jκ2, (iii) hVκ4-1 through hVκ1-6 and Jκ1 through Jκ3, (iv) hVκ4-1 through hVκ1-6 and Jκ1 through Jκ4, (v) hVκ4-1 through hVκ1-6 and Jκ1 through Jκ5, (vi) hVκ3-7 through hVκ1-16, (vii) hVκ1-17 through hVκ2-30, (viii) hVκ3-31 through hVκ2-40, and (ix) a combination thereof.

[00112] In one embodiment, the targeting arms that direct the vector to an endogenous or modified immunoglobulin locus are identical or substantially identical to a sequence at the endogenous or modified immunoglobulin locus.

[00113] In one aspect, use of a nucleic acid construct as described herein for the manufacture of a mouse, a cell, or a therapeutic protein (*e.g.*, an antibody or other antigen-binding protein) is provided.

[00114] In one aspect, use of a nucleic acid sequence from a mouse as described herein to make a cell line for the manufacture of a human therapeutic is provided. In one embodiment, the human therapeutic is a binding protein comprising a human light chain variable sequence (*e.g.*, derived from a human Vλ or human Vκ segment) fused with a human heavy chain constant sequence. In one embodiment, the human therapeutic comprises a first polypeptide that is a human λ or κ immunoglobulin light chain, and a second polypeptide that comprises a human Vλ or human Vκ variable sequence fused with a human heavy chain constant sequence.

[00115] In one aspect, an expression system is provided, comprising a mammalian cell transfected with a DNA construct that encodes a polypeptide that comprises a somatically mutated human V_L domain fused with a human C_H domain.

[00116] In one embodiment, the expression system further comprises a nucleotide sequence that encodes an immunoglobulin V_L domain fused with a human C_L domain, wherein the V_L domain fused with the human C_L domain is a cognate light chain with the V_L domain fused with the human C_H domain.

[00117] In one embodiment, the mammalian cell is selected from a CHO cell, a COS cell, a Vero cell, a 293 cell, and a retinal cell that expresses a viral gene (*e.g.*, a PER.C6™ cell).

[00118] In one aspect, a method for making a binding protein is provided, comprising obtaining a nucleotide sequence encoding a V_L domain from a gene encoding a V_L region fused to a C_H region from a cell of a mouse as described herein, and cloning the nucleotide sequence encoding the V_L region sequence in frame with a gene encoding a human C_H region to form a human binding protein sequence, expressing the human binding protein sequence in a suitable cell.

[00119] In one embodiment, the mouse has been immunized with an antigen of interest, and the V_L region fused to the C_H region specifically binds (*e.g.*, with a K_D in the micromolar, nanomolar, or picomolar range) an epitope of the antigen of interest. In one embodiment, nucleotide sequence encoding the V_L region fused to the C_H region is somatically mutated in the mouse.

[00120] In one embodiment, the suitable cell is selected from a B cell, a hybridoma, a quadroma, a CHO cell, a COS cell, a 293 cell, a HeLa cell, and a human retinal cell expressing a viral nucleic acid sequence (*e.g.,* a PERC.6™ cell).

[00121] In one embodiment, the C_H region comprises a human IgG isotype. In a specific embodiment, the human IgG is selected from an IgG1, IgG2, and IgG4. In another specific embodiment, the human IgG is IgG1. In another specific embodiment, the human IgG is IgG4. In another specific embodiment, the human IgG4 is a modified IgG4. In one embodiment, the modified IgG4 comprises a substitution in the hinge region. In a specific

embodiment, the modified IgG4 comprises a substitution at amino acid residue 228 relative to a wild-type human IgG4, numbered according to the EU numbering index of Kabat. In a specific embodiment, the substitution at amino acid residue 228 is a S228P substitution, numbered according to the EU numbering index of Kabat.

[00122] In one embodiment, the cell further comprises a nucleotide sequence encoding a V_L domain from a light chain that is cognate to the V_L domain fused to the C_H region, and the method further comprises expressing the nucleotide sequence encoding the cognate V_L domain fused to a human Cκ or Cλ domain.

[00123] In one aspect, a method for making a genetically modified mouse is provided, comprising replacing at an endogenous mouse heavy chain locus one or more immunoglobulin heavy chain gene segments of a mouse with one or more human immunoglobulin light chain gene segments. In one embodiment, the replacement is of all or substantially all functional mouse immunoglobulin heavy chain segments (*i.e.*, V_H, D_H, and J_H segments) with one or more functional human light chain segments (*i.e.*, V_L and J_L segments). In one embodiment, the replacement is of all or substantially all functional mouse heavy chain V_H, D_H, and J_H segments with all or substantially all human Vλ or Vκ segments and at least one Jλ or Jκ segment. In a specific embodiment, the replacement includes all or substantially all functional human Jλ or Jκ segments.

[00124] In one aspect, a method is provided for making a mouse that expresses a polypeptide that comprises a sequence derived from a human immunoglobulin Vλ or Vκ and/or Jλ or Jκ segment fused with a mouse heavy chain constant region, comprising replacing endogenous mouse heavy chain immunoglobulin variable segments (V_H, D_H, and J_H) with at least one human Vλ or Vκ segment and at least one human Jλ or Jκ segment, wherein the replacement is in a pluripotent, induced pluripotent, or totipotent mouse cell to form a genetically modified mouse progenitor cell; the genetically modified mouse progenitor cell is introduced into a mouse host; and, the mouse host comprising the genetically modified progenitor cell is gestated to form a mouse comprising a genome derived from the genetically modified mouse progenitor cell. In one embodiment, the host is an embryo. In a specific embodiment, the host is selected from a mouse pre-morula (*e.g.,* 8- or 4-cell stage), a tetraploid embryo, an aggregate of embryonic cells, or a blastocyst.

[00125] In one aspect, a method is provided for making a genetically modified mouse as described herein, comprising introducing by nuclear transfer a nucleic acid containing a modification as described herein into a cell, and maintaining the cell under suitable conditions (*e.g.*, including culturing the cell and gestating an embryo comprising the cell in a surrogate mother) to develop into a mouse as described herein.

[00126] In one aspect, a method for making a modified mouse is provided, comprising modifying as described herein a mouse ES cell or pluripotent or totipotent or induced pluripotent mouse cell to include one or more unrearranged immunoglobulin light chain variable gene segments operably linked to an immunoglobulin heavy chain constant sequence, culturing the ES cell, introducing the cultured ES cell into a host embryo to form a chimeric embryo, and introducing the chimeric embryo into a suitable host mouse to develop into a modified mouse. In one embodiment, the one or more unrearranged immunoglobulin light chain variable region gene segments are human λ or human κ gene segments. In one embodiment, the one or more unrearranged immunoglobulin light chain variable region gene segments comprise human Vλ or human Vκ segments and one or more Jλ, Jκ, or J_H segments. In one embodiment, the heavy chain constant gene sequence is a human sequence selected from C_H1, hinge, C_H2, C_H3, and a combination thereof. In one embodiment, the one or more unrearranged immunoglobulin light chain variable gene segments replace all or substantially all functional endogenous mouse heavy chain variable region gene segments at the endogenous mouse heavy chain locus, and the heavy chain constant sequence is a mouse sequence comprising a C_H1, a hinge, a C_H2, and a C_H3.

[00127] In one aspect, an immunoglobulin variable region (VR) (*e.g.*, comprising a human V_L sequence fused with a human J_L, or J_H, or D_H and J_H, or D_H and J_L) made in a mouse as described herein is provided. In a specific embodiment, the immunoglobulin VR is derived from a germline human gene segment selected from a Vκ segment and a Vλ segment, wherein the VR is encoded by a rearranged sequence from the mouse wherein the rearranged sequence is somatically hypermutated. In one embodiment, the rearranged sequence comprises 1 to 5 somatic hypermutations. In one embodiment, the rearranged sequence comprises at least 6, 7, 8, 9, or 10 somatic hypermutations. In one embodiment, the rearranged sequence comprises more than 10 somatic hypermutations. In one embodiment, the rearranged sequence is fused with one or more human or mouse heavy chain constant region sequences (*e.g.*, selected from a human or mouse C_H1, hinge, C_H2, C_H3, and a combination thereof).

[00128] In one aspect, an immunoglobulin variable domain amino acid sequence of a binding protein made in a mouse as described herein is provided. In one embodiment, the VR is fused with one or more human or mouse heavy chain constant region sequences (*e.g.*, selected from a human or mouse C_H1, hinge, C_H2, C_H3, and a combination thereof).

[00129] In one aspect, a light chain variable domain encoded by a nucleic acid sequence derived from a mouse as described herein is provided.

[00130] In one aspect, an antibody or antigen-binding fragment thereof (*e.g.*, Fab, $F(ab)_2$, scFv) made in a mouse as described herein, or derived from a sequence made in a mouse as described herein, is provided.

BRIEF DESCRIPTION OF THE FIGURES

[00131] **FIG. 1A** illustrates a schematic (not to scale) of the mouse heavy chain locus. The mouse heavy chain locus is about 3 Mb in length and contains approximately 200 heavy chain variable (V_H) gene segments, 13 heavy chain diversity (D_H) gene segments and 4 heavy chain joining (J_H) gene segments as well as enhancers (Enh) and heavy chain constant (C_H) regions.

[00132] **FIG. 1B** illustrates a schematic (not to scale) of the human κ light chain locus. The human κ light chain locus is duplicated into distal and proximal contigs of opposite polarity spanning about 440 kb and 600 kb, respectively. Between the two contigs is about 800 kb of DNA that is believed to be free of Vκ gene segments. The human κ light chain locus contains about 76 Vκ gene segments, 5 Jκ gene segments, an intronic enhancer (Enh) and a single constant region (Cκ).

[00133] **FIG. 2** shows a targeting strategy for progressive insertion of 40 human Vκ and 5 human Jκ gene segments into the mouse heavy chain locus. Hygromycin (HYG) and Neomycin (NEO) selection cassettes are shown with recombinase recognition sites (R1, R2, etc.).

[00134] **FIG. 3** shows a modified mouse heavy chain locus comprising human Vκ and Jκ gene segments operably linked to mouse C_H regions.

[00135] **FIG. 4A** shows an exemplary targeting strategy for progressive insertion of human Vλ and a single human Jλ gene segment into the mouse heavy chain locus. Hygromycin (HYG) and Neomycin (NEO) selection cassettes are shown with recombinase recognition sites (R1, R2, etc.).

[00136] **FIG. 4B** shows an exemplary targeting strategy for progressive insertion of human Vλ and four human Jλ gene segments into the mouse heavy chain locus. Hygromycin (HYG) and Neomycin (NEO) selection cassettes are shown with recombinase recognition sites (R1, R2, etc.).

[00137] **FIG. 5A** shows an exemplary targeting strategy for progressive insertion of human Vλ, human D_H and human J_H gene segments into the mouse heavy chain locus. Hygromycin (HYG) and Neomycin (NEO) selection cassettes are shown with recombinase recognition sites (R1, R2, etc.).

[00138] **FIG. 5B** shows an exemplary targeting strategy for progressive insertion of human Vλ, human D_H and human Jκ gene segments into the mouse heavy chain locus.

Hygromycin (HYG) and Neomycin (NEO) selection cassettes are shown with recombinase recognition sites (R1, R2, etc.).

[00139] **FIG. 6A** shows contour plots of splenocytes stained for surface expression of B220 and IgM from a representative wild type (WT) and a representative mouse homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO).

[00140] **FIG. 6B** shows contour plots of splenocytes gated on $CD19^+$ B cells and stained for immunoglobulin D (IgD) and immunoglobulin M (IgM) from a representative wild type (WT) and a representative mouse homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO).

[00141] **FIG. 6C** shows the total number of $CD19^+$ B cells, transitional B cells ($CD19^+IgM^{hi}IgD^{int}$) and mature B cells ($CD19^+IgM^{int}IgD^{hi}$) in harvested spleens from wild type (WT) and mice homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO).

[00142] **FIG. 7A** shows contour plots of bone marrow gated on singlets stained for immunoglobulin M (IgM) and B220 from a wild type mouse (WT) and a mouse homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO). Immature, mature and pro/pre B cells are noted on each of the dot plots.

[00143] **FIG. 7B** shows the total number of pre/pro ($B220^+IgM^-$), immature ($B220^{int}IgM^+$) and mature ($B220^{hi}IgM^+$) B cells in bone marrow isolated from the femurs of wild type mice (WT) and mice homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO).

[00144] **FIG. 7C** shows contour plots of bone marrow gated on $CD19^+$ and stained for $ckit^+$ and $CD43^+$ from a wild type mouse (WT) and a mouse homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO). Pro and pre B cells are noted on each of the dot plots.

[00145] **FIG. 7D** shows the number of pro B ($CD19^+CD43^+ckit^+$) and pre B ($CD19^+CD43^-ckit^-$) cells in bone marrow harvested from the femurs of wild type mice (WT) and mice homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO).

[00146] **FIG. 7E** shows contour plots of bone marrow gated on singlets stained for CD19 and CD43 from a wild type mouse (WT) and a mouse homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO). Immature, pre and pro B cells are noted on each of the dot plots.

[00147] **FIG. 7F** shows histograms of bone marrow gated on pre B cells ($CD19^{+}CD43^{int}$) and expressing immunoglobulin M (IgM) from a wild type mouse (WT) and a mouse homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVk-5hJk HO).

[00148] **FIG. 7G** shows the number of IgM^{+} pre B cells ($CD19^{+}IgM^{+}CD43^{int}$) and immature B cells ($CD19^{+}IgM^{+}CD43^{-}$) in bone marrow harvest from the femurs of wild type (WT) and mice homozygous for six human Vκ and five human Jκ gene segments positioned at the endogenous heavy chain locus (6hVκ-5hJκ HO).

[00149] **FIG.8A** shows contour plots of splenocytes gated on $CD19^{+}$ and stained for $Ig\lambda^{+}$ and $Ig\kappa^{+}$ expression from a mouse containing a wild type heavy chain locus and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segments (WT) and a mouse homozygous for thirty hVκ and five Jκ gene segments at the endogenous heavy chain locus and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segments (30hVκ-5hJκ HO).

[00150] **FIG. 8B** shows contour plots of bone marrow gated on immature ($B220^{int}IgM^{+}$) and mature ($B220^{hi}IgM^{+}$) B cells stained for Igλ and Igκ expression isolated from the femurs of a mouse containing a wild type heavy chain locus and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segments (WT) and a mouse homozygous for thirty hVκ and five Jκ gene segments at the endogenous heavy chain locus and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segments (30hVκ-5hJκ HO).

[00151] **FIG. 9** shows a nucleotide sequence alignment of the Vκ-Jκ-mIgG junction of twelve independent RT-PCR clones amplified from splenocyte RNA of naïve mice homozygous for thirty hVκ and five Jκ gene segments at the mouse heavy chain locus and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segment. Lower case bases indicate non-germline bases resulting from either mutation and/or N addition during recombination. Artificial spaces (periods) are included to properly align the Framework 4 region and show alignment of the mouse heavy chain IgG nucleotide sequence for IgG1, IgG2a/c, and IgG3 primed clones.

DETAILED DESCRIPTION

[00152] The phrase "bispecific binding protein" includes a binding protein capable of selectively binding two or more epitopes. Bispecific binding proteins comprise two different polypeptides that comprise a first light chain variable domain (V_L1) fused with a first C_H region and a second light chain variable domain (V_L2) fused with a second C_H region. In general, the first and the second C_H regions are identical, or they differ by one or more

amino acid substitutions (*e.g.*, as described herein). V_L1 and V_L2 specifically binding different epitopes—either on two different molecules (*e.g.,* antigens) or on the same molecule (*e.g.,* on the same antigen). If a bispecific binding protein selectively binds two different epitopes (a first epitope and a second epitope), the affinity of V_L1 for the first epitope will generally be at least one to two or three or four orders of magnitude lower than the affinity of V_L1 for the second epitope, and vice versa with respect to V_L2. The epitopes recognized by the bispecific binding protein can be on the same or a different target (*e.g.,* on the same or a different antigen). Bispecific binding proteins can be made, for example, by combining a V_L1 and a V_L2 that recognize different epitopes of the same antigen. For example, nucleic acid sequences encoding V_L sequences that recognize different epitopes of the same antigen can be fused to nucleic acid sequences encoding different C_H regions, and such sequences can be expressed in a cell that expresses an immunoglobulin light chain, or can be expressed in a cell that does not express an immunoglobulin light chain. A typical bispecific binding protein has two heavy chains each having three light chain CDRs, followed by (N-terminal to C-terminal) a C_H1 domain, a hinge, a C_H2 domain, and a C_H3 domain, and an immunoglobulin light chain that either does not confer antigen-binding specificity but that can associate with each heavy chain, or that can associate with each heavy chain and that can bind one or more of the epitopes bound by V_L1 and/or V_L2, or that can associate with each heavy chain and enable binding or assist in binding of one or both of the heavy chains to one or both epitopes.

[00153] Therefore, two general types of bispecific binding proteins are (1) V_L1-C_H(dimer), and (2) V_L1-C_H:light chain + V_L2-C_H:light chain, wherein the light chain is the same or different. In either case, the C_H (*i.e.*, the heavy chain constant region) can be differentially modified (*e.g.*, to differentially bind protein A, to increase serum half-life, *etc.*) as described herein, or can be the same.

[00154] The term "cell," when used in connection with expressing a sequence, includes any cell that is suitable for expressing a recombinant nucleic acid sequence. Cells include those of prokaryotes and eukaryotes (single-cell or multiple-cell), bacterial cells (*e.g.,* strains of *E. coli, Bacillus spp., Streptomyces spp., etc.),* mycobacteria cells, fungal cells, yeast cells (*e.g., S. cerevisiae, S. pombe, P. pastoris, P. methanolica, etc.*), plant cells, insect cells (*e.g., SF-9, SF-21, baculovirus-infected insect cells, Trichoplusia ni, etc.*), non-human animal cells, human cells, B cells, or cell fusions such as, for example, hybridomas or quadromas. In some embodiments, the cell is a human, monkey, ape, hamster, rat, or mouse cell. In some embodiments, the cell is eukaryotic and is selected from the following cells: CHO (*e.g.,* CHO K1, DXB-11 CHO, Veggie-CHO), COS (*e.g.,* COS-7), retinal cell, Vero, CV1, kidney (*e.g.,* HEK293, 293 EBNA, MSR 293, MDCK, HaK, BHK), HeLa, HepG2, WI38, MRC 5, Colo205, HB 8065, HL-60, (*e.g.,* BHK21), Jurkat, Daudi, A431 (epidermal), CV-1, U937,

3T3, L cell, C127 cell, SP2/0, NS-0, MMT 060562, Sertoli cell, BRL 3A cell, HT1080 cell, myeloma cell, tumor cell, and a cell line derived from an aforementioned cell. In some embodiments, the cell comprises one or more viral genes, *e.g.* a retinal cell that expresses a viral gene (*e.g.,* a PER.C6™ cell).

[00155] The term "cognate," when used in the sense of "cognate with," *e.g.*, a first V_L domain that is "cognate with" a second V_L domain, is intended to include reference to the relation between two V_L domains from a same binding protein made by a mouse in accordance with the invention. For example, a mouse that is genetically modified in accordance with an embodiment of the invention, *e.g.*, a mouse having a heavy chain locus in which V_H, D_H, and J_H regions are replaced with V_L and J_L regions, makes antibody-like binding proteins that have two identical polypeptide chains made of the same mouse C_H region (*e.g.*, an IgG isotype) fused with a first human V_L domain, and two identical polypeptide chains made of the same mouse C_L region fused with a second human V_L domain. During clonal selection in the mouse, the first and the second human V_L domains were selected by the clonal selection process to appear together in the context of a single antibody-like binding protein. Thus, first and second V_L domains that appear together, as the result of the clonal selection process, in a single antibody-like molecule are referred to as being "cognate." In contrast, a V_L domain that appears in a first antibody-like molecule and a V_L domain that appears in a second antibody-like molecule are *not* cognate, *unless* the first and the second antibody-like molecules have identical heavy chains (*i.e.*, unless the V_L domain fused to the first human heavy chain region and the V_L domain fused to the second human heavy chain region are identical).

[00156] The phrase "complementarity determining region," or the term "CDR," includes an amino acid sequence encoded by a nucleic acid sequence of an organism's immunoglobulin genes that normally (*i.e.,* in a wild-type animal) appears between two framework regions in a variable region of a light or a heavy chain of an immunoglobulin molecule (*e.g.,* an antibody or a T cell receptor). A CDR can be encoded by, for example, a germline sequence or a rearranged or unrearranged sequence, and, for example, by a naïve or a mature B cell or a T cell. In some circumstances (*e.g.,* for a CDR3), CDRs can be encoded by two or more sequences (*e.g.,* germline sequences) that are not contiguous (*e.g.,* in an unrearranged nucleic acid sequence) but are contiguous in a B cell nucleic acid sequence, *e.g.,* as the result of splicing or connecting the sequences (*e.g.,* V-D-J recombination to form a heavy chain CDR3).

[00157] The phrase "gene segment," or "segment" includes reference to a V (light or heavy) or D or J (light or heavy) immunoglobulin gene segment, which includes unrearranged sequences at immunoglobulin loci (in *e.g.*, humans and mice) that can participate in a rearrangement (mediated by, *e.g.,* endogenous recombinases) to form a

rearranged V/J or V/D/J sequence. Unless indicated otherwise, the V, D, and J segments comprise recombination signal sequences (RSS) that allow for V/J recombination or V/D/J recombination according to the 12/23 rule. Unless indicated otherwise, the segments further comprise sequences with which they are associated in nature or functional equivalents thereof (*e.g.*, for V segments promoter(s) and leader(s)).

[00158] The phrase "heavy chain," or "immunoglobulin heavy chain" includes an immunoglobulin heavy chain constant region sequence from any organism, and unless otherwise specified includes a heavy chain variable domain (V_H). V_H domains include three heavy chain CDRs and four framework (FR) regions, unless otherwise specified. Fragments of heavy chains include CDRs, CDRs and FRs, and combinations thereof. A typical heavy chain consists essentially of, following the variable domain (from N-terminal to C-terminal), a C_H1 domain, a hinge, a C_H2 domain, a C_H3 domain, and optionally a C_H4 domain (*e.g.*, in the case of IgM or IgE) and a transmembrane (M) domain (*e.g.*, in the case of membrane-bound immunoglobulin on lymphocytes). A heavy chain constant region is a region of a heavy chain that extends (from N-terminal side to C-terminal side) from outside FR4 to the C-terminal of the heavy chain. Heavy chain constant regions with minor deviations, *e.g.,* truncations of one, two, three or several amino acids from the C-terminal, would be encompassed by the phrase "heavy chain constant region," as well as heavy chain constant regions with sequence modifications, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acid substitutions. Amino acid substitutions can be made at one or more positions selected from, *e.g.* (with reference to EU numbering of an immunoglobulin constant region, *e.g.*, a human IgG constant region), 228, 233, 234, 235, 236, 237, 238, 239, 241, 248, 249, 250, 252, 254, 255, 256, 258, 265, 267, 268, 269, 270, 272, 276, 278, 280, 283, 285, 286, 289, 290, 292, 293, 294, 295, 296, 297, 298, 301, 303, 305, 307, 308, 309, 311, 312, 315, 318, 320, 322, 324, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 337, 338, 339, 340, 342, 344, 356, 358, 359, 360, 361, 362, 373, 375, 376, 378, 380, 382, 383, 384, 386, 388, 389, 398, 414, 416, 419, 428, 430, 433, 434, 435, 437, 438, and 439.

[00159] For example, and not by way of limitation, a heavy chain constant region can be modified to exhibit enhanced serum half-life (as compared with the same heavy chain constant region without the recited modification(s)) and have a modification at position 250 (*e.g.,* E or Q); 250 and 428 (*e.g.,* L or F); 252 (*e.g.,* L/Y/F/W or T), 254 (*e.g.,* S or T), and 256 (*e.g.,* S/R/Q/E/D or T); or a modification at 428 and/or 433 (*e.g.,* L/R/SI/P/Q or K) and/or 434 (*e.g.,* H/F or Y); or a modification at 250 and/or 428; or a modification at 307 or 308 (*e.g.,* 308F, V308F), and 434. In another example, the modification can comprise a 428L (*e.g.,* M428L) and 434S (*e.g.,* N434S) modification; a 428L, 259I (*e.g.,* V259I), and a 308F (*e.g.,* V308F) modification; a 433K (*e.g.,* H433K) and a 434 (*e.g.,* 434Y) modification; a 252,

254, and 256 (*e.g.,* 252Y, 254T, and 256E) modification; a 250Q and 428L modification (*e.g.,* T250Q and M428L); a 307 and/or 308 modification (*e.g.,* 308F or 308P).

[00160] The phrase "light chain" includes an immunoglobulin light chain constant (C_L) region sequence from any organism, and unless otherwise specified includes human κ and λ light chains. Light chain variable (V_L) domains typically include three light chain CDRs and four framework (FR) regions, unless otherwise specified. Generally, a full-length light chain (V_L + C_L) includes, from amino terminus to carboxyl terminus, a V_L domain that includes FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, and a C_L region. Light chains (V_L + C_L) that can be used with this invention include those, *e.g.,* that do not selectively bind either a first or second (in the case of bispecific binding proteins) epitope selectively bound by the binding protein (*e.g.,* the epitope(s) selectively bound by the V_L domain fused with the C_H domain). V_L domains that do not selectively bind the epitope(s) bound by the V_L that is fused with the C_H domain include those that can be identified by screening for the most commonly employed light chains in existing antibody libraries (wet libraries or *in silico*), wherein the light chains do not substantially interfere with the affinity and/or selectivity of the epitope binding domains of the binding proteins. Suitable light chains include those that can bind (alone or in combination with its cognate V_L fused with the C_H region) an epitope that is specifically bound by the V_L fused to the C_H region.

[00161] The phrase "micromolar range" is intended to mean 1-999 micromolar; the phrase "nanomolar range" is intended to mean 1-999 nanomolar; the phrase "picomolar range" is intended to mean 1-999 picomolar.

[00162] The term "non-human animals" is intended to include any vertebrate such as cyclostomes, bony fish, cartilaginous fish such as sharks and rays, amphibians, reptiles, mammals, and birds. Suitable non-human animals include mammals. Suitable mammals include non-human primates, goats, sheep, pigs, dogs, cows, and rodents. Suitable non-human animals are selected from the rodent family including rat and mouse. In one embodiment, the non-human animals are mice.

Mice, Nucleotide Sequences, and Binding Proteins

[00163] Binding proteins are provided that are encoded by elements of immunoglobulin loci, wherein the binding proteins comprise immunoglobulin heavy chain constant regions fused with immunoglobulin light chain variable domains. Further, multiple strategies are provided to genetically modify an immunoglobulin heavy chain locus in a mouse to encode binding proteins that contain elements encoded by immunoglobulin light chain loci. Such genetically modified mice represent a source for generating unique populations of binding

proteins that have an immunoglobulin structure, yet exhibit an enhanced diversity over traditional antibodies.

[00164] Binding protein aspects described herein include binding proteins that are encoded by modified immunoglobulin loci, which are modified such that gene segments that normally (*i.e.*, in a wild-type animal) encode immunoglobulin light chain variable domains (or portions thereof) are operably linked to nucleotide sequences that encode heavy chain constant regions. Upon rearrangement of the light chain gene segments, a rearranged nucleotide sequence is obtained that comprises a sequence encoding a light chain variable domain fused with a sequence encoding a heavy chain constant region. This sequence encodes a polypeptide that has an immunoglobulin light chain variable domain fused with a heavy chain constant region. Thus, in one embodiment, the polypeptide consists essentially of, from N-terminal to C-terminal, a V_L domain, a C_H1 region, a hinge, a C_H2 region, a C_H3 region, and optionally a C_H4 region.

[00165] In modified mice described herein, such binding proteins are made that also comprise a cognate light chain, wherein in one embodiment the cognate light chain pairs with the polypeptide described above to make a binding protein that is antibody-like, but the binding protein comprises a V_L region—not a V_H region—fused to a C_H region.

[00166] In various embodiments, the modified mice make binding proteins that comprise a V_L region fused with a C_H region (a hybrid heavy chain), wherein the V_L region of the hybrid heavy chain exhibits an enhanced degree of somatic hypermutation. In these embodiments, the enhancement is over a V_L region that is fused with a C_L region (a light chain). In some embodiments, a V_L region of a hybrid heavy chain exhibits about 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 4.5-fold, or 5-fold or more somatic hypermutations than a V_L region fused with a C_L region. In some embodiments, the modified mice in response to an antigen exhibit a population of binding proteins that comprise a V_L region of a hybrid heavy chain, wherein the population of binding proteins exhibits an average of about 1.5-fold, 2-fold, 2.5-fold, 3-fold, 3.5-fold, 4-fold, 4.5-fold, 5-fold or more somatic hypermutations in the V_L region of the hybrid heavy chain than is observed in a wild-type mouse in response to the same antigen. In one embodiment, the somatic hypermutations in the V_L region of the hybrid heavy chain comprise one or more or two or more N additions in a CDR3.

[00167] In various embodiments, the binding proteins comprise variable domains encoded by immunoglobulin light chain sequences that comprise a larger number of N additions than observed in nature for light chains rearranged from an endogenous light chain locus, *e.g.*, a binding protein comprising a mouse heavy chain constant region fused with a variable domain derived from human light chain V gene segments and human (light or heavy) J gene segments, wherein the human V and human J segments rearrange to form a rearranged gene that comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions.

[00168] In various embodiments, the mice of the invention make binding proteins that are on average smaller than wild-type antibodies (*i.e.*, antibodies that have a V_H domain), and possess advantages associated with smaller size. Smaller size is realized at least in part through the absence of an amino acid sequence encoded by a D_H region, normally present in a V_H domain. Smaller size can also be realized in the formation of a CDR3 that is derived, *e.g.,* from a $V\kappa$ region and a $J\kappa$ region.

[00169] In another aspect, a mouse and a method is provided for providing a population of binding proteins having somatically hypermutated V_L domains, *e.g.*, somatically mutated human $V\kappa$ domains, and, *e.g.*, human $V\kappa$ domains encoded by rearranged κ variable genes that comprise 1-10 or more N additions. In one embodiment, in the absence of a V_H region for generating antibody diversity, a mouse of the invention will generate binding proteins, *e.g.*, in response to challenge with an antigen, whose V domains are only or substantially V_L domains. The clonal selection process of the mouse therefore is limited to selecting only or substantially from binding proteins that have V_L domains, rather than V_H domains. Somatic hypermutation of the V_L domains will be as frequent, or substantially more frequent (*e.g.*, 2- to 5-fold higher, or more), than in wild-type mice (which also mutate V_L domains with some frequency). The clonal selection process in a mouse of the invention will generate high affinity binding proteins from the modified immunoglobulin locus, including binding proteins that specifically bind an epitope with an affinity in the nanomolar or picomolar range. Sequences that encode such binding proteins can be used to make therapeutic binding proteins containing human variable regions and human constant regions using an appropriate expression system.

[00170] In other embodiments, a mouse according to the invention can be made wherein the mouse heavy chain and/or light chain immunoglobulin loci are disabled, rendered non-functional, or knocked out, and fully human or chimeric human-mouse transgenes can be placed in the mouse, wherein at least one of the transgenes contains a modified heavy chain locus (*e.g.,* having light chain gene segments operably linked to one or more heavy chain gene sequences). Such a mouse may also make a binding protein as described herein.

[00171] In one aspect, a method is provided for increasing the diversity, including by somatic hypermutation or by N additions in a V_L domain, comprising placing an unrearranged light chain V gene segment and an unrearranged J gene segment in operable linkage with a mouse C_H gene sequence, exposing the animal to an antigen of interest, and isolating from the animal a rearranged and somatically hypermutated V(light)/J gene sequence of the animal, wherein the rearranged V(light)/J gene sequence is fused with a nucleotide sequence encoding an immunoglobulin C_H region.

[00172] In one embodiment, the immunoglobulin heavy chain fused with the hypermutated V_L is an IgM; in another embodiment, an IgG; in another embodiment, an IgE; in another embodiment, an IgA.

[00173] In one embodiment, the somatically hypermutated and class-switched V_L domain contains about 2- to 5-fold or more of the somatic hypermutations observed for a rearranged and class-switched antibody having a V_L sequence that is operably linked to a C_L sequence. In one embodiment, the observed somatic hypermutations in the somatically hypermutated V_L domain are about the same in number as observed in a V_H domain expressed from a V_H gene fused to a C_H region.

[00174] In one aspect, a method for making a high-affinity human V_L domain is provided, comprising exposing a mouse of the invention to an antigen of interest, allowing the mouse to develop an immune response to the antigen of interest, and isolating a somatically mutated, class-switched human V_L domain from the mouse that specifically binds the antigen of interest with high affinity.

[00175] In one embodiment, the K_D of a binding protein comprising the somatically mutated, class-switched human V_L domain is in the nanomolar or picomolar range.

[00176] In one embodiment, the binding protein consists essentially of a polypeptide dimer, wherein the polypeptide consists essentially of the somatically mutated, class-switched binding protein comprising a human V_L domain fused with a human C_H region.

[00177] In one embodiment, the binding protein consists essentially of a polypeptide dimer and two light chains, wherein the polypeptide consists essentially of the somatically mutated, class-switched binding protein having a human V_L domain fused with a human C_H region; and wherein each polypeptide of the dimer is associated with a cognate light chain comprising a cognate light chain V_L domain and a human C_L region.

[00178] In one aspect, a method is provided for somatically hypermutating a human V_L gene sequence, comprising placing a human V_L gene segment and a human J_L gene segment in operable linkage with an endogenous mouse C_H gene at an endogenous mouse heavy chain immunoglobulin locus, exposing the mouse to an antigen of interest, and obtaining from the mouse a somatically hypermutated human V_L gene sequence that binds the antigen of interest.

[00179] In one embodiment, the method further comprises obtaining from the mouse a V_L gene sequence from a light chain that is cognate to the human somatically hypermutated human V_L gene sequence that binds the antigen of interest.

V_L Binding Proteins with D_H Sequences

[00180] In various aspects, mice comprising an unrearranged immunoglobulin light chain V gene segment and an unrearranged (*e.g.*, light or heavy) J gene segment also comprise

an unrearranged DH gene segment that is capable of recombining with the J segment to form a rearranged D/J sequence, which in turn is capable of rearranging with the light chain V segment to form a rearranged variable sequence derived from (a) the light chain V segment, (b) the DH segment, and (c) the (*e.g.*, light or heavy) J segment; wherein the rearranged variable sequence is operably linked to a heavy chain constant sequence (*e.g.*, selected from CH1, hinge, CH2, CH3, and a combination thereof; *e.g.*, operably linked to a mouse or human CH1, a hinge, a CH2, and a CH3).

[00181] In various aspects, mice comprising unrearranged human light chain V segments and J segments that also comprise a human D segment are useful, *e.g.*, as a source of increased diversity of CDR3 sequences. Normally, CDR3 sequences arise in light chains from V/J recombination, and in heavy chains from V/D/J recombination. Further diversity is provided by nucleotide additions that occur during recombination (*e.g.*, N additions), and also as the result of somatic hypermutation. Binding characteristics conferred by CDR3 sequences are generally limited to those conferred by the light chain CDR3 sequence, the heavy chain CDR3 sequence, and a combination of the light and the heavy chain CDR3 sequence, as the case may be. In mice as described herein, however, an added source of diversity is available due to binding characteristics conferred as the result of a combination of a first light chain CDR3 (on the heavy chain polypeptide) and a second light chain CDR3 (on the light chain polypeptide). Further diversity is possible where the first light chain CDR3 may contain a sequence derived from a D gene segment, as from a mouse as described herein that comprises an unrearranged V segment from a light chain V region operably linked to a D segment and operably linked to a J segment (light or heavy), employing the RSS engineering as taught here.

[00182] Another source of diversity is the N and/or P additions that can occur in the V(light)/J or V(light)/D/J recombinations that are possible in mice as described. Thus, mice described herein not only provide a different source of diversity (light chain-light chain) but also a further source of diversity due to the addition of, *e.g.,* 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more N additions in a rearranged V(light)/J or a rearranged V(light)/D/J gene in a mouse as described herein.

[00183] In various aspects the use of a D gene segment operably linked to a J gene segment and a light chain V gene segment provides an enhanced diversity. Operable linkage of a D_H segment in this instance will require that that D segment is capable of recombining with the J segment with which it is recited. Thus, the D segment will require to have juxtaposed a downstream RSS that is matched to the RSS juxtaposed upstream of the J segment such that the D segment and the J segment may rearrange. Further, the D segment will require an appropriate RSS juxtaposed upstream that is matched to the RSS

juxtaposed downstream of the V segment such that the rearranged D/J segment and the V segment may rearrange to form a gene encoding a variable domain.

[00184] An RSS, or a recombination signal sequence, comprises a conserved nucleic acid heptamer sequence separated, by 12 base pairs (bp) or 23 base pairs (bp) of unconserved sequence, from a conserved nucleic acid nonamer sequence. RSS's are used by recombinases to achieve joining of immunoglobulin gene segments during the rearrangement process following the 12/23 rule. According to the 12/23 rule, a gene segment juxtaposed with an RSS having a 12 bp (unconserved) spacer rearranges with a gene segment juxtaposed with an RSS having a 23 bp (unconserved) spacer; *i.e.,* rearrangements between gene segments each having an RSS with a 12 bp spacer, or each having an RSS with a 23 bp spacer, are generally not observed.

[00185] In the case of the λ light chain locus, variable gene segments (Vλ gene segments) are flanked downstream (with respect to the direction of transcription of the V sequence) with an RSS having a 23-mer spacer, and joining gene segments (Jλ gene segments) are flanked upstream (with respect to the direction of transcription of the J sequence) with an RSS having a 12-mer spacer. Thus, Vλ and Jλ segments are flanked with RSS's that are compatible under the 12/23 rule, and therefore are capable of recombine during rearrangement.

[00186] At the κ locus in a wild-type organism, however, each functional Vκ segment is flanked downstream with an RSS having a 12-mer spacer. Jκ segments, therefore, have 23-mer spaces juxtaposed on the upstream side of the Jκ segment. At the heavy chain locus, V_H gene segments are juxtaposed downstream with an RSS having a 23-mer spacer, followed by D_H segment juxtaposed upstream and downstream with a 12-mer spacer, and J_H segments each with a 23-mer segment juxtaposed on the upstream side of the J_H segment. At the heavy chain locus, D/J recombination occurs first, mediated by the downstream D_H RSS with the 12-mer spacer and the upstream J_H RSS with the 23-mer spacer, to yield an intermediate rearranged D-J sequence having an RSS juxtaposed on the upstream side that has an RSS with a 12-mer spacer. The rearranged D-J segment having the RSS with the 12-mer juxtaposed on the upstream side then rearranges with the V_H segment having the RSS with the 23-mer juxtaposed on its downstream side to form a rearranged V/D/J sequence.

[00187] In one embodiment, a Vλ segment is employed at the heavy chain locus with a J gene segment that is a Jλ segment, wherein the Vλ segment comprises an RSS juxtaposed on the downstream side of the Vλ sequence, and the RSS comprises a 23-mer spacer, and the J segment is a Jλ segment with an RSS juxtaposed on its upstream side having a 12-mer spacer (*e.g.,* as found in nature).

[00188] In one embodiment, a Vλ segment is employed at the heavy chain locus with a J gene segment that is a Jκ or a J_H gene segment, wherein the Vλ sequence has juxtaposed on its downstream side an RSS comprising a 23-mer spacer, and the Jκ or J_H segment has juxtaposed on its upstream side an RSS comprising a 12-mer spacer.

[00189] In one embodiment, a Vλ segment is employed at the heavy chain locus with a D_H gene segment and a J gene segment. In one embodiment, the Vλ segment comprises an RSS juxtaposed on the downstream side of the Vλ sequence with an RSS having a 23-mer spacer; the D_H segment comprises an RSS juxtaposed on the upstream side and on the downstream side of the D_H sequence with an RSS having a 12-mer spacer; and a J segment having an RSS juxtaposed on its upstream side having a 23-mer spacer, wherein the J segment is selected from a Jλ, a Jκ, and a J_H.

[00190] In one embodiment, a Vκ segment is employed at the heavy chain locus with a J gene segment (with no intervening D segment), wherein the Vκ segment has an RSS juxtaposed on the downstream side of the Vκ segment that comprises a 12-mer spaced RSS, and the J segment has juxtaposed on its upstream side a 23-mer spaced RSS, and the Jκ segment is selected from a Jκ segment, a Jλ segment, and a J_H segment. In one embodiment, the V segment and/or the J segment are human.

[00191] In one embodiment, the Vκ segment is employed at the heavy chain locus with a D segment and a J segment, wherein the Vκ segment has an RSS juxtaposed on the downstream side of the Vκ segment that comprises a 12-mer spaced RSS, the D segment has juxtaposed on its upstream and downstream side a 23-mer spaced RSS, and the J segment has juxtaposed on its upstream side a 12-mer spaced RSS. In one embodiment, the J segment is selected from a Jκ segment, a Jλ segment, and a J_H segment. In one embodiment, the V segment and/or the J segment are human.

[00192] A Jλ segment with an RSS having a 23-mer spacer juxtaposed at its upstream end, or a Jκ or J_H segment with an RSS having a 12-mer spacer juxtaposed at its upstream end, is made using any suitable method for making nucleic acid sequences that is known in the art. A suitable method for making a J segment having an RSS juxtaposed upstream wherein the RSS has a selected spacer (*e.g.*, either 12-mer or 23-mer) is to chemically synthesize a nucleic acid comprising the heptamer, the nonamer, and the selected spacer and fuse it to a J segment sequence that is either chemically synthesized or cloned from a suitable source (*e.g.,* a human sequence source), and employ the fused J segment sequence and RSS in a targeting vector to target the RSS-J to a suitable site.

[00193] A D segment with a 23-mer spaced RSS juxtaposed upstream and downstream can be made by any method known in the art. One method comprises chemically synthesizing the upstream 23-mer RSS and D segment sequence and the downstream 23-

mer RSS, and placing the RSS-flanked D segment in a suitable vector. The vector may be directed to replace one or more mouse D segments with a human D segment with 12-mer RSS sequences juxtaposed on the upstream and downstream sides, or directed to be inserted into, *e.g.*, a humanized locus at a position between a human V segment and a human or mouse J segment.

[00194] Suitable nonamers and heptamers for RSS construction are known in the art (e.g., see Janeway's Immunobiology, 7th ed., Murphy et al., (2008, Garland Science, Taylor & Francis Group, LLC) at page 148, Fig. 4.5, incorporated by reference). Suitable nonconserved spacer sequences include, *e.g.*, spacer sequences observed in RSS sequences at human or mouse immunoglobulin loci.

Bispecific-Binding Proteins

[00195] The binding proteins described herein, and nucleotide sequences encoding them, can be used to make multispecific binding proteins, *e.g.*, bispecific binding proteins. In this aspect, a first polypeptide consisting essentially of a first V_L domain fused with a C_H region can associate with a second polypeptide consisting essentially of a second V_L domain fused with a C_H region. Where the first V_L domain and the second V_L domain specifically bind a different epitope, a bispecific-binding molecule can be made using the two V_L domains. The C_H region can be the same or different. In one embodiment, *e.g.*, one of the C_H regions can be modified so as to eliminate a protein A binding determinant, whereas the other heavy chain constant region is not so modified. This particular arrangement simplifies isolation of the bispecific binding protein from, *e.g.*, a mixture of homodimers (*e.g.,* homodimers of the first or the second polypeptides).

[00196] In one aspect, the methods and compositions described herein are used to make bispecific-binding proteins. In this aspect, a first V_L that is fused to a C_H region and a second V_L that is fused to a C_H region are each independently cloned in frame with a human IgG sequence of the same isotype (*e.g.*, a human IgG1, IgG2, IgG3, or IgG4). The first V_L specifically binds a first epitope, and the second V_L specifically binds a second epitope. The first and second epitopes may be on different antigens, or on the same antigen.

[00197] In one embodiment, the IgG isotype of the C_H region fused to the first V_L and the IgG isotype of the C_H region fused to the second V_L are the same isotype, but differ in that one IgG isotype comprises at least one amino acid substitution. In one embodiment, the at least one amino acid substitution renders the heavy chain bearing the substitution unable or substantially unable to bind protein A as compared with the heavy chain that lacks the substitution.

[00198] In one embodiment, the first C_H region comprises a first C_H3 domain of a human IgG selected from IgG1, IgG2, and IgG4; and the second C_H region comprises a second C_H3

domain of a human IgG selected from IgG1, IgG2, and IgG4, wherein the second C_H3 domain comprises a modification that reduces or eliminates binding of the second C_H3 domain to protein A.

[00199] In one embodiment, the second C_H3 domain comprises a 435R modification, numbered according to the EU index of Kabat. In another embodiment, the second C_H3 domain further comprises a 436F modification, numbered according to the EU index of Kabat.

[00200] In one embodiment, the second C_H3 domain is that of a human IgG1 that comprises a modification selected from the group consisting of D356E, L358M, N384S, K392N, V397M, and V422I, numbered according to the EU index of Kabat.

[00201] In one embodiment, the second C_H3 domain is that of a human IgG2 that comprises a modification selected from the group consisting of N384S, K392N, and V422I, numbered according to the EU index of Kabat.

[00202] In one embodiment, the second C_H3 domain is that of a human IgG4 comprising a modification selected from the group consisting of Q355R, N384S, K392N, V397M, R409K, E419Q, and V422I, numbered according to the EU index of Kabat.

[00203] In one embodiment, the binding protein comprises C_H regions having one or more modifications as recited herein, wherein the constant region of the binding protein is nonimmunogenic or substantially nonimmunogenic in a human. In a specific embodiment, the C_H regions comprise amino acid sequences that do not present an immunogenic epitope in a human. In another specific embodiment, the binding protein comprises a C_H region that is not found in a wild-type human heavy chain, and the C_H region does not comprise a sequence that generates a T-cell epitope.

EXAMPLES

[00204] The following examples are provided so as to describe how to make and use methods and compositions of the invention, and are not intended to limit the scope of what the inventors regard as their invention. Unless indicated otherwise, temperature is indicated in Celsius, and pressure is at or near atmospheric.

Example I

Introduction of Light Chain Gene Segments Into A Heavy Chain Locus

[00205] Various targeting constructs were made using VELOCIGENE® genetic engineering technology (see, *e.g.*, US Pat. No. 6,586,251 and Valenzuela, D.M., Murphy, A.J., Frendewey, D., Gale, N.W., Economides, A.N., Auerbach, W., Poueymirou, W.T., Adams, N.C., Rojas, J., Yasenchak, J., Chernomorsky, R., Boucher, M., Elsasser, A.L., Esau, L., Zheng, J., Griffiths, J.A., Wang, X., Su, H., Xue, Y., Dominguez, M.G., Noguera, I.,

Torres, R., Macdonald, L.E., Stewart, A.F., DeChiara, T.M., Yancopoulos, G.D. (2003). High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. Nat Biotechnol 21, 652-659) to modify mouse genomic Bacterial Artificial Chromosome (BAC) libraries. Mouse BAC DNA was modified by homologous recombination to inactivate the endogenous mouse heavy chain locus through targeted deletion of V_H, D_H and J_H gene segments for the ensuing insertion of unrearranged human germline κ light chain gene sequences (top of FIG 2).

[00206] Briefly, the mouse heavy chain locus was deleted in two successive targeting events using recombinase-mediated recombination. The first targeting event included a targeting at the 5' end of the mouse heavy chain locus using a targeting vector comprising from 5' to 3' a 5' mouse homology arm, a recombinase recognition site, a neomycin cassette and a 3' homology arm. The 5' and 3' homology arms contained sequence 5' of the mouse heavy chain locus. The second targeting event included a targeting at the 3' end of the mouse heavy chain locus in the region of the J_H gene segments using a second targeting vector that contained from 5' to 3' a 5' mouse homology arm, a 5' recombinase recognition site, a second recombinase recognition site, a hygromycin cassette, a third recombinase recognition site, and a 3' mouse homology arm. The 5' and 3' homology arms contained sequence flanking the mouse J_H gene segments and 5' of the intronic enhancer and constant regions. Positive ES cells containing a modified heavy chain locus targeted with both targeting vectors (as described above) were confirmed by karyotyping. DNA was then isolated from the double-targeted ES cells and subjected to treatment with a recombinase thereby mediating the deletion of genomic DNA of the mouse heavy chain locus between the 5' recombinase recognition site in the first targeting vector and the 5' recombinase recognition site in the second targeting vector, leaving a single recombinase recognition site and the hygromycin cassette flanked by two recombinase recognition sites (see top of FIG. 2). Thus a modified mouse heavy chain locus containing intact C_H genes was created for progressively inserting human κ germline gene segments in a precise manner using targeting vectors described below.

[00207] Four separate targeting vectors were engineered to progressively insert 40 human Vκ gene segments and five human Jκ gene segments into the inactivated mouse heavy chain locus (described above) using standard molecular techniques recognized in the art (FIG. 2). The human κ gene segments used for engineering the four targeting constructs are naturally found in proximal contig of the germline human κ light chain locus (FIG. 1B and Table 1).

[00208] A ~110,499 bp human genomic fragment containing the first six human Vκ gene segments and five human Jκ gene segments was engineered to contain a PI-SceI site 431

bp downstream (3') of the human Jκ5 gene segment. Another PI-SceI site was engineered at the 5' end of a ~7,852 bp genomic fragment containing the mouse heavy chain intronic enhancer, the IgM switch region (Sμ) and the IgM gene of the mouse heavy chain locus. This mouse fragment was used as a 3' homology arm by ligation to the ~110.5 kb human fragment, which created a 3' junction containing, from 5' to 3', ~110.5 kb of genomic sequence of the human κ light chain locus containing the first six consecutive Vκ gene segments and five Jκ gene segments, a PI-SceI site, ~7,852 bp of mouse heavy chain sequence containing the mouse intronic enhancer, Sμ and the mouse IgM constant gene. Upstream (5') from the human Vκ1-6 gene segment was an additional 3,710 bp of human κ sequence before the start of the 5' mouse homology arm, which contained 19,752 bp of mouse genomic DNA corresponding to sequence 5' of the mouse heavy chain locus. Between the 5' homology arm and the beginning of the human κ sequence was a neomycin cassette flanked by three recombinase recognition sites (see Targeting Vector 1, FIG. 2). The final targeting vector for the first insertion of human κ sequence from 5' to 3' included a 5' homology arm containing ~20 kb of mouse genomic sequence 5' of the heavy chain locus, a first recombinase recognition site (R1), a neomycin cassette, a second recombinase recognition site (R2), a third recombinase recognition site (R3), ~110.5 kb of human genomic κ sequence containing the first six consecutive human Vκ gene segments and five human Jκ gene segments, a PI-SceI site, and a 3' homology arm containing ~8 kb of mouse genomic sequence including the intronic enhancer, Sμ and the mouse IgM constant gene (see FIG. 2, Targeting Vector 1). Homologous recombination with this targeting vector created a modified mouse heavy chain locus containing six human Vκ gene segments and five human Jκ gene segments operably linked to the endogenous mouse heavy chain constant genes which, upon recombination, leads to the formation of a hybrid heavy chain (*i.e.*, a human Vκ domain and a mouse C_H region).

Table 1

Targeting Vector	Size of Human κ Sequence	Human κ Gene Segments Added	
		Vκ	Jκ
1	~110.5 kb	4-1, 5-2, 7-3, 2-4, 1-5, 1-6	1 – 5
2	~140 kb	3-7, 1-8, 1-9, 2-10, 3-11, 1-12, 1-13, 2-14, 3-15, 1-16	-
3	~161 kb	1-17, 2-18, 2-19, 3-20, 6-21, 1-22, 1-23, 2-24, 3-25, 2-26, 1-27, 2-28, 2-29, 2-30	-
4	~90 kb	3-31, 1-32, 1-33, 3-34, 1-35, 2-36, 1-37, 2-38, 1-39, 2-40	-

[00209] Introduction of ten additional human Vκ gene segments into a hybrid heavy chain locus. A second targeting vector was engineered for introduction of 10 additional human Vκ gene segments to the modified mouse heavy chain locus described above (see FIG. 2, Targeting Vector 2). A 140,058 bp human genomic fragment containing 12 consecutive human Vκ gene segments from the human κ light chain locus was engineered with a 5' homology arm containing mouse genomic sequence 5' of the mouse heavy chain locus and a 3' homology arm containing human genomic κ sequence. Upstream (5') from the human Vκ1-16 gene segment was an additional 10,170 bp of human κ sequence before the start of the 5' mouse homology arm, which was the same 5' homology arm used for construction of Targeting Vector 1 (see FIG. 2). Between the 5' homology arm and the beginning of the human κ sequence was a hygromycin cassette flanked by recombinase recognition sites. The 3' homology arm included a 31,165 bp overlap of human genomic κ sequence corresponding to the equivalent 5' end of the ~110.5 kb fragment of human genomic κ sequence of Targeting Vector 1 (FIG. 2). The final targeting vector for the insertion of 10 additional human Vκ gene segments from 5' to 3' included a 5' homology arm containing ~20 kb of mouse genomic sequence 5' of the heavy chain locus, a first recombinase recognition site (R1), a hygromycin cassette, a second recombinase recognition site (R2) and ~140 kb of human genomic κ sequence containing 12 consecutive human Vλ gene segments, ~31 kb of which overlaps with the 5' end of the human κ sequence of Targeting Vector 1 and serves as the 3' homology arm for this targeting construct. Homologous recombination with this targeting vector created a modified mouse heavy chain locus containing 16 human Vκ gene segments and five human Jκ gene

segments operably linked to the mouse heavy chain constant genes which, upon recombination, leads to the formation of a hybrid heavy chain.

[00210] Introduction of fourteen additional human Vκ gene segments into a hybrid heavy chain locus. A third targeting vector was engineered for introduction of 14 additional human Vκ gene segments to the modified mouse heavy chain locus described above (see FIG. 2, Targeting Vector 3). A 160,579 bp human genomic fragment containing 15 consecutive human Vκ gene segments was engineered with a 5' homology arm containing mouse genomic sequence 5' of the mouse heavy chain locus and a 3' homology arm containing human genomic κ sequence. Upstream (5') from the human Vκ2-30 gene segment was an additional 14,687 bp of human κ sequence before the start of the 5' mouse homology arm, which was the same 5' homology used for the previous two targeting vectors (described above, see also FIG. 2). Between the 5' homology arm and the beginning of the human κ sequence was a neomycin cassette flanked by recombinase recognition sites. The 3' homology arm included a 21,275 bp overlap of human genomic κ sequence corresponding to the equivalent 5' end of the ~140 kb fragment of human genomic κ sequence of Targeting Vector 2 (FIG. 2). The final targeting vector for the insertion of 14 additional human Vκ gene segments from 5' to 3' included a 5' homology arm containing ~20 kb of mouse genomic sequence 5' of the mouse heavy chain locus, a first recombinase recognition site (R1), a neomycin cassette, a second recombinase recognition site (R2) and ~161 kb of human genomic κ sequence containing 15 human Vκ gene segments, ~21 kb of which overlaps with the 5' end of the human κ sequence of Targeting Vector 2 and serves as the 3' homology arm for this targeting construct. Homologous recombination with this targeting vector created a modified mouse heavy chain locus containing 30 human Vκ gene segments and five human Jκ gene segments operably linked to the mouse heavy chain constant genes which, upon recombination, leads to the formation of a chimeric κ heavy chain.

[00211] Introduction of ten additional human Vκ gene segments into a hybrid heavy chain locus. A fourth targeting vector was engineered for introduction of 10 additional human Vκ gene segments to the modified mouse heavy chain locus described above (see FIG. 2, Targeting Vector 4). A 90,398 bp human genomic fragment containing 16 consecutive human Vκ gene segments was engineered with a 5' homology arm containing mouse genomic sequence 5' of the mouse heavy chain locus and a 3' homology arm containing human genomic κ sequence. Upstream (5') from the human Vκ2-40 gene segment was an additional 8,484 bp of human κ sequence before the start of the 5' mouse homology arm, which was the same 5' homology as the previous targeting vectors

(described above, see also FIG. 2). Between the 5' homology arm and the beginning of the human κ sequence was a hygromycin cassette flanked by recombinase recognition sites. The 3' homology arm included a 61,615 bp overlap of human genomic κ sequence corresponding to the equivalent 5' end of the ~160 kb fragment of human genomic κ sequence of Targeting Vector 3 (FIG. 2). The final targeting vector for the insertion of 10 additional human Vκ gene segments from 5' to 3' included a 5' homology arm containing ~20 kb of mouse genomic sequence 5' of the mouse heavy chain locus, a first recombinase recognition site (R1), a hygromycin cassette, a second recombinase recognition site (R2) and ~90 kb of human genomic κ sequence containing 16 human Vκ gene segments, ~62 kb of which overlaps with the 5' end of the human κ sequence of Targeting Vector 3 and serves as the 3' homology arm for this targeting construct. Homologous recombination with this targeting vector created a modified mouse heavy chain locus containing 40 human Vκ gene segments and five human Jκ gene segments operably linked to the mouse heavy chain constant genes which, upon recombination, leads to the formation of a chimeric κ heavy chain (FIG. 3).

[00212] Using a similar approach as described above, other combinations of human light chain variable domains in the context of mouse heavy chain constant regions are constructed. Additional light chain variable domains may be derived from human Vλ and Jλ gene segments (FIG. 4A and 4B).

[00213] The human λ light chain locus extends over 1,000 kb and contains over 80 genes that encode variable (V) or joining (J) segments. Among the 70 Vλ gene segments of the human λ light chain locus, anywhere from 30-38 appear to be functional gene segments according to published reports. The 70 Vλ sequences are arranged in three clusters, all of which contain different members of distinct V gene family groups (clusters A, B and C). Within the human λ light chain locus, over half of all observed Vλ domains are encoded by the gene segments 1-40, 1-44, 2-8, 2-14, and 3-21. There are seven Jλ gene segments, only four of which are regarded as generally functional Jλ gene segments—Jλ1, Jλ2, Jλ3, and Jλ7. In some alleles, a fifth Jλ-Cλ gene segment pair is reportedly a pseudo gene (Cλ6). Incorporation of multiple human Jλ gene segments into a hybrid heavy chain locus, as described herein, is constructed by de novo synthesis. In this way, a genomic fragment containing multiple human Jλ gene segments in germline configuration is engineered with multiple human Vλ gene segments and allow for normal V-J recombination in the context of a heavy chain constant region.

[00214] Coupling light chain variable domains with heavy chain constant regions represents a potentially rich source of diversity for generating unique V_L binding proteins with

human V_L regions in non-human animals. Exploiting this diversity of the human λ light chain locus (or human κ locus as described above) in mice results in the engineering of unique hybrid heavy chains and gives rise to another dimension of binding proteins to the immune repertoire of genetically modified animals and their subsequent use as a next generation platform for the generation of therapeutics.

[00215] Additionally, human D_H and J_H (or Jκ) gene segments can be incorporated with either human Vκ or Vλ gene segments to construct novel hybrid loci that will give rise, upon recombination, to novel engineered variable domains (FIG. 5A and 5B). In this latter case, engineering combinations of gene segments that are not normally contained in a single locus would require specific attention to the recombination signal sequences (RSS) that are associated with respective gene segments such that normal recombination can be achieved when they are combined into a single locus. For example, V(D)J recombination is known to be guided by conserved noncoding DNA sequences, known as heptamer and nonamer sequences that are found adjacent to each gene segment at the precise location at which recombination takes place. Between these noncoding DNA sequences are nonconserved spacer regions that either 12 or 23 base pairs (bp) in length. Generally, recombination only occurs at gene segments located on the same chromosome and those gene segments flanked by a 12-bp spacer can be joined to a gene segment flanked by a 23-bp spacer, *i.e.* the 12/23 rule, although joining two of D_H gene segments (each flanked by 12-bp spacers) has been observed in a small proportion of antibodies. To allow for recombination between gene segments that do not normally have compatible spacers (*e.g.,* Vκ and a D_H or D_H and Jλ), unique, compatible spacers are synthesized in adjacent locations with the desired gene segments for construction of unique hybrid heavy chains that allow for successful recombination to form unique heavy chains containing light chain variable regions.

[00216] Thus, using the strategy outlined above for incorporation of human κ light chain gene segments into an endogenous heavy chain locus allows for the use of other combinations of human λ light chain gene segments as well as specific human heavy chain gene segments (*e.g.,* D_H and J_H) and combinations thereof.

Example II

Identification of Targeted ES cells Bearing Human Light Chain Gene Segments at an Endogenous Heavy Chain Locus

[00217] The targeted BAC DNA made in the foregoing Examples was used to electroporate mouse ES cells to created modified ES cells for generating chimeric mice that express V_L binding proteins (*i.e.,* human κ light chain gene segments operably linked to mouse heavy chain constant regions). ES cells containing an insertion of unrearranged

human κ light chain gene segments were identified by the quantitative PCR assay, TAQMAN® (Lie and Petropoulos, 1998. Curr. Opin. Biotechnology 9:43-48). Specific primers sets and probes were design for insertion of human κ sequences and associated selection cassettes, loss of mouse heavy chain sequences and retention of mouse sequences flanking the endogenous heavy chain locus.

[00218] ES cells bearing the human κ light chain gene segments can be transfected with a construct that expresses a recombinase in order to remove any undesired selection cassette introduced by the insertion of the targeting construct containing human κ gene segments. Optionally, the selection cassette may be removed by breeding to mice that express the recombinase (*e.g.*, US 6,774,279). Optionally, the selection cassette is retained in the mice.

Example III

Generation and Analysis of Mice Expressing V_L Binding Proteins

[00219] Targeted ES cells described above were used as donor ES cells and introduced into an 8-cell stage mouse embryo by the VELOCIMOUSE® method (see, *e.g.*, US Pat. No. 7,294,754 and Poueymirou, W.T., Auerbach, W., Frendewey, D., Hickey, J.F., Escaravage, J.M., Esau, L., Dore, A.T., Stevens, S., Adams, N.C., Dominguez, M.G., Gale, N.W., Yancopoulos, G.D., DeChiara, T.M., Valenzuela, D.M. (2007). F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. Nat Biotechnol 25, 91-99). VELOCIMICE® (F0 mice fully derived from the donor ES cell) independently bearing human κ gene segments at the mouse heavy chain locus were identified by genotyping using a modification of allele assay (Valenzuela *et al.*, *supra*) that detected the presence of the unique human κ gene segments at the endogenous heavy chain locus (*supra*). Pups are genotyped and a pup heterozygous for the hybrid heavy chain gene locus is selected for characterizing expression of V_L binding proteins.

[00220] Flow Cytometry. The introduction of human κ light chain gene segments into the mouse heavy chain locus was carried out in an F1 ES line (F1H4; Valenzuela *et al.* 2007, *supra*) derived from 129S6/SvEvTac and C57BL/6NTac heterozygous embryos that further contained an *in situ* replacement of the mouse κ light chain gene segments with human κ light chain gene segments (US 6,596,541). The human κ light chain germline variable gene segments are targeted to the 129S6 allele, which carries the IgM^a haplotype, whereas the unmodified mouse C576BL/6N allele bears the IgM^b haplotype. These allelic forms of IgM can be distinguished by flow cytometry using antibodies specific to the polymorphisms found in the IgM^a or IgM^b alleles. Heterozygous mice bearing human κ light chain gene segments at the endogenous heavy chain locus as described in Example I were evaluated for expression of human V_L binding proteins using flow cytometry.

[00221] Briefly, blood was drawn from groups of mice (n=6 per group) and grinded using glass slides. C57BL/6 and Balb/c mice were used as control groups. Following lysis of red blood cells (RBCs) with ACK lysis buffer (Lonza Walkersville), cells were resuspended in BD Pharmingen FACS staining buffer and blocked with anti-mouse CD16/32 (BD Pharmingen). Lymphocytes were stained with anti-mouse IgM^b-FITC (BD Pharmingen), anti-mouse IgM^a-PE (BD Pharmingen), anti-mouse CD19 (Clone 1D3; BD Biosciences), and anti-mouse CD3 (17A2; BIOLEGEND®) followed by fixation with BD CYTOFIX™ all according to the manufacturer's instructions. Final cell pellets were resuspended in staining buffer and analyzed using a BD FACSCALIBUR™ and BD CELLQUEST PRO™ software. Table 2 sets forth the average percent values for B cells ($CD19^+$), T cells ($CD3^+$), hybrid heavy chain ($CD19^+IgM^{a+}$), and wild type heavy chain ($CD19^+IgM^{b+}$) expression observed in groups of animals bearing each genetic modification.

[00222] In a similar experiment, B cell contents of the spleen, blood and bone marrow compartments from mice homozygous for six human Vκ and five human Jκ gene segments operably linked to the mouse heavy chain constant region (described in Example I, FIG. 2) were analyzed for progression through B cell development using flow cytometry of various cell surface markers.

[00223] Briefly, two groups (n=3 each, 8 weeks old females) of wild type and mice homozygous for six human Vκ and five human Jκ gene segments operably linked to the mouse heavy chain constant region were sacrificed and blood, spleens and bone marrow were harvested. Blood was collected into microtainer tubes with EDTA (BD Biosciences). Bone marrow was collected from femurs by flushing with complete RPMI medium (RPMI medium supplemented with fetal calf serum, sodium pyruvate, Hepes, 2-mercaptoethanol, non-essential amino acids, and gentamycin). RBCs from spleen and bone marrow preparations were lysed with ACK lysis buffer (Lonza Walkersville), followed by washing with complete RPMI medium.

[00224] Cells ($1x10^6$) were incubated with anti-mouse CD16/CD32 (2.4G2, BD) on ice for ten minutes, followed by labeling with the following antibody cocktail for thirty minutes on ice: anti-mouse FITC-CD43 (1B11, BIOLEGEND®), PE-ckit (2B8, BIOLEGEND®), PeCy7-IgM (II/41, EBIOSCIENCE®), PerCP-Cy5.5-IgD (11-26c.2a, BIOLEGEND®), APC-eFluor 780-B220 (RA3-6B2, EBIOSCIENCE®), APC-CD19 (MB19-1, EBIOSCIENCE®). Bone marrow: immature B cells ($B220^{int}IgM^+$), mature B cells ($B220^{hi}IgM^+$), pro B cells ($CD19^+ckit^+CD43^+$), pre B cells ($CD19^+ckit^-CD43^-$), pre-B cells ($CD19^+CD43^{int}IgM^{+/-}$), immature B cells ($CD19^+CD43^-IgM^{+/-}$). Blood and spleen: B cells ($CD19^+$), mature B cells ($CD19^+IgM^{int}IgD^{hi}$), transitional/immature B cells ($CD19^+IgM^{hi}IgD^{int}$).

[00225] Following staining, cells were washed and fixed in 2% formaldehyde. Data acquisition was performed on a LSRII flow cytometer and analyzed with FLOWJO™ software (Tree Star, Inc.). FIG. 6A, 6B and 6C show the results for the splenic compartment. FIG. 7A – 7G show the results for the bone marrow compartment. The results obtained for the blood compartment from each group of mice demonstrated similar results as compared to the results from the splenic compartment from each group (data not shown).

[00226] In a similar experiment, B cell contents of the spleen, blood and bone marrow compartments from mice homozygous for thirty human Vκ and five human Jκ gene segments operably linked to the mouse heavy chain constant region (described in Example I, FIG. 2) were analyzed for progression through B cell development using flow cytometry of various cell surface markers.

[00227] Briefly, two groups (N=3 each, 6 week old females) of mice containing a wild-type heavy chain locus and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segments (WT) and mice homozygous for thirty hVκ and five Jκ gene segments and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segments (30hVκ-5hJκ HO) were sacrificed and spleens and bone marrow were harvested. Bone marrow and splenocytes were prepared for staining with various cell surface markers (as described above).

[00228] Cells (1 x 10^6) were incubated with anti-mouse CD16/CD32 (2.4G2, BD Biosciences) on ice for ten minutes, followed by labeling with bone marrow or splenocyte panels for thirty minutes on ice. Bone marrow panel: anti-mouse FITC-CD43 (1B11, BIOLEGEND®), PE-ckit (2B8, BIOLEGEND®), PeCy7-IgM (II/41, EBIOSCIENCE®), APC-CD19 (MB19-1, EBIOSCIENCE®). Bone marrow and spleen panel: anti-mouse FITC-Igκ (187.1 BD Biosciences), PE-Igλ (RML-42, BIOLEGEND®), PeCy7-IgM (II/41, EBIOSCIENCE®), PerCP-Cy5.5-IgD (11-26c.2a, BIOLEGEND®), Pacific Blue-CD3 (17A2, BIOLEGEND®), APC-B220 (RA3-6B2, EBIOSCIENCE®), APC-H7-CD19 (ID3, BD). Bone marrow: immature B cells ($B220^{int}IgM^+$), mature B cells ($B220^{hi}IgM^+$), pro B cells ($CD19^+ckit^+CD43^+$), pre B cells (CD19+ckit-CD43-), immature $Ig\kappa^+$ B cells ($B220^{int}IgM^+Ig\kappa^+Ig\lambda^-$), immature $Ig\lambda^+$ B cells ($B220^{int}IgM^+Ig\kappa^-Ig\lambda^+$), mature $Ig\kappa^+$ B cells ($B220^{hi}IgM^+Ig\kappa^+Ig\lambda^-$), mature $Ig\lambda^+$ B cells ($B220^{hi}IgM^+Ig\kappa^-Ig\lambda^+$). Spleen: B cells ($CD19^+$), mature B cells ($CD19^+IgD^{hi}IgM^{int}$), transitional/immature B cells ($CD19^+IgD^{int}IgM^{hi}$). Bone marrow and spleen: $Ig\kappa^+$ B cells ($CD19^+Ig\kappa^+Ig\lambda^-$), $Ig\lambda^+$ B cells ($CD19^+ Ig\kappa^-Ig\lambda^+$).

[00229] Following staining, cells were washed and fixed in 2% formaldehyde. Data acquisition was performed on a LSRII flow cytometer and analyzed with FLOWJO™ software (Tree Star, Inc.). The results demonstrated similar staining patterns and cell

populations for all three compartments as compared to mice homozygous for six human Vκ and five human Jκ gene segments (described above). However, these mice demonstrated a loss in endogenous λ light chain expression in both the splenic and bone marrow compartments (FIG. 8A and 8B, respectively), despite the endogenous λ light chain locus being intact in these mice. This may reflect an inability of rearranged human κ light chain domains, in the context of heavy chain constant regions, to pair or associate with murine λ light chain domains, leading to deletion of $Ig\lambda^+$ cells.

[00230] Isotype Expression. Total and surface (*i.e.,* membrane bound) immunoglobulin M (IgM) and immunoglobulin G1 (IgG1) was determined for mice homozygous for human heavy and κ light chain variable gene loci (VELCOIMMUNE® Humanized Mice, see US 7,105,348) and mice homozygous for six human Vκ and 5 human Jκ gene segments engineered into the endogenous heavy chain locus (6hVκ-5hJκ HO) by a quantitative PCR assay using TAQMAN® probes (as described above in Example II).

[00231] Briefly, $CD19^+$ B cells were purified from the spleens of groups of mice (n=3 to 4 mice per group) using mouse CD19 Microbeads (Miltenyi Biotec) according to manufacturer's instructions. Total RNA was purified using the RNEASY™ Mini kit (Qiagen). Genomic RNA was removed using an RNase-free DNase on-column treatment (Qiagen). About 200 ng mRNA was reverse-transcribed into cDNA using the First Stand cDNA Synthesis kit (Invitrogen) and then amplified with the TAQMAN® Universal PCR Master Mix (Applied Biosystems) using the ABI 7900 Sequence Detection System (Applied Biosystems). Unique primer/probe combinations were employed to specifically determine expression of total, surface (*i.e.,* transmembrane) and secreted forms of IgM and IgG1 isotypes (Table 3). Relative expression was normalized to the mouse κ constant region (mCκ).

Table 2

Mouse Genotype	% CD3	% CD19	% IgM[a]	% IgM[b]
C57BL/6	22	63	0	100
Balb/c	11	60	100	0
6hVκ-5hJκ HET	43	30	7	85
16hVκ-5hJκ HET	33	41	7	81

Table 3

Isotype	Sequence (5'-3')	SEQ ID NOs:
Surface IgM	sense: GAGAGGACCG TGGACAAGTC antisense: TGACGGTGGT GCTGTAGAAG probe: ATGCTGAGGA GGAAGGCTTT GAGAACCT	1 2 3
Total IgM	sense: GCTCGTGAGC AACTGAACCT antisense: GCCACTGCAC ACTGATGTC probe: AGTCAGCCAC AGTCACCTGC CTG	4 5 6
Surface IgG1	sense: GCCTGCACAA CCACCATAC antisense: GAGCAGGAAG AGGCTGATGA AG probe: AGAAGAGCCT CTCCCACTCT CCTGG	7 8 9
Total IgG1	sense: CAGCCAGCGG AGAACTACAA G antisense: GCCTCCCAGT TGCTCTTCTG probe: AACACTCAGC CCATCATGGA CACA	10 11 12
Cκ	sense: TGAGCAGCAC CCTCACGTT antisense: GTGGCCTCAC AGGTATAGCT GTT probe: ACCAAGGACG AGTATGAA	13 14 15

[00232] The results from the quantitative TAQMAN® PCR assay demonstrated a decrease in total IgM and total IgG1. However, the ratio of secreted versus surface forms of IgM and IgG1 appeared normal as compared to VELCOIMMUNE® humanized mice (data not shown).

[00233] **Human κ gene segment usage and Vκ-Jκ junction analysis.** Naïve mice homozygous for thirty hVκ and five Jκ gene segments and a replacement of the endogenous Vκ and Jκ gene segments with human Vκ and Jκ gene segments (30hVκ-5hJκ HO) were analyzed for unique human Vκ-Jκ rearrangements on mouse heavy chain (IgG) by reverse transcription polymerase chain reaction (RT-PCR) using RNA isolated from splenocytes.

[00234] Briefly, spleens were harvested and perfused with 10 mL RPMI-1640 (Sigma) with 5% HI-FBS in sterile disposable bags. Each bag containing a single spleen was then placed into a STOMACHER™ (Seward) and homogenized at a medium setting for 30 seconds. Homogenized spleens were filtered using a 0.7μm cell strainer and then pelleted with a centrifuge (1000 rpm for 10 minutes) and RBCs were lysed in BD PHARM LYSE™ (BD Biosciences) for three minutes. Splenocytes were diluted with RPMI-1640 and centrifuged again, followed by resuspension in 1 mL of PBS (Irvine Scientific). RNA was isolated from pelleted splenocytes using standard techniques known in the art.

[00235] RT-PCR was performed on splenocyte RNA using primers specific for human hVκ gene segments and the mouse IgG. The mouse IgG primer was designed such that it was capable of amplifying RNA derived from all mouse IgG isotypes. PCR products were gel-purified and cloned into pCR2.1-TOPO TA vector (Invitrogen) and sequenced with

primers M13 Forward (GTAAAACGAC GGCCAG; SEQ ID NO:16) and M13 Reverse (CAGGAAACAG CTATGAC; SEQ ID NO:17) located within the vector at locations flanking the cloning site. Human Vκ and Jκ gene segment usage among twelve selected clones are shown in Table 4. FIG. 9 sets forth the nucleotide sequence of the hVκ-hJκ-mIgG junction for the twelve selected RT-PCR clones.

[00236] As shown in this Example, mice homozygous for six human Vκ and five human Jκ gene segments or homozygous for thirty human Vκ and five human Jκ gene segments operably linked to the mouse heavy chain constant region demonstrated expression human light chain variable regions from a modified heavy chain locus containing light chain variable gene segments in their germline configuration. Progression through the various stages of B cell development was observed in these mice, indicating multiple productive recombination events involving light chain variable gene segments from an endogenous heavy chain locus and expression of such hybrid heavy chains (*i.e.,* human light chain variable region linked to a heavy chain constant region) as part of the antibody repertoire.

Table 4

Clone	Hybrid Heavy Chain			SEQ ID NO:
	Vκ	Jκ	C_H	
1E	1-5	4	IgG2A/C	18
1G	1-9	4	IgG2A/C	19
1A	1-16	5	IgG3	20
2E	1-12	2	IgG1	21
1C	1-27	4	IgG2A/C	22
2H	2-28	1	IgG1	23
3D	3-11	4	IgG1	24
3A	3-20	4	IgG2A/C	25
4B	4-1	5	IgG2A/C	26
4C	4-1	2	IgG3	27
5A	5-2	2	IgG2A/C	28
5D	5-2	1	IgG1	29

Example IV

Propagation of Mice Expressing V_L Binding Proteins

[00237] To create a new generation of V_L binding proteins, mice bearing the unrearranged human κ gene segments can be bred to another mouse containing a deletion of the other endogenous heavy chain allele. In this manner, the progeny obtained would express only hybrid heavy chains as described in Example I. Breeding is performed by standard techniques recognized in the art and, alternatively, by commercial companies, *e.g.*, The

Jackson Laboratory. Mouse strains bearing a hybrid heavy chain locus are screened for presence of the unique hybrid heavy chains and absence of traditional mouse heavy chains.

[00238] Alternatively, mice bearing the unrearranged human κ gene segments at the mouse heavy chain locus can be optimized by breeding to other mice containing one or more deletions in the mouse light chain loci (κ and λ). In this manner, the progeny obtained would express unique human κ heavy chain only antibodies as described in Example I. Breeding is similarly performed by standard techniques recognized in the art and, alternatively, by commercial companies, *e.g.*, The Jackson Laboratory. Mouse strains bearing a hybrid heavy chain locus and one or more deletions of the mouse light chain loci are screened for presence of the unique hybrid heavy chains containing human κ light chain domains and mouse heavy chain constant domains and absence of endogenous mouse light chains.

[00239] Mice bearing an unrearranged hybrid heavy chain locus are also bred with mice that contain a replacement of the endogenous mouse κ light chain variable gene locus with the human κ light chain variable gene locus (see US 6,596,541, Regeneron Pharmaceuticals, The VELOCIMMUNE® Humanized Mouse Technology). The VELOCIMMUNE® Humanized Mouse includes, in part, having a genome comprising human κ light chain variable regions operably linked to endogenous mouse κ light chain variable constant region loci such that the mouse produces antibodies comprising a human κ light chain variable domain and a mouse heavy chain constant domain in response to antigenic stimulation. The DNA encoding the variable regions of the light chains of the antibodies can be isolated and operably linked to DNA encoding the human light chain constant regions. The DNA can then be expressed in a cell capable of expressing the fully human light chain of the antibody. Upon a suitable breeding schedule, mice bearing a replacement of the endogenous mouse κ light chain with the human κ light chain locus and an unrearranged hybrid heavy chain locus is obtained. Unique V_L binding proteins containing somatically mutated human $V\kappa$ domains can be isolated upon immunization with an antigen of interest.

Example V

Generation of V_L Binding Proteins

[00240] After breeding mice that contain the unrearranged hybrid heavy chain locus to various desired strains containing modifications and deletions of other endogenous Ig loci (as described in Example IV), selected mice can be immunized with an antigen of interest.

[00241] Generally, a VELOCIMMUNE® humanized mouse containing at least one hybrid heavy chain locus is challenged with an antigen, and cells (such as B-cells) are recovered from the animal (*e.g.*, from spleen or lymph nodes). The cells may be fused with a myeloma

cell line to prepare immortal hybridoma cell lines, and such hybridoma cell lines are screened and selected to identify hybridoma cell lines that produce antibodies containing hybrid heavy chains specific to the antigen used for immunization. DNA encoding the human $V\kappa$ regions of the hybrid heavy chains may be isolated and linked to desirable constant regions, *e.g.*, heavy chain and/or light chain. Due to the presence of human $V\kappa$ gene segments fused to the mouse heavy chain constant regions, a unique antibody-like repertoire is produced and the diversity of the immunoglobulin repertoire is dramatically increased as a result of the unique antibody format created. This confers an added level of diversity to the antigen specific repertoire upon immunization. The resulting cloned antibody sequences may be subsequently produced in a cell, such as a CHO cell. Alternatively, DNA encoding the antigen-specific V_L binding proteins or the variable domains may be isolated directly from antigen-specific lymphocytes (*e.g.*, B cells).

[00242] Initially, high affinity V_L binding proteins are isolated having a human $V\kappa$ region and a mouse constant region. As described above, the V_L binding proteins are characterized and selected for desirable characteristics, including affinity, selectivity, epitope, *etc.* The mouse constant regions are replaced with a desired human constant region to generate unique fully human V_L binding proteins containing somatically mutated human $V\kappa$ domains from an unrearranged hybrid heavy chain locus of the invention. Suitable human constant regions include, for example wild type or modified IgG1 or IgG4 or, alternatively $C\kappa$ or $C\lambda$.

[00243] Separate cohorts of mice containing a replacement of the endogenous mouse heavy chain locus with six human $V\kappa$ and five human $J\kappa$ gene segments (as described in Example I) and a replacement of the endogenous $V\kappa$ and $J\kappa$ gene segments with human $V\kappa$ and $J\kappa$ gene segments were immunized with a human cell-surface receptor protein (Antigen X). Antigen X is administered directly onto the hind footpad of mice with six consecutive injections every 3-4 days. Two to three micrograms of Antigen X are mixed with 10 µg of CpG oligonucleotide (Cat # tlrl-modn - ODN1826 oligonucleotide ; InVivogen, San Diego, CA) and 25 µg of Adju-Phos (Aluminum phosphate gel adjuvant, Cat# H-71639-250; Brenntag Biosector, Frederikssund, Denmark) prior to injection. A total of six injections are given prior to the final antigen recall, which is given 3-5 days prior to sacrifice. Bleeds after the 4th and 6th injection are collected and the antibody immune response is monitored by a standard antigen-specific immunoassay.

[00244] When a desired immune response is achieved splenocytes are harvested and fused with mouse myeloma cells to preserve their viability and form hybridoma cell lines. The hybridoma cell lines are screened and selected to identify cell lines that produce Antigen X-specific V_L binding proteins. Using this technique several anti-Antigen X-specific

V_L binding proteins (*i.e.,* binding proteins possessing human Vκ domains in the context of mouse heavy and light chain constant domains) are obtained.

[00245] Alternatively, anti-Antigen X V_L binding proteins are isolated directly from antigen-positive B cells without fusion to myeloma cells, as described in U.S. 2007/0280945A1, herein specifically incorporated by reference in its entirety. Using this method, several fully human anti-Antigen X V_L binding proteins (*i.e.,* antibodies possessing human Vκ domains and human constant domains) were obtained.

[00246] Human κ Gene Segment Usage. To analyze the structure of the anti-Antigen X V_L binding proteins produced, nucleic acids encoding the human Vκ domains (from both the heavy and light chains of the V_L binding protein) were cloned and sequenced using methods adapted from those described in US 2007/0280945A1 (*supra*). From the nucleic acid sequences and predicted amino acid sequences of the antibodies, gene usage was identified for the hybrid heavy chain variable region of selected V_L binding proteins obtained from immunized mice (described above). Table 5 sets for the gene usage of human Vκ and Jκ gene segments from selected anti-Antigen X V_L binding proteins, which demonstrates that mice according to the invention generate antigen-specific V_L binding proteins from a variety of human Vκ and Jκ gene segments, due to a variety of rearrangements at the endogenous heavy chain and κ light chain loci both containing unrearranged human Vκ and Jκ gene segments. Human Vκ gene segments rearranged with a variety of human Jκ segments to yield unique antigen-specific V_L binding proteins.

Table 5

Antibody	Hybrid Heavy Chain		Light Chain	
	Vκ	Jκ	Vκ	Jκ
A	4-1	3	3-20	1
B	4-1	3	3-20	1
C	4-1	3	3-20	1
D	4-1	3	3-20	1
E	4-1	3	3-20	1
F	4-1	3	3-20	1
G	4-1	3	3-20	1
H	4-1	3	3-20	1
I	4-1	3	3-20	1
J	1-5	3	1-33	3
K	4-1	3	3-20	1
L	4-1	3	1-9	3
M	4-1	1	1-33	4
N	4-1	1	1-33	3

O	1-5	1		1-9	2
P	1-5	3		1-16	4
Q	4-1	3		3-20	1
R	4-1	3		3-20	1
S	1-5	1		1-9	2
T	1-5	1		1-9	2
U	5-2	2		1-9	3
V	1-5	2		1-9	2
W	4-1	1		1-33	4

[00247] **Enzyme-linked immunosorbent assay (ELISA).** Human V_L binding proteins raised against Antigen X were tested for their ability to block binding of Antigen X's natural ligand (Ligand Y) in an ELISA assay.

[00248] Briefly, Ligand Y was coated onto 96-well plates at a concentration of 2 µg/mL diluted in PBS and incubated overnight followed by washing four times in PBS with 0.05% Tween-20. The plate was then blocked with PBS (Irvine Scientific, Santa Ana, CA) containing 0.5% (w/v) BSA (Sigma-Aldrich Corp., St. Louis, MO) for one hour at room temperature. In a separate plate, supernatants containing anti-Antigen X V_L binding proteins were diluted 1:10 in buffer. A mock supernatant with the same components of the V_L binding proteins was used as a negative control. The extracellular domain (ECD) of Antigen X was conjugated to the Fc portion of mouse IgG2a (Antigen X-mFc). Antigen X-mFc was added to a final concentration of 0.150 nM and incubated for one hour at room temperature. The V_L binding protein/Antigen X-mFc mixture was then added to the plate containing Ligand Y and incubated for one hour at room temperature. Detection of Antigen X-mFc bound to Ligand Y was determined with Horse-Radish Peroxidase (HRP) conjugated to anti-Penta-His antibody (Qiagen, Valencia, CA) and developed by standard colorimetric response using tetramethylbenzidine (TMB) substrate (BD Biosciences, San Jose, CA) neutralized by sulfuric acid. Absorbance was read at OD450 for 0.1 sec. Background absorbance of a sample without Antigen X was subtracted from all samples. Percent blocking was calculated for >250 (three 96 well plates) Antigen X-specific V_L binding proteins by division of the background-subtracted MFI of each sample by the adjusted negative control value, multiplying by 100 and subtracting the resulting value from 100.

[00249] The results showed that several V_L binding proteins isolated from mice immunized with Antigen X specifically bound the extracellular domain of Antigen X fused to the Fc portion of mouse IgG2a (data not shown).

[00250] **Affinity Determination.** Equilibrium dissociation constants (K_D) for selected Antigen X-specific V_L binding protein supernatants were determined by SPR (Surface Plasmon Resonance) using a BIACORE™ T100 instrument (GE Healthcare). All data were

obtained using HBS-EP (10mM HEPES, 150mM NaCl, 0.3mM EDTA, 0.05% Surfactant P20, pH 7.4) as both the running and sample buffers, at 25°C.

[00251] Briefly, V_L binding proteins were captured from crude supernatant samples on a CM5 sensor chip surface previously derivatized with a high density of anti-human Fc antibodies using standard amine coupling chemistry. During the capture step, supernatants were injected across the anti-human Fc surface at a flow rate of 3 µL/min, for a total of 3 minutes. The capture step was followed by an injection of either running buffer or analyte at a concentration of 100 nM for 2 minutes at a flow rate of 35 µL/min. Dissociation of antigen from the captured V_L binding protein was monitored for 6 minutes. The captured V_L binding protein was removed by a brief injection of 10 mM glycine, pH 1.5. All sensorgrams were double referenced by subtracting sensorgrams from buffer injections from the analyte sensorgrams, thereby removing artifacts caused by dissociation of the V_L binding protein from the capture surface. Binding data for each V_L binding protein was fit to a 1:1 binding model with mass transport using BIAcore T100 Evaluation software v2.1.

[00252] The binding affinities of thirty-four selected V_L binding proteins varied, with all exhibiting a K_D in the nanomolar range (1.5 to 130 nM). Further, about 70% of the selected V_L binding proteins (23 of 34) demonstrated single digit nanomolar affinity. $T^{1/2}$ measurements for these selected V_L binding proteins demonstrated a range of about 0.2 to 66 minutes. Of the thirty-four V_L binding proteins, six showed greater than 3 nM affinity for Antigen X (1.53, 2.23, 2.58, 2.59, 2.79, and 2.84). The affinity data is consistent with the V_L binding proteins resulting from the combinatorial association of rearranged human light chain variable domains linked to heavy and light chain constant regions (described in Table 4) being high-affinity, clonally selected, and somatically mutated. The V_L binding proteins generated by the mice described herein comprise a collection of diverse, high-affinity unique binding proteins that exhibit specificity for one or more epitopes on Antigen X.

[00253] In another experiment, selected human V_L binding proteins raised against Antigen X were tested for their ability to block binding of Antigen X's natural ligand (Ligand Y) to Antigen X in a LUMINEX® bead-based assay (data not shown). The results demonstrated that in addition to specifically binding the extracellular domain of Antigen X with affinities in the nanomolar range (described above), selected V_L binding proteins were also capable of binding Antigen X from cynomolgus monkey (*Macaca fascicularis*).

The claims defining the invention are as follows:

1. A mouse, comprising in its germline an immunoglobulin hybrid chain locus comprising at least one unrearranged light chain variable region (V_L) gene segment, at least one unrearranged light chain joining (J_L) gene segment, and an immunoglobulin heavy chain constant region nucleic acid sequence capable of associating with a light chain constant region,

wherein each of the unrearranged V_L and J_L gene segments comprise recombination signal sequences that allow the unrearranged V_L and J_L gene segments to recombine such that the mouse further comprises a rearranged immunoglobulin hybrid gene comprising light chain variable region (V_L/J_L) nucleotide sequence operably linked with the immunoglobulin heavy chain constant region nucleic acid sequence.

2. The mouse of claim 1, wherein the at least one unrearranged light chain V_L gene segment is selected from an unrearranged human κ gene segment, an unrearranged human λ gene segment, and a combination thereof.

3. The mouse of claim 1, wherein the heavy chain constant region nucleic acid sequence is selected from the group consisting of a C_H1 sequence, a hinge sequence, a C_H2 sequence, a C_H3 sequence, and a combination thereof.

4. The mouse of claim 1, wherein the at least one unrearranged light chain V_L gene segment and the at least one unrearranged J_L gene segment respectively replace an endogenous mouse heavy chain V_H gene segment and an endogenous mouse heavy chain J_H gene segment at the endogenous mouse heavy chain locus.

5. The mouse of claim 4, wherein the at least one unrearranged light chain V_L gene segment replaces all or substantially all functional mouse heavy chain V_H gene segments of the endogenous mouse heavy chain locus.

6. The mouse of claim 4, wherein the at least one unrearranged J_L gene segment replaces all or substantially all functional mouse heavy chain J_L gene segments of the endogenous mouse heavy chain locus.

7. The mouse of claim 6, wherein the at least one unrearranged light chain V_L gene segment and the unrearranged light chain J_L gene segment are operably linked and the mouse lacks a functional heavy chain D_H gene segment between the at least one unrearranged light chain V_L gene segment and the at least one unrearranged light chain J_L gene segment.

8. The mouse of claim 7, wherein the at least one unrearranged light chain V_L gene segment is an unrearranged human κ gene segment and the unrearranged light chain J_L gene segment is an unrearranged human κ gene segment.

9. The mouse of claim 1, comprising a B cell that comprises in its genome the rearranged immunoglobulin hybrid gene, and wherein the rearranged hybrid gene comprises a rearranged human κ variable region operably linked to a mouse constant region gene.

10. The mouse of claim 1, wherein the rearranged immunoglobulin hybrid gene is at an endogenous mouse immunoglobulin heavy chain locus.

11. The mouse of claim 1, further comprising in its germline an unrearranged human light chain V_L gene segment and an unrearranged human light chain joining J_L gene segment operably linked to a light chain constant gene.

12. The mouse of claim 11, wherein the light chain constant gene is a mouse light chain constant gene.

13. The mouse of claim 12, wherein the unrearranged human light chain V_L gene segment is an unrearranged human κ V_L gene segment.

14. The mouse of claim 13, wherein the mouse light chain constant gene is a mouse κ light chain constant gene.

15. A mouse comprising a hybrid immunoglobulin comprising a first polypeptide comprising a first human light chain variable region sequence fused with an immunoglobulin heavy chain constant region, and a second polypeptide comprising a second human light chain variable region fused with an immunoglobulin light chain constant region.

16. The mouse of claim 15, wherein the first human light chain variable region sequence comprises a human κ variable region sequence, and the second human light chain variable region is selected from a human κ variable region and a human λ variable region.

17. The mouse of claim 16, wherein the immunoglobulin heavy chain constant region is selected from a human heavy chain constant region and a mouse heavy chain constant region.

18. The mouse of claim 16, wherein the immunoglobulin light chain constant region is selected from a human light chain constant region and a mouse light chain constant region.

19. The mouse of claim 16, wherein the first polypeptide is expressed from a modified endogenous mouse immunoglobulin heavy chain locus that lacks a functional endogenous heavy chain V gene segment.

20. The mouse of claim 19, wherein the second polypeptide is expressed from a modified endogenous mouse immunoglobulin light chain locus that lacks a functional endogenous light chain V gene segment.

21. A mouse, comprising a replacement in the germline of the mouse at an endogenous mouse immunoglobulin heavy chain locus of all or substantially all functional endogenous mouse heavy chain variable (V_H) gene segments with at least six or more unrearranged human light chain V_L gene segments and one or more unrearranged human light chain J_L gene segments, wherein each of the unrearranged human light chain V_L and J_L gene

segments comprise recombination signal sequences that allow the unrearranged V_L and J_L gene segments to recombine such that the mouse further comprises a rearranged immunoglobulin hybrid gene comprising light chain variable region (V_L/J_L) nucleotide sequence operably linked to an immunoglobulin heavy chain constant region, wherein the mouse is incapable of expressing an immunoglobulin heavy chain derived from a heavy chain V gene segment, and wherein the mouse comprises a splenic B cell population ($B220^+/IgM^+$) that is at least about 75% the size of a splenic B cell population ($B220^+/IgM^+$) of a wild-type mouse.

22. Use of a mouse according to any one of claims 1 to 21 to produce a hybrid immunoglobulin binding protein that comprises a human light chain variable domain operably linked to heavy chain constant region.

23. Use of a mouse according to any one of claims 1 to 21 to produce an antigen binding protein comprising a hybrid immunoglobulin chain and a light chain, wherein the hybrid immunoglobulin chain comprises a light chain variable domain fused with a heavy chain constant region, and wherein the light chain comprises a light chain variable domain fused with a light chain constant region.

24. The use according to claim 23, wherein the antigen binding protein comprises human light chain variable domains.

25. Use of a mouse according to any one of claims 1 to 21 to produce a bispecific antibody.

26. A cell or tissue derived from a mouse according to any one of claims 1 to 21.

27. A cell according to claim 26, wherein the cell is selected from an ES cell, a B cell, and a hybridoma.