

(51) International Patent Classification:
C40B 30/06 (2006.01)(21) International Application Number:
PCT/US2013/031442(22) International Filing Date:
14 March 2013 (14.03.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/616,091 27 March 2012 (27.03.2012) US(71) Applicant: **DANA-FARBER CANCER INSTITUTE, INC.** [US/US]; 450 Brookline Avenue, Boston, MA 02215 (US).(72) Inventors: **RUPRECHT, Ruth, M.**; 127 Winthrop Road, #6, Brookline, MA 02445-4408 (US). **ESSONO, Sosthene**; 32 Clinton St., Malden, MA 02148 (US). **SHOLUKH, Anton**; 396 Langley Rd., Apt. #4, Newton, MA 02459 (US). **VYAS, Hemant**; 15 Harbor Point Blvd., Apt. 207, Dorchester, MA 02125 (US).(74) Agent: **SMITH, Deann, F.**; Foley Hoag LLP, Seaport West, 155 Seaport Blvd., Boston, MA 02210-2600 (US).

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

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

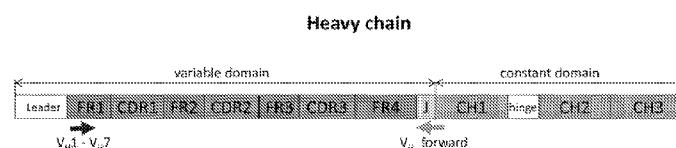
Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: COMPOSITIONS, KITS, AND METHODS FOR THE GENERATION OF IMMUNOGLOBULIN SEQUENCES

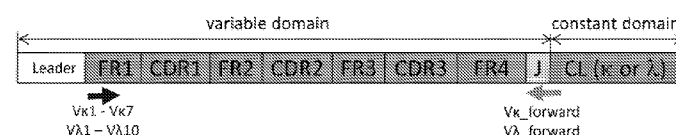
Figure 1

A



B

Light chain

(57) Abstract: The present invention is based, in part, on the discovery of certain oligonucleotide sequences derived from specific regions of the rhesus monkey (*Macaca mulatta*) immunoglobulin gene repertoire sequences useful for amplifying the light chain, heavy chain, and full-length immunoglobulin genes encoding the entire *Macaca mulatta* immunoglobulin gene repertoire. Thus, the present invention allows for the generation of diverse and complex libraries of *Macaca mulatta* immunoglobulin gene sequences encoding full-length immunoglobulins, as well as fragments and variants thereof (e.g., scFvs).

COMPOSITIONS, KITS, AND METHODS FOR THE GENERATION OF IMMUNOGLOBULIN SEQUENCES

Cross-Reference to Related Applications

- 5 This application claims the benefit of U.S. Provisional Application No. 61/616,091, filed on March 27, 2012; the entire content of which application is incorporated herein in its entirety by this reference.

Background of the Invention

- 10 The tremendous diversity and specificity of antibodies generated by mammalian immune systems has led to great biomedical investment toward identifying specific antibodies useful for diagnostic, prognostic, and therapeutic applications. Accordingly, there has been a dramatic increase in the general effort to identify full immunological repertoires from which to select particularly useful antibodies (*e.g.*, hybridoma technology,
15 phage display, and lymphocyte antibody recovery methods). Antibodies have long been viewed as potential agents for therapeutic interventions and targeted drug delivery, largely with a view to exploiting the combination of high specificity and affinity of the antibody-antigen interaction. Hence, the number of antibody variable heavy and variable light chain gene sequences, their organisation into sequence specific families and their relative
20 utilisation is of interest with respect to antibody engineering.

- Unfortunately, current methods for generating immunologically diverse antibody libraries have been hampered by limitations in techniques such as hybridoma generation, which prevent the cloning of full immunological repertoires. The inability to use humans for the production of antibodies hampers scientific interrogation of human immunological
25 diversity. In addition, a highly complex and large number of primers are usually required to amplify a diverse set of mammalian immunoglobulin gene sequences given the tremendous number of such sequences maintained within the host genome and transcriptome.

- In view of the above, it is clear that there remains a need in the art for compositions,
30 kits and methods for identifying and cloning the full repertoire of mammalian immunoglobulins. In addition, there is a need in the art for compositions (*e.g.*, primer sets) that will efficiently recover a diverse and complex set of variable and constant light and heavy chains, ensuring cloning of these regions from single or small numbers of B cells (or

other antibody producing cells). Finally, there is a need in the art to generate diverse and complex libraries of immunoglobulins and/or fragments and variants thereof.

Summary of the Invention

5 The present invention is based, in part, on the discovery of compositions, kits and methods for identifying and cloning light chain, heavy chain, and full-length immunoglobulin genes covering the entire rhesus monkey (*Macaca mulatta*) immunoglobulin gene repertoire.

 In one aspect, a nucleic acid molecule is provided having a nucleotide sequence
10 selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.

 In another aspect, provides a nucleic acid molecule is provided having a nucleotide sequence comprising a sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10, wherein the oligonucleotide is capable of annealing to a *Macaca mulatta* immunoglobulin sequence, or a portion thereof.

15 In still another aspect, a nucleic acid molecule is provided comprising a nucleotide sequence which is at least 80% identical to a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.

 In yet another aspect, a nucleic acid molecule is provided comprising a fragment of at least 8 nucleotides of a nucleic acid molecule having a nucleotide sequence selected from
20 the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.

 In another aspect, a set of primers is provided selected from the group consisting of two or more oligonucleotides of oligonucleotides shown in Tables 1-3 and 8-10.

 In still another aspect, an oligonucleotide is provided for amplifying a nucleic acid sequence in a sample, wherein the sample contains a nucleotide sequence selected from the
25 group consisting of oligonucleotides shown in Tables 1-3 and 8-10, or complement thereof. In one embodiment, the sample is from a rhesus monkey (*e.g.*, splenocytes, lymph nodes, lymphocytes, and the like).

 In yet another aspect, an isolated VL kappa sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby is provided, which is selected from the
30 group of sequences set forth in Tables 4-7.

 In another aspect, an isolated VL lambda sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby is provided, which is selected from the group of sequences set forth in Tables 4-7.

In still another aspect, an isolated VH sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby is provided, which is selected from the group of sequences set forth in Tables 4-7 and 12.

5 In yet another aspect, an isolated CL lambda sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby is provided, which is selected from the group of sequences set forth in Tables 4-7.

In another aspect, an isolated CH lambda sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby is provided, which is selected from the group of sequences set forth in Tables 4-7.

10 In still another aspect, a library of nucleic acid molecules encoding a plurality of antigen binding polypeptides is provided, wherein the library is made by a method comprising the steps of: a.) isolating nucleic acid molecules encoding an immunoglobulin from a rhesus monkey; b.) amplifying the variable regions of the immunoglobulin's heavy chain and the lambda and/or kappa light chains using PCR with a set of primers designed
15 for the variable regions, wherein primers for the VL kappa chains are represented by any one of SEQ ID NOs: 1-18, primers for the VL lambda chains are represented by any one of SEQ ID NOs: 19-38, primers for the VH chains are represented by any one of SEQ ID NOs: 39-50, and/or their analogues or combination, wherein the amplified variable regions incorporate a flexible linker into the 3' end of the V_L amplicons and the 5' end of the V_H
20 amplicons or vice versa; and c.) using the flexible linker, randomly linking the V_H and V_L amplicons. In one embodiment, the library further comprises scFv DNA constructs generated from randomly combined VH and VL chains with each scFv chain constituting a library member. In another embodiment, the library is a library of display packages displaying the scFv chains, wherein a library member comprises a nucleic acid encoding
25 scFv chain, and the scFv chain is displayed from the package. In still another embodiment, the population of nucleic acids are cloned into a DNA construct (*e.g.*, a plasmid, phagemid, or expression cassette). In yet another embodiment, the DNA construct is a phage display vector or bacteriophage or yeast that express individual scFv to form phage or yeast display libraries expressing the generated scFv constructs. In another embodiment, a nucleic acid
30 sequence and amino acid sequence encoded thereby representing said VH chain is selected from the group of sequences set forth in Tables 4-6 and 12 and a nucleic acid sequence and amino acid sequence encoded thereby representing said VL chain is selected from the group of sequences set forth in Tables 4-7. In still another embodiment, a library member further

comprises a nucleic acid segment encoding a tag linked to the nucleic acid encoding the ScFv chain, wherein the tag is the same in different library members. In yet another embodiment, the method further comprises contacting library members with an agent having specific affinity for the tag and isolating a subpopulation of library members that
5 bind to the agent.

In yet another aspect, a nucleic acid sequence encoding an scFv isolated from libraries described herein are provided. In one embodiment, the nucleic acid sequence and amino acid sequence encoded thereby representing said VH chain is selected from the group of sequences set forth in Tables 4-7 and 12. In another embodiment, the nucleic acid
10 sequence and amino acid sequence encoded thereby representing said VL chain is selected from the group of sequences set forth in Tables 4-7. In another aspect, an isolated scFv polypeptide encoded by the nucleic acid is provided.

In another aspect, a kit for the detection of *Macaca mulatta* immunoglobulin sequences is provided, wherein the kit comprises at least one nucleic acid molecule,
15 wherein the nucleic acid molecule comprises a nucleotide sequence which is at least 80% identical to a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10. Similarly, a kit for the detection of *Macaca mulatta* immunoglobulin sequences, comprising a set of oligonucleotides is provided, wherein the oligonucleotides comprise a nucleotide sequence which is at least 80% identical to a
20 nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10. In one embodiment, the kit(s) further comprise at least one labeled oligonucleotide for detecting amplified *Macaca mulatta* immunoglobulin nucleic acid. In another embodiment, the kit(s) further comprise a component selected from the group consisting of a plurality of nucleotides and a nucleic acid polymerase. In still another
25 embodiment, the kit(s) further comprise at least one other component for conducting a polymerase amplification reaction (*e.g.*, a thermostable DNA polymerase). In yet another embodiment, the kit(s) comprise instructions for use.

In still another aspect, a method for amplifying a target nucleic acid is provided, wherein the method comprises: combining a target nucleic acid under conditions which
30 allow for an amplification reaction to occur with: a) one or more nucleic acid primer sequences which are at least 80% identical to the sequences set forth as oligonucleotides shown in Tables 1-3 and 8-10; b) a nucleic acid polymerase ; and c) a plurality of nucleotides, thereby resulting in an amplified target nucleic acid. In one embodiment, the

method further comprises step for generating an scFv, wherein the amplified target nucleic acid is combined under conditions which allow for an amplification reaction to occur with:

a) one or more nucleic acid primer sequences which are at least 80% identical to the sequences set forth as SEQ ID NOs: 51-52; b) a nucleic acid polymerase; and c) a plurality

5 of nucleotides, thereby generating an scFv. In another embodiment, the target nucleic acid is from the *Macaca mulatta* genome or transcriptome (e.g., a *Macaca mulatta* immunoglobulin sequence). In still another embodiment, the polymerase is selected from the group consisting of reverse transcriptase and thermostable DNA polymerase. In yet another embodiment, the amplified *Macaca mulatta* immunoglobulin nucleic acid is
10 sequenced and/or cloned into a vector. In another aspect, the vector is provided.

In yet another aspect, a method for preparing single-chain variable fragment (scFv) nucleic acid molecules is provided, wherein the method comprises the steps of: a.) isolating nucleic acid molecules encoding an immunoglobulin from a rhesus monkey; b.) amplifying the variable regions of the immunoglobulin's heavy chain and the lambda and/or kappa light
15 chains using PCR with a set of primers designed for the variable regions, wherein primers for the VL kappa chains are represented by any one of SEQ ID NOs: 1-18, primers for the VL lambda chains are represented by any one of SEQ ID NOs: 19-38, primers for the VH chains are represented by any one of SEQ ID NOs: 39-50, and/or their analogues or combination, wherein the amplified variable regions incorporate a flexible linker into the 3'
20 end of the V_L amplicons and the 5' end of the V_H amplicons or vice versa; and c.) using the flexible linker, randomly linking the V_H and V_L amplicons. In one embodiment, the isolated nucleic acid molecules are selected from the group consisting of genomic DNA, RNA, and cDNA. In another embodiment, the nucleic acid molecules encoding an immunoglobulin from a rhesus monkey are obtained from rhesus monkey splenocytes or
25 lymphocytes. In still another embodiment, the rhesus monkey is immunized with a specific antigen. In yet another embodiment, the antigen is a bacterial, fungal, parasitic or viral antigen. In another embodiment, the amplified variable regions incorporate enzyme restriction sites into the 5' end of the V_L amplicons and the 3' end of the V_H amplicons. In still another embodiment, the method further comprises the step of cloning the single chain
30 variable fragment into a DNA construct. In yet another embodiment, the DNA construct is a plasmid, phagemid, or expression cassette.

In another aspect, a method for mutagenizing the plurality of scFv described herein is provided, wherein the method comprises a.) mutagenizing genes encoding the individual

scFv members of the library; and b.) expressing the genes to produce mutagenized chimeric scFv. In one embodiment, the method further comprises the step of screening said mutagenized chimeric scFv antibodies to select for a desired structure or function.

In still another aspect, a method of identifying an antigen-specific scFv is provided, wherein the method comprises the step of panning the library against a target antigen, identifying an antigen-specific scFv; and isolating the identified antigen-specific scFv. In one embodiment, panning is performed using phage display.

In yet another aspect, a method for specifically detecting *Macaca mulatta* immunoglobulin nucleic acids in a sample is provided, wherein the method comprises: a) contacting said sample with one or more nucleic acid sequences which are at least 80% identical to the sequences set forth in Tables 1-3 and 8-10, under conditions such that said *Macaca mulatta* immunoglobulin nucleic acids can hybridize with said primers; b) reverse transcribing and amplifying said nucleic acids to obtain amplified *Macaca mulatta* immunoglobulin nucleic acids; and c) detecting the presence of said amplified *Macaca mulatta* immunoglobulin nucleic acids. In one embodiment, detecting the presence of said amplified *Macaca mulatta* immunoglobulin nucleic acids comprises: a) contacting said amplified *Macaca mulatta* immunoglobulin nucleic acids with a labeled oligonucleotide to obtain labeled *Macaca mulatta* immunoglobulin nucleic acids; and b) identifying said labeled nucleic acids. In another embodiment, amplification of said nucleic acids is accomplished by nucleic acid sequence based amplification (NASBA), a polymerase chain reaction (PCR), transcription mediated amplification (TMA) or Ligase chain reaction. In still another embodiment, the method further comprises the step of sequencing the amplified *Macaca mulatta* immunoglobulin nucleic acids. In yet another embodiment, the method further comprises evaluating the sequence for mutations.

In another aspect, the use of an scFv described herein for the manufacture of a medicament for the treatment and/or prophylaxis of a disease involving aberrant expression or aberrant activity of an antigen recognised by said antibody is provided.

In still another aspect, a method of screening for and/or diagnosis or prognosis of a disease in a subject, and/or monitoring the effectiveness of therapy for said disease is provided, wherein the method comprises the step of detecting and/or quantifying in a biological sample obtained from said subject, the expression of an antigen recognised by an scFv described herein.

In any aspect of the present invention, certain modifications, additions, and the like may be made. For example, sequences represented in the form a SEQ ID NO may be written for convenience, but does not include substitution of functionally equivalent sequences described herein (*e.g.*, alternative compatible oligonucleotides described in the
 5 Tables). In one embodiment, a vector is provided wherein the vector comprises a nucleic acid molecule described herein. In another embodiment, a nucleic acid molecule described herein (*e.g.*, an oligonucleotide) further comprises a label. In one embodiment, the label is selected from the group consisting of a fluorescent group, digoxigenin, biotin, radioactive labels, chemiluminescent groups, enzymes, antibodies, luminescent agents, precipitating
 10 agents, and dyes. In another embodiment, the label is selected from the group consisting of a fluorescent group, digoxigenin biotin, radioactive labels, chemiluminescent groups, enzymes) antibodies, luminescent agents, precipitating agents, and dyes.

Brief Description of the Drawings

15 **Figures 1A-1B** show schematic diagrams of sub-group specific primers useful for amplifying the light chain (Figure 1A) and heavy chain (Figure 1B) variable domain immunoglobulin gene sequences from *Macaca mulatta*. FR1 - FR4, framework regions 1 - 4; CDR1-34, complementarity-determining regions 1-3; C, constant regions of the light chains (κ or λ); CH1, constant region 1 of the heavy chain.

20 **Figures 2A-2B** show schematic diagrams of constant region-specific primers useful for amplifying the heavy chain (Figure 2A) and light chain (Figure 2B) constant domain immunoglobulin gene sequences from *Macaca mulatta*. Figure 2A uses the following abbreviations: C γ GSP1, constant region γ gene-specific primer 1 (only one primer for all 4 isotypes); C γ CT, primer used for the amplification of the 3' end of the constant region;
 25 C γ GSP2, C γ gene-specific forward primer used for nested PCR; C γ 1h to C γ 4h, forward primers specific for the hinge regions in the four different *Macaca mulatta* γ heavy chains. Figure 2B uses the following abbreviations: C λ _CT and C κ _CT, primers used for the amplification of the 3' end of the light chain constant region; C λ _GSP1 and C κ _GSP1, gene-specific primer 1 to constant region of κ or λ light chain; C λ _GSP2 and C κ _GSP2,
 30 gene-specific primer 2 to constant region of κ or λ light chain.

Figure 3 shows exemplary amplification results of *Macaca mulatta* VH immunoglobulin genes with primers specific for leader sequences listed in Tables 8-10. The upper panel represents results of VH gene RT-PCR using mRNA isolated from bone

marrow B cells of 6 rhesus monkeys. Reverse transcription was performed with random hexamer primers according to the manufacturer protocol for SuperScript III™ reverse transcriptase (Invitrogen) and PCR was done as described in Example 2. Forward primers were leader-specific primers of Set I, II and III (Tables 8, 9, and 10, respectively) and reverse primer was γ -PCR1 listed in Table 3. The lower panel shows gel analysis of VH gene amplification performed according to the same procedure but using total RNA isolated from PBMC of an individual rhesus monkey. L, 100bp+ DNA ladder; VH, VH primers specific for different VH gene subfamilies as labeled in the Group or Note columns of Tables 8, 9, and 10; M, mixture of all primers of the set; O, positive control; and C, negative control.

Figure 4 shows exemplary amplification results of *Macaca mulatta* VH immunoglobulin genes with primers specific for leader sequences listed in Tables 8-10. Total RNA was isolated from PBMC of two naive rhesus macaques, RGj-11 (upper gel) and RCf-12 (lower gel) by Trizol. Reverse transcription was performed according to manufacturer instruction for SuperScript III™ reverse transcriptase (Invitrogen). PCR was performed under conditions described in Example 2 using primers of set I, II and III (Tables 8, 9, and 10, respectively) as forward primers and γ -PCR1 listed in Table 3 as reverse primer. L, 100bp+ DNA ladder; VH, VH primers specific for different VH gene subfamilies as labeled in the Group or Note columns of Tables 8, 9, and 10; M, mixture of all primers of the set; O, positive control; and C, negative control.

Detailed Description of the Invention

The present invention is based, in part, on the discovery of certain oligonucleotide sequences derived from specific regions of the rhesus monkey (*Macaca mulatta*) immunoglobulin gene repertoire useful for amplifying the light chain, heavy chain, and full-length immunoglobulin genes encoding the entire *Macaca mulatta* immunoglobulin gene repertoire. Thus, the present invention allows for the generation of diverse and complex libraries of *Macaca mulatta* immunoglobulin gene sequences encoding full-length immunoglobulins, as well as fragments and variants thereof (e.g., scFvs).

Accordingly, compositions and methods are provided for amplifying, cloning, and/or isolating the full repertoire of *Macaca mulatta* immunoglobulin genes. In addition, compositions and methods are provided for generating diverse and complex libraries of

Macaca mulatta-based immunoglobulins, as well as fragments and variants thereof (*e.g.*, scFvs).

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. (See *e.g.*, Maniatis, Fitch & Sambrook, *Molecular Cloning; A Laboratory Manual* (1982); *DNA Cloning, Volumes I and II* (D. N Glover ed. 1985); *Oligonucleotide Synthesis* (M. J. Gait ed, 1984); *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); the series, *Methods in Enzymology* (Academic Press, Inc.), particularly Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds. , respectively)). All patents, patent applications, and publications mentioned herein, both *supra* and *infra*, are hereby incorporated herein by reference.

I. Definitions

The articles “a” and “an” are used herein to refer to one or to more than one (*i.e.* to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

The term “about” as used herein means in quantitative terms plus or minus 5%, or in another embodiment plus or minus 10%, or in another embodiment plus or minus 15%, or in another embodiment plus or minus 20%.

The terms “amplification” or “amplify” include the reactions necessary to increase the number of copies of a nucleic acid sequence (*e.g.*, a DNA sequence). For the purposes of this invention, amplification refers to the *in vitro* exponential increase in copy number of a target nucleic acid sequence, such as that mediated by a polymerase amplification reaction such as *e.g.*, PCR, however, other amplification reactions encompassed by the invention include, *e.g.*, RT-PCR (see, *e.g.*, U. S. P. N. 4,683, 202; Mullis et al.), and the ligase chain reaction (Barany, *Proc. Natl. Acad. Sci. USA* 88: 189- 193 (1991)).

Unless otherwise specified here within, the terms “antibody” and “antibodies” and “immunoglobulins” broadly encompass naturally-occurring forms of antibodies (*e.g.* IgG, IgA, IgM, IgE) and recombinant antibodies such as single-chain antibodies, chimeric and humanized antibodies and multi-specific antibodies, as well as fragments and derivatives of all of the foregoing, which fragments and derivatives have at least an antigenic binding site. Antibody derivatives can comprise a protein or chemical moiety conjugated to an antibody. Moreover, “antibody” can refer to endogenously generate immunoglobulins produced by an

organism, such as *Macaca mulatta*, or by engineered variants thereof, such as through enhanced diversity introduced by scientific manipulation according to methods known in the art and described further herein.

The term “body fluid” refers to fluids that are excreted or secreted from the body as well as fluid that are normally not (*e.g.* amniotic fluid, aqueous humor, bile, blood and blood plasma, cerebrospinal fluid, cerumen and earwax, cowper’s fluid or pre-ejaculatory fluid, chyle, chyme, stool, female ejaculate, interstitial fluid, intracellular fluid, lymph, menses, breast milk, mucus, pleural fluid, pus, saliva, sebum, semen, serum, sweat, synovial fluid, tears, urine, vaginal lubrication, vitreous humor, vomit).

As used herein, the term “coding region” refers to regions of a nucleotide sequence comprising codons which are translated into amino acid residues, whereas the term “noncoding region” refers to regions of a nucleotide sequence that are not translated into amino acids (*e.g.*, 5' and 3' untranslated regions).

“Complementary” refers to the broad concept of sequence complementarity between regions of two nucleic acid strands or between two regions of the same nucleic acid strand. It is known that an adenine residue of a first nucleic acid region is capable of forming specific hydrogen bonds (“base pairing”) with a residue of a second nucleic acid region which is antiparallel to the first region if the residue is thymine or uracil. Similarly, it is known that a cytosine residue of a first nucleic acid strand is capable of base pairing with a residue of a second nucleic acid strand which is antiparallel to the first strand if the residue is guanine. A first region of a nucleic acid is complementary to a second region of the same or a different nucleic acid if, when the two regions are arranged in an antiparallel fashion, at least one nucleotide residue of the first region is capable of base pairing with a residue of the second region. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, when the first and second portions are arranged in an antiparallel fashion, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion. More preferably, all nucleotide residues of the first portion are capable of base pairing with nucleotide residues in the second portion.

A molecule is “fixed” or “affixed” to a substrate if it is covalently or non-covalently associated with the substrate such that the substrate can be rinsed with a fluid (*e.g.* standard

saline citrate, pH 7.4) without a substantial fraction of the molecule dissociating from the substrate.

“Homologous” as used herein, refers to nucleotide sequence similarity between two regions of the same nucleic acid strand or between regions of two different nucleic acid strands. When a nucleotide residue position in both regions is occupied by the same nucleotide residue, then the regions are homologous at that position. A first region is homologous to a second region if at least one nucleotide residue position of each region is occupied by the same residue. Homology between two regions is expressed in terms of the proportion of nucleotide residue positions of the two regions that are occupied by the same nucleotide residue. By way of example, a region having the nucleotide sequence 5'-ATTGCC-3' and a region having the nucleotide sequence 5'-TATGGC-3' share 50% homology. Preferably, the first region comprises a first portion and the second region comprises a second portion, whereby, at least about 50%, and preferably at least about 75%, at least about 90%, or at least about 95% of the nucleotide residue positions of each of the portions are occupied by the same nucleotide residue. More preferably, all nucleotide residue positions of each of the portions are occupied by the same nucleotide residue.

As used herein, the term “host cell” is intended to refer to a cell into which a nucleic acid of the invention, such as a recombinant expression vector of the invention, has been introduced. The terms “host cell” and “recombinant host cell” are used interchangeably herein. It should be understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications can occur in succeeding generations due to either mutation or environmental influences, such progeny can not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

The term “humanized antibody”, as used herein, is intended to include antibodies made by a non-human cell having variable and constant regions which have been altered to more closely resemble antibodies that would be made by a human cell. For example, by altering the non-human antibody amino acid sequence to incorporate amino acids found in human germline immunoglobulin sequences. The humanized antibodies of the invention can include amino acid residues not encoded by human germline immunoglobulin sequences (*e.g.*, mutations introduced by random or site-specific mutagenesis *in vitro* or by somatic mutation *in vivo*), for example in the CDRs. The term “humanized antibody”, as used herein, also includes antibodies in which CDR sequences derived from the germline of

another mammalian species, such as a mouse, have been grafted onto human framework sequences.

The terms “hybridize” or “hybridization” are art-known and include the hydrogen bonding of complementary DNA and/or RNA sequences to form a duplex molecule.

5 As used herein, the term “immune cell” refers to cells that play a role in the immune response. Immune cells are of hematopoietic origin, and include lymphocytes, such as B cells and T cells; natural killer cells; myeloid cells, such as monocytes, macrophages, eosinophils, mast cells, basophils, and granulocytes.

The term “immunoglobulin polypeptide” as used herein also includes “antibody,”
10 “antibody fragment,” and “antigen-binding portion” of an antibody (or simply “antibody portion”). The term “antigen-binding portion”, as used herein, refers to one or more fragments of an antibody that retain the ability to specifically bind to an antigen. It is well known that the antigen-binding function of an antibody can be performed by fragments of a full-length antibody. Examples of binding fragments encompassed within the term
15 “antigen-binding portion” of an antibody include (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a F(ab')₂ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the VH and CH1 domains; (iv) a Fv fragment consisting of the VL and VH domains of a single arm of an antibody, (v) a dAb fragment (Ward *et al.*,
20 (1989) *Nature* 341:544-546), which consists of a VH domain; and (vi) an isolated complementarity determining region (CDR). Furthermore, although the two domains of the Fv fragment, VL and VH, are coded for by separate genes, they can be joined, using recombinant methods, by a synthetic linker that enables them to be made as a single protein chain in which the VL and VH regions pair to form monovalent polypeptides (known as
25 single chain Fv (scFv); see *e.g.*, Bird *et al.* (1988) *Science* 242:423-426; and Huston *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:5879-5883; and Osbourn *et al.* 1998, *Nature Biotechnology* 16: 778). Such single chain antibodies are also intended to be encompassed within the term “antigen-binding portion” of an antibody. Any VH and VL sequences of specific scFv can be linked to human immunoglobulin constant region cDNA or genomic
30 sequences, in order to generate expression vectors encoding complete IgG polypeptides or other isotypes. VH and VL can also be used in the generation of Fab, Fv or other fragments of immunoglobulins using either protein chemistry or recombinant DNA technology. Other forms of single chain antibodies, such as diabodies are also encompassed. Diabodies are

bivalent, bispecific antibodies in which VH and VL domains are expressed on a single polypeptide chain, but using a linker that is too short to allow for pairing between the two domains on the same chain, thereby forcing the domains to pair with complementary domains of another chain and creating two antigen binding sites (see *e.g.*, Holliger, P., *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:6444-6448; Poljak, R. J., *et al.* (1994) *Structure* 2:1121-1123).

Still further, an immunoglobulin polypeptide of the present invention or antigen-binding portion thereof can be part of larger immunoadhesion polypeptides, formed by covalent or noncovalent association of the immunoglobulin polypeptide with one or more other proteins or peptides. Examples of such immunoadhesion polypeptides include use of the streptavidin core region to make a tetrameric scFv polypeptide (Kipriyanov, S.M., *et al.* (1995) *Human Antibodies and Hybridomas* 6:93-101) and use of a cysteine residue, a marker peptide and a C-terminal polyhistidine tag to make bivalent and biotinylated scFv polypeptides (Kipriyanov, S.M., *et al.* (1994) *Mol. Immunol.* 31:1047-1058).

Immunoglobulin polypeptide portions, such as Fab and F(ab')₂ fragments, can be prepared from whole immunoglobulin polypeptide using conventional techniques, such as papain or pepsin digestion, respectively, of whole antibodies. Moreover, antibodies, antibody portions and immunoadhesion polypeptides can be obtained using standard recombinant DNA techniques, as described herein.

Immunoglobulin polypeptides can be polyclonal or monoclonal; xenogeneic, allogeneic, or syngeneic; or modified forms thereof (*e.g.* humanized, chimeric, etc.). They can also be fully human. Preferably, antibodies of the invention bind specifically or substantially specifically to target antigen polypeptides or fragments thereof. The terms “monoclonal”, as used herein, refers to a population of immunoglobulin polypeptides that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of an antigen, whereas the term “polyclonal” refers to a population of antibody polypeptides that contain multiple species of antigen binding sites capable of interacting with a particular antigen.

As used herein, the term “inhibit” includes the decrease, limitation, or blockage, of, for example a particular action, function, or interaction.

As used herein, the term “interaction,” when referring to an interaction between two molecules, refers to the physical contact (*e.g.*, binding) of the molecules with one another.

Generally, such an interaction results in an activity (which produces a biological effect) of one or both of said molecules.

An “isolated antibody,” as used herein, is intended to refer to an antibody that is substantially free of other antibodies having different antigenic specificities. Moreover, an
5 isolated antibody can be substantially free of other cellular material and/or chemicals.

As used herein, an “isolated protein” refers to a protein that is substantially free of other proteins, cellular material, separation medium, and culture medium when isolated from cells or produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. An “isolated” or “purified” protein or biologically
10 active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the antibody, polypeptide, peptide or fusion protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language “substantially free of cellular material” includes preparations of *Macaca mulatta* immunoglobulin polypeptide sequences
15 polypeptide or fragment thereof, in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language “substantially free of cellular material” includes preparations of *Macaca mulatta* immunoglobulin protein or fragment thereof, having less than about 30% (by dry weight) of non-*Macaca mulatta* immunoglobulin protein (also referred to herein as a “contaminating
20 protein”), more preferably less than about 20% of non-*Macaca mulatta* immunoglobulin protein, still more preferably less than about 10% of non-*Macaca mulatta* immunoglobulin protein, and most preferably less than about 5% non-*Macaca mulatta* immunoglobulin protein. When antibody, polypeptide, peptide or fusion protein or fragment thereof, *e.g.*, a biologically active fragment thereof, is recombinantly produced, it is also preferably
25 substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The term “kit” is any manufacture (*e.g.*, a package or container) comprising at least one reagent, *e.g.*, a primer, for specifically amplifying and/or sequencing a portion of the
30 *Macaca mulatta* immunoglobulin sequence repertoire, the manufacture being promoted, distributed, or sold as a unit for performing the methods of the present invention. A kit can also include instructions for use.

The term “label” refers to a molecular moiety capable of detection including, by way of example, without limitation, radioactive labels which can be incorporated by known methods (*e.g.*, nick translation or kinasing), radioactive isotopes, biotin, fluorescent groups, chemiluminescent groups (*e.g.*, dioxetanes, particularly triggered dioxetanes), digoxigenin, 5 enzymes, antibodies, luminescent agents, precipitating agents, dyes, and the like.

The term “nucleotides” refers to any nucleotide (including modified nucleotides, *e.g.*, methylated or biotinylated nucleotides) that can be incorporated into a nucleic acid by a polymerase.

The term “nucleic acid” includes DNA molecules (*e.g.*, cDNA or genomic DNA), 10 RNA molecules (*e.g.*, mRNA), and analogs of the DNA or RNA generated using nucleotide analogs or using nucleic acid chemistry. Typical modifications include methylation, biotinylation, and other art-known modifications. In addition, the nucleic acid molecule can be single-stranded or double-stranded.

The term “polymerase” includes any one of, or a mixture of, the nucleotide 15 polymerizing enzymes *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, reverse transcriptase where the template is RNA and the extension product is DNA, or a thermostable DNA polymerase. The term “thermostable DNA polymerase” includes a thermostable DNA polymerase isolated from *Thermus aquaticus*, *Thermus thermophilus*, *Thermus filiformis*, *Thermus flavus*, *Pyrococcus* 20 *furiosus*, *Thermococcus litoralis*, a *Thermotoga* species, or a recombinant form thereof.

The term “probe” refers to a structure comprised of a polynucleotide that is capable of selectively binding or forming a hybrid structure with an intended target sequence, due to complementarity of at least one sequence in the probe with a sequence in the target region. The polynucleotide regions of probes can be composed of DNA and/or RNA and/or 25 synthetic nucleotide analogs. Included within probes are “capture probes,” “blocking probes,” and “label probes.” The term “primer” or “nucleic acid primer” or “nucleic acid primer sequence” includes single-stranded oligonucleotides that, typically, are between about 4 to about 100 bases, or alternatively between about 17 to 30 bases, or alternatively 20 or more bases, and are designed to hybridize with a corresponding template nucleic acid. 30 Primer molecules can be complementary to either the sense or the anti-sense strand of a template nucleic acid and are typically used as complementary pairs that flank a nucleic acid region of interest. Probes can be either synthesized by one skilled in the art, or derived from appropriate biological preparations. For purposes of detection of the target molecule,

probes can be specifically designed to be labeled, as described herein. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA, proteins, antibodies, and organic molecules.

The term “sample” or “biological sample” refers to a sample of tissue or fluid isolated from an individual, including but not limited to, blood, plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, urine, blood cells, tumors, amniotic fluid, organs, genomic DNA, RNA, or cDNA in solution or bound to a substrate, and also samples of *in vitro* cell culture constituents including, but not limited to, conditioned medium resulting from the growth of cells in cell culture medium, putatively virally infected cells, recombinant cells, and cell components, *e.g.*, chromosome(s), organelles, paraffin embedded tissues, or membranes isolated from a cell.

The “sense strand” of a nucleic acid contains the sequence that has sequence homology to that of mRNA. The “anti-sense strand” contains a sequence which is complementary to that of the “sense strand.” The term “target region” refers to a region of the nucleic acid that is to be amplified and/or detected. The term “target sequence” refers to a sequence with which a probe or primer will form a stable hybrid under desired conditions.

The term “subject” refers in one embodiment to an animal in need of therapy for, or susceptible to, a condition or its sequelae. The subject can include dogs, cats, pigs, cows, sheep, goats, horses, rats, mice, monkeys, and humans. The term “subject” does not exclude an individual that is normal in all respects.

The term “target nucleic acid” or “template” includes any nucleic acid intended to be copied in, *e.g.*, a polymerase amplification reaction, such as PCR.

The term “targeting polynucleotide sequence” as used herein, refers to a polynucleotide sequence which is comprised of nucleotides which are complementary to a target nucleotide sequence such that the sequence is of sufficient length and complementarity with the target sequence to form a duplex which has sufficient stability for the purpose intended.

The language “substantially free of chemical precursors or other chemicals” includes preparations of antibody, polypeptide, peptide or fusion protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language “substantially free of chemical precursors or other chemicals” includes preparations of antibody, polypeptide, peptide or

fusion protein having less than about 30% (by dry weight) of chemical precursors or non-antibody, polypeptide, peptide or fusion protein chemicals, more preferably less than about 20% chemical precursors or non-*Macaca mulatta* immunoglobulin polypeptide, peptide or fusion protein chemicals, still more preferably less than about 10% chemical precursors or non-antibody, polypeptide, peptide or fusion protein chemicals, and most preferably less than about 5% chemical precursors or non-antibody, polypeptide, peptide or fusion protein chemicals.

A “transcribed polynucleotide” or “nucleotide transcript” is a polynucleotide (*e.g.* an mRNA, hnRNA, a cDNA, or an analog of such RNA or cDNA) which is complementary to or homologous with all or a portion of a mature mRNA made by transcription of a marker of the invention and normal post-transcriptional processing (*e.g.* splicing), if any, of the RNA transcript, and reverse transcription of the RNA transcript.

As used herein, the term “T cell” includes CD4+ T cells and CD8+ T cells. The term T cell also includes both T helper 1 type T cells and T helper 2 type T cells. The term “antigen presenting cell” includes professional antigen presenting cells (*e.g.*, B lymphocytes, monocytes, dendritic cells, Langerhans cells) as well as other antigen presenting cells (*e.g.*, keratinocytes, endothelial cells, astrocytes, fibroblasts, oligodendrocytes).

As used herein, the term “vector” refers to a nucleic acid capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid”, which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “recombinant expression vectors” or simply “expression vectors”. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, “plasmid” and “vector” can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral

vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

There is a known and definite correspondence between the amino acid sequence of a particular protein and the nucleotide sequences that can code for the protein, as defined by the genetic code (shown below). Likewise, there is a known and definite correspondence between the nucleotide sequence of a particular nucleic acid and the amino acid sequence encoded by that nucleic acid, as defined by the genetic code.

GENETIC CODE

10	Alanine (Ala, A)	GCA, GCC, GCG, GCT
	Arginine (Arg, R)	AGA, ACG, CGA, CGC, CGG, CGT
	Asparagine (Asn, N)	AAC, AAT
	Aspartic acid (Asp, D)	GAC, GAT
	Cysteine (Cys, C)	TGC, TGT
15	Glutamic acid (Glu, E)	GAA, GAG
	Glutamine (Gln, Q)	CAA, CAG
	Glycine (Gly, G)	GGA, GGC, GGG, GGT
	Histidine (His, H)	CAC, CAT
	Isoleucine (Ile, I)	ATA, ATC, ATT
20	Leucine (Leu, L)	CTA, CTC, CTG, CTT, TTA, TTG
	Lysine (Lys, K)	AAA, AAG
	Methionine (Met, M)	ATG
	Phenylalanine (Phe, F)	TTC, TTT
	Proline (Pro, P)	CCA, CCC, CCG, CCT
25	Serine (Ser, S)	AGC, AGT, TCA, TCC, TCG, TCT
	Threonine (Thr, T)	ACA, ACC, ACG, ACT
	Tryptophan (Trp, W)	TGG
	Tyrosine (Tyr, Y)	TAC, TAT
	Valine (Val, V)	GTA, GTC, GTG, GTT
30	Termination signal (end)	TAA, TAG, TGA

An important and well known feature of the genetic code is its redundancy, whereby, for most of the amino acids used to make proteins, more than one coding nucleotide triplet can be employed (illustrated above). Therefore, a number of different

nucleotide sequences can code for a given amino acid sequence. Such nucleotide sequences are considered functionally equivalent since they result in the production of the same amino acid sequence in all organisms (although certain organisms can translate some sequences more efficiently than they do others). Moreover, occasionally, a methylated variant of a
5 purine or pyrimidine can be found in a given nucleotide sequence. Such methylations do not affect the coding relationship between the trinucleotide codon and the corresponding amino acid.

In view of the foregoing, the nucleotide sequence of a DNA or RNA coding for a fusion protein or polypeptide of the invention (or any portion thereof) can be used to derive
10 the fusion protein or polypeptide amino acid sequence, using the genetic code to translate the DNA or RNA into an amino acid sequence. Likewise, for fusion protein or polypeptide amino acid sequence, corresponding nucleotide sequences that can encode the fusion protein or polypeptide can be deduced from the genetic code (which, because of its redundancy, will produce multiple nucleic acid sequences for any given amino acid
15 sequence). Thus, description and/or disclosure herein of a nucleotide sequence which encodes a fusion protein or polypeptide should be considered to also include description and/or disclosure of the amino acid sequence encoded by the nucleotide sequence. Similarly, description and/or disclosure of a fusion protein or polypeptide amino acid sequence herein should be considered to also include description and/or disclosure of all
20 possible nucleotide sequences that can encode the amino acid sequence.

I. Isolated Oligonucleotide Nucleic Acid Molecules

An animal's immunological repertoire must be extremely diverse in order to recognize the billions of potential pathogens to which it can be exposed throughout its
25 lifetime. In order to generate such extreme diversity within the immunoglobulin molecules, the VH and VL chains that together form the antigen-binding site must themselves be highly diverse. The germline DNA that encodes the VH chain of immunoglobulins is comprised of multiple different copies of variable (V), diverse (D) and junctional (J) genes. During B cell development, one variable (V), one diverse (D) and one junctional (J) gene
30 are randomly selected from the germline DNA, recombined and then transcribed to generate an RNA transcript that encodes for an almost unique VH chain. To increase diversity further, random nucleotides are inserted between the V, D and J sequences. Similar events occur in the generation of the VL immunoglobulin chains where V and J genes are randomly selected and rearranged, with random nucleotides inserted between the

V and J genes, to generate VL chains with almost unique amino acid sequences. Since the antigen-binding site of an immunoglobulin is comprised of the combination of one VH and one VL chain, further diversity in antigen specificity results from this pairing. These processes result in a complex and diverse set of immunoglobulin encoding gene sequences.

5 Accordingly, in one aspect, the present invention features novel, isolated oligonucleotide (primer) sequences corresponding directly to or derived from the *Macaca mulatta* genome set forth herein as oligonucleotides shown in Tables 1-3 and 8-10. An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Preferably,
10 an "isolated" nucleic acid molecule is free of sequences (preferably protein-encoding sequences) which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. A nucleic acid molecule of the present invention can be isolated using standard molecular biology techniques and the sequence information in the database records
15 described herein. Using all or a portion of such nucleic acid sequences, nucleic acid molecules of the invention can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, ed., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

 In another embodiment, the present invention features an oligonucleotide selected
20 from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10. In still another embodiment, the oligonucleotides of the invention are at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to the nucleotide sequences set forth in oligonucleotides shown in Tables 1-3 and 8-10. In yet
25 another embodiment, the oligonucleotides of oligonucleotides shown in Tables 1-3 and 8-10 are at least 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80 or more nucleotides in length. It will be appreciated that the 5' end can contain greater variability (*e.g.*, nucleic acid
30 substitutions), yet remain functional (*e.g.*, able to anneal to immunoglobulin gene sequences of the *Macaca mulatta* genome). Moreover, a skilled artisan will appreciate that certain elements contained within some oligonucleotides of oligonucleotides shown in Tables 1-3 and 8-10 (*e.g.*, restriction enzyme sites, cloning sites, overlapping linker sites,

etc.) can readily be altered, as described further below, because they are not used to bind and amplify *Macaca mulatta* immunoglobulin gene sequences *per se*. It will be appreciated that the nucleic acid sequences of the present invention need not consist only of the sequence which is complementary to the targeted *Macaca mulatta* immunoglobulin sequences sequence. Thus, the nucleic acid sequences of the present invention can contain in addition, nucleotide sequences or other moieties which are suitable for the purposes for which the nucleic acid sequences are used. For example, use of the nucleic acid sequences as primers can be used for the amplification of *Macaca mulatta* immunoglobulin sequences via PCR, they can contain sequences which, when in duplex, form restriction enzyme sites which facilitate the cloning of the amplified sequences.

In one embodiment, the invention provides a combination of one or more oligonucleotides of the present invention. In yet another embodiment, the invention provides a set of oligonucleotides, also referred to herein as “primer pairs” and “nucleic acid primer sequences,” selected from the group consisting of two or more of the oligonucleotides of the present invention. In still another embodiment, the invention provides oligonucleotides which are able to amplify an immunoglobulin gene sequence from *Macaca mulatta* having a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10, or complement thereof.

In still another embodiment, the present invention features an oligonucleotide primer set comprising or consisting of:

- (a) the oligonucleotide of SEQ ID NO: 9 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 1-8 or complement(s) thereof, including variant(s) thereof;
- (b) the oligonucleotide of SEQ ID NO: 9 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 10-17 or complement(s) thereof, including variant(s) thereof;
- (c) the oligonucleotide of SEQ ID NO: 18 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 1-8 or complement(s) thereof, including variant(s) thereof; and/or
- (d) the oligonucleotide of SEQ ID NO: 18 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 10-17 or complement(s) thereof, including variant(s) thereof, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2.

The use of any one or more of the primer sets defined in (a)-(d) in an amplification reaction (*e.g.*, PCR-based reaction) enables the isolation of a nucleic acid molecule encoding *Macaca mulatta* immunoglobulin light chain kappa variable (V_{κ}) domain encoding sequences. Moreover, the skilled artisan will readily recognize that primers shown in Tables 1-3 and 8-10 can similarly and accordingly be used to generate an oligonucleotide primer set comprising or consisting of particular primers and that such primer sets can be used in an amplification reaction (*e.g.*, PCR-based reaction) to isolate a nucleic acid molecule encoding *Macaca mulatta* immunoglobulin light chain kappa variable (V_{κ}) domain encoding sequences.

In yet another embodiment, the present invention features an oligonucleotide primer set comprising or consisting of:

(e) the oligonucleotide of SEQ ID NO: 28 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 19-27 or complement(s) thereof, including variant(s) thereof;

(f) the oligonucleotide of SEQ ID NO: 28 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 29-37 or complement(s) thereof, including variant(s) thereof;

(g) the oligonucleotide of SEQ ID NO: 38 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 19-27 or complement(s) thereof, including variant(s) thereof; and/or

(h) the oligonucleotide of SEQ ID NO: 38 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 29-37 or complement(s) thereof, including variant(s) thereof, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2.

The use of any one or more of the primer sets defined in (e)-(h) in an amplification reaction (*e.g.*, PCR-based reaction) enables the isolation of a nucleic acid comprising *Macaca mulatta* immunoglobulin light chain lambda variable (V_{λ}) domain encoding sequences. Moreover, the skilled artisan will readily recognize that primers shown in Tables 1-3 and 8-10 can similarly and accordingly be used to generate an oligonucleotide primer set comprising or consisting of particular primers and that such primer sets can be used in an amplification reaction (*e.g.*, PCR-based reaction) to isolate a nucleic acid molecule encoding *Macaca mulatta* immunoglobulin light chain lambda variable (V_{λ}) domain encoding sequences.

In another embodiment, the present invention features an oligonucleotide primer set comprising or consisting of:

- (i) the oligonucleotide of SEQ ID NO: 44 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 39-43 or complement(s) thereof, including variant(s) thereof;
- (j) the oligonucleotide of SEQ ID NO: 44 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 45-49 or complement(s) thereof, including variant(s) thereof;
- (k) the oligonucleotide of SEQ ID NO: 50 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 39-43 or complement(s) thereof, including variant(s) thereof; and/or
- (l) the oligonucleotide of SEQ ID NO: 50 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 45-49 or complement(s) thereof, including variant(s) thereof, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2.

The use of any one or more of the primer sets defined in (i)-(l) in an amplification reaction (*e.g.*, PCR-based reaction) enables the isolation of a nucleic acid comprising *Macaca mulatta* immunoglobulin heavy chain variable (V_H) domain encoding sequences. Moreover, the skilled artisan will readily recognize that primers shown in Tables 1-3 and 8-10 can similarly and accordingly be used to generate an oligonucleotide primer set comprising or consisting of particular primers and that such primer sets can be used in an amplification reaction (*e.g.*, PCR-based reaction) to isolate a nucleic acid molecule encoding *Macaca mulatta* immunoglobulin heavy chain variable (V_H) domain encoding sequences.

In still another embodiment, the present invention features an oligonucleotide primer set comprising or consisting of: the oligonucleotide of SEQ ID NO: 51 or complement thereof, including variant(s) thereof, and the oligonucleotide of SEQ ID NO: 52 or complement thereof, including variant(s) thereof, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2.

The use of a primer set defined in (m) in an amplification reaction (*e.g.*, PCR-based reaction) enables the isolation of a nucleic acid comprising single chain fragment variable (scFv) encoding sequences. In addition, the use of a primer set defined in (m) in an amplification reaction enables the generation of a library of *Macaca mulatta* scFv

immunoglobulins that is diverse and complex. Moreover, the skilled artisan will readily recognize that primers shown in Tables 1-3 and 8-10 can similarly and accordingly be used to generate an oligonucleotide primer set comprising or consisting of particular primers and that such primer sets can be used in an amplification reaction (*e.g.*, PCR-based reaction) to isolate a nucleic acid molecule comprising single chain fragment variable (scFv) encoding sequences.

In yet another embodiment, the present invention features an oligonucleotide primer set comprising or consisting of:

(n) the oligonucleotide of SEQ ID NO: 53 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 54-55 or complement(s) thereof, including variant(s) thereof, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2.

The use of any one or more of the primer sets defined in (n) in an amplification reaction (*e.g.*, PCR-based reaction) enables the isolation of a nucleic acid comprising *Macaca mulatta* immunoglobulin light chain kappa constant (C_{κ}) domain encoding sequences.

In another embodiment, the present invention features an oligonucleotide primer set comprising or consisting of:

(o) the oligonucleotide of SEQ ID NO: 56 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 57-58 or complement(s) thereof, including variant(s) thereof, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2.

The use of any one or more of the primer sets defined in (o) in an amplification reaction (*e.g.*, PCR-based reaction) enables the isolation of a nucleic acid comprising *Macaca mulatta* immunoglobulin light chain lambda constant (C_{λ}) domain encoding sequences.

In still another embodiment, the present invention features an oligonucleotide primer set comprising or consisting of:

(p) the oligonucleotide of SEQ ID NO: 59 or complement thereof, including variant(s) thereof, and at least one oligonucleotide selected from the group consisting of SEQ ID NOs: 60-65 or complement(s) thereof, including variant(s) thereof, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2.

The use of any one or more of the primer sets defined in (p) in an amplification reaction (*e.g.*, PCR-based reaction) enables the isolation of a nucleic acid comprising *Macaca mulatta* immunoglobulin heavy chain constant (C_γ) domain encoding sequences. Some subsets of the primer sets defined in (p) allow for the isolation of *Macaca mulatta* immunoglobulin heavy chain constant (C_γ) domain encoding sequences spanning all four different *Macaca mulatta* immunoglobulin γ heavy chains. Other subsets of the primer sets defined in (p) allow for the isolation of *Macaca mulatta* immunoglobulin heavy chain constant (C_γ) domain encoding specific immunoglobulin γ heavy chains (*e.g.*, one of the four γ heavy chain types).

In the sequences described herein, the letter “Y” represents the bases C or T; “R” represents A or G; “W” represents A or T; “S” represents C or G; “K” represents T or G; and “M” represents A or C.

It will be appreciated that “at least one primer” is not limiting and can be, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more primers.

Similarly, “two or more primers” is not limiting and can be, for example, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more primers.

It will be understood by one skilled in the art that the additional primer set as required can be a forward or a reverse primer set depending upon whether the primer set provided herein is a forward or a reverse primer set. Such additional primer sets can be designed to anneal to any desired region of a desired *Macaca mulatta* immunoglobulin sequence. It will also be understood that suitable use of primer sets requires that forward and reverse primers are used.

It will also be apparent to one skilled in the art that modifications to the oligonucleotides described herein can be made as desired. For example, the sequence of one or more oligonucleotides can be altered (*e.g.*, shortened, lengthened, and/or modified in sequence). In one embodiment, the restriction sites can be altered or removed altogether. In another embodiment, the linker region between the light chain and heavy chain variable regions in the scFv sequences and oligonucleotides of the present invention determining such linker region(s) described herein can be altered (*e.g.*, shortened, lengthened, and/or modified in sequence). Suitable annealing temperatures for using such primers can be selected as known in the art, for example but without limitation, lowering the annealing temperature can allow less stringent priming and result in a PCR product where previously, at higher temperatures, no product was isolated.

In another embodiment, the oligonucleotides of the present invention comprise a label for detection. Such labels can be, *e.g.*, radioactive labels which can be incorporated by known methods (*e.g.*, nick translation or kinasing), radioactive isotopes, biotin, fluorescent groups, chemiluminescent groups (*e.g.*, dioxetanes, particularly triggered dioxetanes), digoxigenin, enzymes, antibodies, luminescent agents, precipitating agents, dyes, combinations thereof, and the like.

In still another embodiment, the oligonucleotides of the present invention can comprise a promoter-primer, wherein a 5' portion of the sequence includes a promoter sequence.

In another aspect, oligonucleotide combinations of the present invention, for example, including at least one forward and one reverse primer, which together can be used for amplification and/or sequencing of a *Macaca mulatta* immunoglobulin sequence of the present invention, can be suitably packaged in a kit. In one embodiment, nested pairs of amplification and sequencing primers are provided. In another embodiment, the invention provides a kit for the detection of particular subtypes of the V_K , V_L , V_H , C_K , C_L , and C_H classes of *Macaca mulatta* immunoglobulin sequences. In still another embodiment, the kit comprises a set of primers selected from the group consisting of the oligonucleotides of the present invention. The primers in such kits can be labeled or unlabeled. The kit can also include additional reagents such as reagents for performing an amplification (*e.g.*, PCR) reaction, a reverse transcriptase for conversion of RNA to cDNA for amplification, DNA polymerases, dNTP and ddNTP feedstocks. The kit can also include instructions for use.

II. Immunoglobulin-Encoding Nucleic Acids

Immunoglobulin sequences provided using the oligonucleotides and methods of the invention described herein include sequences of functionally active fragments, derivatives or analogues and can be, but are not limited to, polyclonal, monoclonal, bi-, tri- or tetra-valent antibodies, humanised or chimeric antibodies, single chain antibodies, Fab fragments, Fab' and Fab'₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such immunoglobulin variants can be readily engineered by a skilled artisan based upon the *Macaca mulatta* immunoglobulin sequences described herein. Methods of making such immunoglobulin-encoding nucleic acids are further described herein and well-known to the skilled artisan given the oligonucleotide sequences provided herein.

It will be appreciated that the invention provides methods for amplifying a target nucleic acid, *e.g.*, *Macaca mulatta* immunoglobulin sequences, by combining a target nucleic acid under conditions which allow for an amplification reaction to occur with one or more oligonucleotide sequences of the present invention and other necessary amplification agents such as a nucleic acid polymerase and a plurality of nucleotides. In one embodiment, the amplification of a target nucleic acid sequence results in an increased amount of the amplified target nucleic acid. In another embodiment, the target nucleic acid sequence is from the *Macaca mulatta* genome. In still another embodiment, a nucleic acid polymerase is used and can be selected from the group consisting of reverse transcriptase and thermostable DNA polymerase. In other embodiments of the invention, the amplified *Macaca mulatta* immunoglobulin sequences of the present invention are sequenced, evaluated for mutations, cloned into expression vectors, pooled into libraries, and/or expressed as polypeptides. In yet another embodiment of the invention, the amplification of the nucleic acids can be accomplished by nucleic acid sequence-based amplification (NASBA), Transcription Mediated Amplification (TMA), polymerase chain reaction (PCR), other target amplification methods, signal amplification methods or probe amplification methods, such as ligase chain reaction.

A DNA template can be prepared using standard means known in the art, for example by cloning rearranged genomic DNA, or mRNA via cDNA (reviewed by *e.g.* Mountain and Adair, 1992, Biotech. Gen. Eng. Rev. 10:1-142; Ehlich and Kuppers 1995, Curr. Opin. Immunol. 7:281-284). In one embodiment, RNA obtained from a single antibody-producing cell by lysis is used to prepare a DNA template. In one embodiment, the antibody-producing cells can be cultured to increase the cell number. In another embodiment, a single antibody producing cell, *e.g.*, a B cell or a hybridoma cell, can be cultured to increase the cell number before preparation of cDNA from mRNA for use as a template for cloning in order to determine the sequence of, in particular, the variable heavy and light chain region sequences of an immunoglobulin of interest. In another embodiment, multiple antibody producing cells (*e.g.*, splenocytes, lymph nodes, peripheral blood lymphocytes, etc.) are used as a source for deriving template cDNA, which cells can or can not be cultured to increase cell number. In one embodiment, the multiple antibody producing cells are clonal. Alternatively, in other embodiments the cells are not clonal. In another embodiment, the cDNA is derived from splenocytes of a healthy organism (*e.g.*, *Macaca mulatta*) and/or such an organism that has been exposed to a viral, bacterial,

parasite-derived or tumor antigen, a toxin, toxoid, a self-antigen (auto-antigen) administered with an adjuvant that allows breaking of immunological tolerance, an oncogene (*e.g.*, a human oncogene) administered with an adjuvant that allow breaking of immunological tolerance, or their combination. DNA templates prepared from multiple antibody
5 producing cells increases the complexity and diversity of the resulting immunoglobulin libraries generated from the DNA templates. Such complexity and diversity can further be realized by multiplexing PCR primer set(s) according to methods well-known in the art.

It will be appreciated that a sample used in the methods of the present invention can be a biological sample, *e.g.*, from a rhesus monkey. Such samples can include, without
10 limitation, blood, plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, urine, blood cells, tumors, organs, genomic DNA, RNA, or cDNA in solution or bound to a substrate, and also samples of *in vitro* cell culture constituents including, but not limited to, conditioned medium resulting from the growth of cells in cell culture medium, putatively virally
15 infected cells, recombinant cells, and cell components, *e.g.*, chromosome (s), organelles, paraffin embedded tissue, or membranes isolated from a cell.

Reverse transcription and amplification (*e.g.*, PCR) reagents include reagents as known in the art, such as without limitation, DNA polymerase, dNTPs, buffers such as but not limited to Tris buffers, cations such as Mg^{2+} , detergents, for example but not limited, to
20 Nonidet NP-40, reducing agents such as dithiothreitol or mercaptoethanol and RNAasin.

In particular, when cloning an antibody from a single or small numbers of antibody producing cells, it can be desirable to perform the minimum numbers of PCRs necessary. Therefore, in some embodiments, any one of primer set(s) (a) through (p) provided herein and any combinations thereof are utilised in a single PCR. Alternatively, one or more
25 PCRs can be performed utilising a portion of the primers present within a primer set, for example but without limitation, one or two or more primers from any one of primer set(s) described herein (*e.g.*, such as sets (a) through (p) provided herein as non-limiting examples) and any combinations thereof plus additional primers where required, can be utilised in a single PCR.

30 As stated above, it can be desirable to amplify a PCR product to increase specific yield and to add restriction sites to facilitate cloning, *e.g.*, into an expression vector. Thus, the invention further provides oligonucleotide primer sets for performing nested PCR. These primer sets can alternatively be used to perform a first round PCR. In some

embodiments, the primer(s) consist of or comprise at least one or more of the sequences provided herein as SEQ ID NOs: 10-18, 29-38, and/or 45-50, wherein the SEQ ID NOs are listed (*e.g.*, in order from top to bottom) in Tables 1 to 2. These primer sets have the following restriction sites incorporated: NcoI (for SEQ ID NOs: 10-17). In other
5 embodiments, and as described further in the Examples, the primer(s) consist of or comprise at least one or more of the sequences provided herein in Tables 1-3 and 8-10. As stated above, it will be understood by one skilled in the art that any restriction site desired can be incorporated into the primers, for example by replacing the restriction site encoded with coding for the desired restriction site. In another embodiment the upstream restriction
10 site and bases for annealing to the vector can be omitted.

The invention described herein also provides methods for the isolation of *Macaca mulatta* immunoglobulin sequences described herein. Accordingly, provided is a method for the isolation of nucleic acid(s) consisting of or comprising the V_K , V_L , V_H , C_K , C_L , and C_H (such as $C_{\gamma 1H}$, $C_{\gamma 2H}$, $C_{\gamma 3H}$, $C_{\gamma 4H}$) and/or scFv immunoglobulin sequence(s) of *Macaca*
15 *mulatta*, wherein the primer set comprises or consists of primer set(s) described herein, such as those selected from the group consisting of (a) through (p) defined herein as non-limiting examples, or any combination thereof. In some embodiments, any of said primer set(s), such as those defined in (a) through (p) above, can be used in a second PCR, using template derived from a first PCR. Amplified VH and VL chains can further be randomly
20 recombined in one embodiment, using a flexible linker to generate a combinatorial library of scFvs that can be cloned in one embodiment, into a phagemid and expressed on the surface of bacteriophage. The size and diversity of the generated library is evaluated in one embodiment, using cloning and sequencing techniques as will be understood by a skilled artisan.

25 In another aspect, any one of said primer set(s) defined herein, such as those in (a) through (p) above as non-limiting examples, can be used in the methods having a first and a second PCR. For example, in one embodiment, the method comprises two PCRs; a first PCR and a second PCR using as template, DNA from the first PCR. In some embodiments, the PCRs are performed individually. Alternatively, said PCRs are performed jointly. The
30 complexity and diversity of the immunoglobulin sequences to be isolated and/or sequenced can further be realized by multiplexing PCR primer set(s) according to methods well-known in the art. In further embodiments, more than two PCRs can be used, for example, but not limited to three, four, five, six or more PCRs. Thus, one skilled in the art will

appreciate that if, for example, fewer numbers of cycles are used then one or more additional PCRs can be performed to produce the amplification required.

According to such methods and in one aspect, provided herein is an isolated VH sequence selected from the group set forth in Tables 4-7 and 12. In another embodiment, provided herein is an isolated V_K sequence selected from the group set forth in Tables 4-7, an isolated V_L sequence selected from the group set forth in Table 4-7, an isolated V_H sequence selected from the group set forth in Tables 4-7 and 12, an isolated C_K sequence selected from the group set forth in Tables 4-7, an isolated C_L sequence selected from the group set forth in Tables 4-7, an isolated C_H sequence selected from the group set forth in Tables 4-7, and/or an isolated scFv sequence selected from the group set forth in Tables 4-7. These novel immunoglobulin sequences contain *Macaca mulatta* framework regions and complementarity determining regions which can be used to engineer rhesus monkey-like antibodies. In one embodiment, the VH and VL sequences provided herein or homologous sequences are used to arrive at the scFv nucleic acid and polypeptide libraries provided herein. In another embodiment, provided herein is a nucleic acid sequence encoding a single chain fragment variable antibody (scFv) isolated from the scFv nucleic acid library, comprising any combination of a VH and VL chain sequence, wherein in other embodiments, said VH and VL sequences are linked via a flexible linker and/or in any random order, *i.e.*, VH- linker- VL, or VL- linker- VH and different VH and VL sequences thereof.

A nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence, wherein the full length nucleic acid sequence comprises a marker of the invention or which encodes a polypeptide corresponding to a marker of the invention. Such nucleic acid molecules can be used, for example, as a probe or primer. The probe/primer typically is used as one or more substantially purified oligonucleotides. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, preferably about 15, more preferably about 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, or 400 or more consecutive nucleotides of a nucleic acid of the invention.

Probes based on the sequence of a nucleic acid molecule of the invention can be used to detect transcripts or genomic sequences corresponding to one or more markers of the invention. The probe comprises a label group attached thereto, *e.g.*, a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as

part of a diagnostic test kit for identifying cells or tissues which mis-express the protein, such as by measuring levels of a nucleic acid molecule encoding the protein in a sample of cells from a subject, *e.g.*, detecting mRNA levels or determining whether a gene encoding the protein has been mutated or deleted.

5 The invention further encompasses nucleic acid molecules that differ, due to degeneracy of the genetic code, from the nucleotide sequence of nucleic acid molecules encoding a protein which corresponds to a marker of the invention, and thus encode the same protein.

 In another embodiment, an isolated immunoglobulin-encoding nucleic acid
10 molecule of the invention is at least 7, 15, 20, 25, 30, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 550, 650, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000, 3500, 4000, 4500, or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule corresponding to a marker of the invention or to a nucleic acid molecule encoding a protein corresponding to a marker of the invention.
15 As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, 75%, 80%, preferably 85%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in sections 6.3.1-6.3.6 of *Current Protocols in Molecular Biology*, John Wiley &
20 Sons, N.Y. (1989). A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 50-65°C.

 A skilled artisan will appreciate that sequence changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein,
25 without altering the biological activity of the protein encoded thereby. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues
30 that are not conserved or only semi-conserved among homologs of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologs of various species (*e.g.*, murine and human) may be essential for activity and thus would not be likely targets for alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a polypeptide of the invention that contain changes in amino acid residues that are not essential for activity. Such polypeptides differ in amino acid sequence from the naturally-occurring proteins which correspond to the markers of the invention, yet retain
5 biological activity. In one embodiment, such a protein has an amino acid sequence that is at least about 40% identical, 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more identical to the amino acid sequence of one of the proteins which correspond to the markers of the invention.

An isolated nucleic acid molecule encoding a variant protein can be created by
10 introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of nucleic acids of the invention, such that one or more amino acid residue substitutions, additions, or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more
15 predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains
20 (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), non-polar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation
25 mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA
30 libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of polypeptides. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into

replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the
5 frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of *Macaca mulatta* immunoglobulin polypeptide sequences (Arkin and Youvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delagrave *et al.* (1993) *Protein Eng.* 6(3):327-331). In one embodiment, cell based assays can be exploited to analyze a variegated polypeptide library. For example, a library of expression
10 vectors can be transfected into a cell line which ordinarily synthesizes *Macaca mulatta* immunoglobulin polypeptide sequences. The transfected cells are then cultured such that the full length polypeptide and a particular mutant polypeptide are produced and the effect of expression of the mutant on the full length polypeptide activity in cell supernatants can be detected, *e.g.*, by any of a number of functional assays. Plasmid DNA can then be
15 recovered from the cells which score for inhibition, or alternatively, potentiation of full length polypeptide activity, and the individual clones further characterized.

In another embodiment, the methods provided herein enable a skilled artisan to generate an scFv library and screen the library against viral, tumor, or bacterial antigens or any other protein target of interest (such as but not limited to human immunodeficiency
20 virus (HIV) and related SIV antigens) to which the sequences of VH and VL chains generated by the methods provided herein specifically bind to, thus allowing the skilled artisan to subsequently isolate a specific scFv and use it diagnostically, therapeutically and prophylactically *in vivo*. In another embodiment, the isolated scFv, antibody or fragment thereof comprises a VH and VL region provided herein or homologous regions thereof, in
25 any combination.

III. Immunoglobulin Polypeptides

In another aspect, provided herein are isolated immunoglobulin polypeptides. An "isolated" or "purified" protein or biologically active portion thereof is substantially free of
30 cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is

substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of heterologous protein (also referred to herein as a "contaminating protein"). When the protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*,
5 culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, *i.e.*, it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of the protein have less than about 30%, 20%, 10%, 5% (by
10 dry weight) of chemical precursors or compounds other than the polypeptide of interest.

The amino acid sequences disclosed herein will enable those of skill in the art to produce polypeptides corresponding peptide sequences and sequence variants thereof. Such polypeptides can be produced in prokaryotic or eukaryotic host cells by expression of polynucleotides encoding the peptide sequence, frequently as part of a larger polypeptide.
15 Alternatively, such peptides can be synthesized by chemical methods. Methods for expression of heterologous proteins in recombinant hosts, chemical synthesis of polypeptides, and *in vitro* translation are well known in the art and are described further in Maniatis *et al. Molecular Cloning: A Laboratory Manual* (1989), 2nd Ed., Cold Spring Harbor, N.Y.; Berger and Kimmel, *Methods in Enzymology*, Volume 152, Guide to
20 *Molecular Cloning Techniques* (1987), Academic Press, Inc., San Diego, Calif.; Merrifield, J. (1969) *J. Am. Chem. Soc.* 91:501; Chaiken I. M. (1981) *CRC Crit. Rev. Biochem.* 11: 255; Kaiser *et al.* (1989) *Science* 243:187; Merrifield, B. (1986) *Science* 232:342; Kent, S. B. H. (1988) *Annu. Rev. Biochem.* 57:957; and Offord, R. E. (1980) *Semisynthetic Proteins*, Wiley Publishing, which are incorporated herein by reference).

25 Accordingly, in one aspect, the isolated immunoglobulin polypeptide is an isolated VH sequence selected from the group set forth in Tables 4-7 and 12. In another embodiment, provided herein is an isolated V κ L sequence selected from the group set forth in Tables 4-7, an isolated V λ L sequence selected from the group set forth in Tables 4-7, an isolated VH sequence selected from the group set forth in Tables 4-7 and 12, an isolated
30 C κ L sequence selected from the group set forth in Tables 4-7, an isolated C λ L sequence selected from the group set forth in Tables 4-7, an isolated CH sequence selected from the group set forth in Tables 4-7, and/or an isolated scFv sequence selected from the group set forth in Table 4-7 and 12.

It will be understood that biologically active fragments and/or homologs of the immunoglobulin polypeptide sequences of the present invention (*e.g.*, presented in Tables 4-7 and 12) are also part of the present invention.

Biologically active fragments or portions of a polypeptide corresponding to an immunoglobulin polypeptide of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the protein corresponding to the marker (*e.g.*, Tables 4-7 and 12), which include fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the corresponding protein (*e.g.*, binding to a target antigen). A biologically active portion of a protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide of the invention. In another embodiment, the fragment is an N-terminal fragment. In another embodiment, the fragment is a C-terminal fragment. In one embodiment, the fragment of this invention is an intrasequential section of the protein, peptide, or nucleic acid. In another embodiment, the fragment is a functional intrasequential section of the protein, peptide or nucleic acid. In another embodiment, the fragment is a functional intrasequential section within the protein, peptide or nucleic acid. In another embodiment, the fragment is an N-terminal functional fragment. In one embodiment, the fragment is a C-terminal functional fragment.

As used herein, the term “homology,” “homolog” or “homologous” refers to sequence identity, or refers to structural identity, or functional identity. In another embodiment, by using the term “homology” and the other like forms, it is to be understood that any molecule, whether nucleic acid or peptide, that functions similarly, and/or contains sequence identity, and/or is conserved structurally so that it approximates the reference sequence, is to be considered as part of this invention. In another embodiment, the terms “homology”, “homologue” or “homologous”, in any instance, indicate that the sequence referred to, whether an amino acid sequence, or a nucleic acid sequence, exhibits at least 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater correspondence with the indicated sequence. Homologs also encompass molecular species produced by conjugating chemical groups to the amino

residue side chains of the native proteins or fragments thereof, wherein said chemical groups do not form part of the naturally-occurring amino acid residues present in said native proteins. For example, the term is meant to include native polypeptides (either degradation products, synthetically synthesized peptides or recombinant peptides) and
5 peptidomimetics (typically, synthetically synthesized peptides), such as peptoids and semipeptoids which are peptide analogs, which can have, for example, modifications rendering the polypeptides more stable while in a body or more capable of penetrating into cells. Such modifications include, but are not limited to N terminal, C terminal or peptide bond modification, including, but not limited to, backbone modifications, and residue
10 modification, each of which represents an additional embodiment of the invention. Methods for preparing peptidomimetic compounds are well known in the art and are specified, for example, in Quantitative Drug Design, CA. Ramsden Gd., Chapter 17.2, F. Choplin Pergamon Press (1992).

Similarly, in one embodiment, the reference to a correspondence to a particular
15 sequence includes both direct correspondence, as well as homology to that sequence as herein defined. Accordingly and in one embodiment, the term “non-homologous” refers the amino acid sequence or nucleic acid sequence exhibits no more than 70% correspondence with the indicated sequence. In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 65-74% correspondence with the indicated sequence.
20 In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 55-64% correspondence with the indicated sequence. In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 45-54% correspondence with the indicated sequence. In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 35-44% correspondence with the
25 indicated sequence. In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 35-44% correspondence with the indicated sequence. In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 15-34% correspondence with the indicated sequence. In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 5-14% correspondence
30 with the indicated sequence. In another embodiment, the amino acid sequence or nucleic acid sequence exhibits no more than 0.1-4% correspondence with the indicated sequence. In another embodiment, the term “non-homologous” can be used interchangeably with the term “low sequence similarity”.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = # of identical positions/total # of positions (*e.g.*, overlapping positions) x100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *Comput Appl Biosci*, 4:11-7. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Yet another useful algorithm for

identifying regions of local sequence similarity and alignment is the FASTA algorithm as described in Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448. When using the FASTA algorithm for comparing nucleotide or amino acid sequences, a PAM120 weight residue table can, for example, be used with a k -tuple value of 2.

5 The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

 In one aspect, isolated scFv are provided, comprising an isolated heavy (V_H) chain sequence, an isolated light (V_L) chain sequence and a flexible linker. In another
10 embodiment, the V_H and/or V_L chain is selected from a group of nucleic acid sequences and amino acid sequences encoded by those *Macaca mulatta* immunoglobulin sequences amplified using an appropriate primer set(s) (e.g., selected from the group consisting of (a) through (p) as non-limiting examples) described herein, or any combination thereof. In still another embodiment, the term “light chain” (V_L) refers to two distinct types, called kappa
15 (k) or lambda (λ) based on the amino acid sequence of the constant domains. In another embodiment, the term “heavy chain” (V_H) when used in reference to an antibody refers to distinct “classes”), based on the amino acid sequence of the heavy chain constant domain and several of these can be further divided into “subclasses” (isotypes), e.g., IgG γ_1 , IgG γ_2 , IgG γ_3 and IgG γ_4 , etc. The subunit structures and three-dimensional configurations of
20 different classes of immunoglobulins are well known.

 In one embodiment, the term “scFvs” refers to the smallest antibody fragments that maintain the antigen specificity and binding affinity of the whole antibody and are comprised in another embodiment of an immunoglobulin V_L (variable light) and V_H (variable heavy) chain joined by a flexible linker. scFvs are significantly smaller than
25 intact antibodies and can bind to antigens with comparable affinities to the intact antibody molecule. In general, the small size of the scFv chains provided in the libraries described herein that are generated by the methods described herein, endows them with excellent tissue and tumor penetrating properties making them highly attractive as targeting agents for infectious pathogens and tumor cells.

30 In one embodiment, the flexible linker can be used to randomly link the V_H and V_L amplicons in the methods described herein, as shown in Tables 4-7 and 12. In one embodiment, the term “linker” or “flexible linker” refers to any heterologous polypeptide of at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more amino acids in

length, which when inserted between the carboxy-terminal end of V_H and the amino-terminal end of V_L yields a functional protein capable of forming a functional scFV. In another embodiment, the flexible linker is one known in the art including, but not limited, to a serine-glycine linker. The term “flexible linker nucleic acid” refers in another
5 embodiment to a nucleic acid encoding the peptide linker. The term “linker”, “linker sequence”, “spacer”, “tethering sequence” or grammatical equivalents thereof refer to a molecule or group of molecules (such as a monomer or polymer) that connects two molecules and often serves to place the two molecules in a preferred configuration. A number of strategies can be used to covalently link molecules together. These include, but
10 are not limited to, polypeptide linkages between N- and C- terminus of proteins or protein domains, linkage via disulfide bonds, and linkage via chemical cross-linking reagents. In one aspect of this embodiment, the linker is a peptide bond, generated by recombinant techniques or peptide synthesis. In another embodiment the linker is a cysteine linker. In yet another embodiment it is a multi-cysteine linker. Choosing a suitable linker for a
15 specific case where two polypeptide chains are to be connected depends on various parameters, including but not limited to the nature of the two polypeptide chains (*e.g.*, whether they naturally oligomerize), the distance between the N- and the C-termini to be connected if known, and/or the stability of the linker towards proteolysis and oxidation. Furthermore, the linker can contain amino acid residues that provide flexibility. Thus, the
20 linker peptide can predominantly include the following amino acid residues: Gly, Ser, Ala, or Thr. The linker peptide should have a length that is adequate to link two molecules in such a way that they assume the correct conformation relative to one another so that they retain the desired activity. Suitable lengths for this purpose include at least one and not more than 30 amino acid residues. In one embodiment, the linker is from about 1 to 30
25 amino acids in length. In another embodiment, the linker is from about 1 to 15 amino acids in length. In addition, the amino acid residues selected for inclusion in the linker peptide should exhibit properties that do not interfere significantly with the activity of the scFv polypeptide. Thus, the linker peptide on the whole should not exhibit a charge that would be inconsistent with the activity of the polypeptide, or interfere with internal folding, or
30 form bonds or other interactions with amino acid residues in one or more of the monomers that would seriously impede the binding of receptor monomer domains. Useful linkers include glycine-serine polymers, glycine- alanine polymers, alanine- serine polymers, and other flexible linkers such as the tether for the shaker potassium channel, and a large variety

of other flexible linkers, as will be appreciated by those in the art. Suitable linkers can also be identified by screening databases of known three-dimensional structures for naturally occurring motifs that can bridge the gap between two polypeptide chains. In one embodiment, the linker is not immunogenic when administered in a human subject. Thus
5 linkers can be chosen such that they have low immunogenicity or are thought to have low immunogenicity. For example, a linker can be chosen that exists naturally in a human. In another embodiment the linker has the sequence of the hinge region of an antibody, that is the sequence that links the antibody Fab and Fc regions; alternatively the linker has a sequence that comprises part of the hinge region, or a sequence that is substantially similar
10 to the hinge region of an antibody. In another embodiment, the linker is an IgG hinge region such as an IgG1, an IgG3 or a fragment thereof. Another way of obtaining a suitable linker is by optimizing a simple linker, *e.g.*, (Gly4Ser)_n, through random mutagenesis.

The variability in sequence is concentrated in those regions called complementarity determining regions (CDRs) while the more highly conserved regions within the variable
15 domain are called framework regions (FR). The CDRs of the light and heavy chains contain within them the amino acids which are largely responsible for the interaction of the antibody with antigen and the sequences of which determine the specificity of the antibody or antibody fragment.

The rearranged nucleotide sequences of the VH and VL chains are almost unique
20 and contained within each chain are three complementarity determining regions (CDRs). Folding of the protein chains brings these 3 CDRs together to form the antigen-binding site of each chain. In other embodiments, the term "CDR" will comprise regions as described by Kabat et al., J. Biol. Chem. 252, 6609-6616 (1977) and Kabat et al., Sequences of protein of immunological interest. (1991), and Chothia and Lesk, /. Mol. Biol. 196:901-917
25 (1987) and MacCallum et al., /. Mol. Biol. 262:732-745 (1996); each of which is incorporated herein in its entirety by this reference. The amino acids of the CDRs of the variable domains were initially defined by Kabat, based on sequence variability, to consist of amino acid residues 31-35B (H1), 50-65 (H2), and 95-102 (H3) in the human heavy chain variable domain (VH) and amino acid residues 24-34 (L1), 50-56 (L2), and 89-97 (L3) in
30 the human light chain variable domain (VL), using Kabat's numbering system for amino acid residues of an antibody. See Kabat et al., sequences of proteins of immunological interest, US Dept. Health and Human Services, NIH, USA (5th ed. 1991). Surrounding the CDRs are the less diverse framework regions (FR1-FR4) of each chain and these are highly

conserved between mammalian species. In one embodiment, the term “framework region” or “FR” are those variable domain residues other than the hypervariable region residues.

The framework regions for humans and mice have been precisely defined. See, *e.g.*, Kabat, E. A. et al., Sequences of Proteins of Immunological Interest, US Dept. Health and Human Services, National Institutes of Health, USA (5^{sup}.th ed. 1991).

Each variable domain typically has four FRs identified as FR1, FR2, FR3 and FR4. In some embodiments, “FR” also refers to an antibody variable region comprising amino acid residues abutting or proximal to, but outside of the CDR regions, *i.e.*, regions which directly interact with the antigen, acting as the recognition element of the antibody molecule within the variable region of an antibody. In one embodiment, the term “framework region” is intended to mean each domain of the framework that is separated by the CDRs. In some embodiments, the sequences of the framework regions of different light or heavy chains are relatively conserved within a species. The combined heavy and light chain framework regions of an antibody serve to position and align the CDRs for proper binding to the antigen.

In one embodiment, the term “binds” or “binding” or grammatical equivalents, refer to the compositions having affinity for each other. “Specific binding” is where the binding is selective between two molecules. A particular example of specific binding is that which occurs between an antibody and an antigen. Typically, specific binding can be distinguished from non-specific when the dissociation constant (K_D) is less than about 1×10^{-5} M or less than about 1×10^{-6} M or 1×10^{-7} M. Specific binding can be detected, for example, by ELISA, immunoprecipitation, coprecipitation, with or without chemical crosslinking, two-hybrid assays and the like. Appropriate controls can be used to distinguish between “specific” and “non-specific” binding.

In some embodiments, VH and VL chains are randomly combined to generate single chain molecules that recapitulate the naturally occurring antigen-binding site of the original antibody, where in other embodiments, the synthesized single chain molecules are isolated based on the antigens they recognize and are used therapeutically in a number of targeted therapeutic approaches.

In another embodiment, an isolated antibody, antibody fragment or scFv provided herein finds use in various *in vivo* and *in vitro* applications such as, but not limited to diagnostics, antibody imaging, ameliorating symptoms associated with a disease, preventing and treating diseases treatable by antibody-based therapy.

It will also be understood that the *Macaca mulatta* immunoglobulin polypeptides (e.g., scFv) of the present invention provided herein can be conjugated to a therapeutic moiety to a therapeutic agent, such as a cytotoxic agent, a radionuclide or drug moiety to modify a given biological response. The therapeutic agent is not to be construed as limited to classical chemical therapeutic agents. For example, the therapeutic agent can be a drug moiety which can be a protein or polypeptide possessing a desired biological activity. Such moieties can include, for example and without limitation, antiviral peptides, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin, cantansinoid (DMI), a protein such as tumour necrosis factor, α -interferon, β -interferon, nerve growth factor, platelet derived growth factor or tissue plasminogen activator, a thrombotic agent or an anti-angiogenic agent, e.g. angiostatin or endostatin; angiogenin, gelonin, dolstatins, minor groove-binders, bis-iodo-phenol mustard, or, a biological response modifier such as a lymphokine, interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), nerve growth factor (NGF) or other growth factor.

Therapeutic agents also include cytotoxins or cytotoxic agents including any agent that is detrimental to (e.g. kills) cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vinca alkaloids, e.g. vincristine, vinblastine, 4-desacetylvinblastine-3-carbohydrazide, vindesine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents also include, but are not limited to, anti-folates (e.g. aminopterin and methotrexate), antimetabolites (e.g. methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine, 5-fluoro-2'-deoxyuridine), alkylating agents (e.g. mechlorethamine, thioepa chlorambucil, melphalan, caraiustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g. daunorubicin (formerly daunomycin) and doxorubicin, adriamycin, idarubicin, morpholinodoxorubicin, epirubicin, doxorubicin hydrazides), antibiotics (e.g. dactinomycin (formerly actinomycin), bleomycin, mithramycin, anthramycin (AMC), calicheamicins or duocarmycins, CC-1065, enediyenes, neocarzinostatin), and anti-mitotic agents (e.g. vincristine and vinblastine). See Garnett, 2001, Advanced drug Delivery Reviews 53:171 -216 for further details.

Other therapeutic moieties can include radionuclides such as ^{131}I , ^{111}In and ^{90}Y , Lu^{177} , Bismuth 213 , Bismuth 212 , Californium 252 , Iridium 192 and Tungsten 187 and Rhenium 188 , ^{211}At astatine; or drugs such as but not limited to, alkylphosphocholines, topoisomerase I inhibitors, taxoids and suramin. Techniques for conjugating such therapeutic agents to antibodies are well known in the art (see, *e.g.* Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld et al., eds., 1985 pp. 243-56, ed. Alan R. Liss, Inc; Hellstrom et al., "Antibodies For Drug Delivery", in *Controlled Drug Delivery*, 2nd Ed., Robinson et al., eds., 1987, pp. 623- 53, Marcel Dekker, Inc.; Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*; Pinchera et al., 1985, eds., pp. 475-506; "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin et al. (eds.), 1985, pp. 303-16, Academic Press; Thorpe et al., 1982 "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.*, 62:119-58 and Dubowchik et al., 1999, *Pharmacology and Therapeutics*, 83, 67-123).

In some embodiments, the *Macaca mulatta* immunoglobulin polypeptides (*e.g.*, scFv) of the present invention provided herein can comprise a fusion protein wherein the immunoglobulin polypeptides are operatively linked to an effector or reporter molecule, as can be prepared by standard chemical or recombinant DNA procedures. For example, an effector group is a polymer molecule, which is useful for increasing immunoglobulin polypeptide half-life *in vivo*.

IV. Libraries

In another aspect, the present invention provides libraries of antibodies or antibody variable domains and/or antibody constant domains. In another embodiment, provided herein are scFv libraries comprising the VH and VL sequences linked by a flexible linker for the identification and targeting of antigens. In another embodiment, expression of scFv libraries on the surface of bacteriophage or yeast (a scFv phage display or yeast library as the ones described herein) allows for the rapid screening of millions of scFvs and the selection of those scFvs that bind a particular antigen.

Antigens bound by scFvs are identified in one embodiment, and isolated in another embodiment, using standard biochemical techniques with the novel primer sets described

herein. In one embodiment, the resulting phage or yeast display is used successfully to discover novel therapeutic targets and isolate antibody fragments that bind them.

In some embodiments, the library is a nucleic acid library, a phage display library, a yeast display library, or an oligopeptide library. In some embodiments, the process yields
5 an ScFv fragment library, a FR library, a VH library, a VL library, a VH and VL library, a CDR library or an Fab fragment library. In another embodiment, the libraries/methods of the present invention arrive at a *Macaca mulatta* scFv library, used to identify and target certain antigens.

In another embodiment, the libraries described herein, further comprise scFv DNA
10 constructs generated from randomly combined VH and VL chains with each scFv chain constituting a library member. In another embodiment, the population of nucleic acids sequences of the scFv DNA constructs generated from randomly combined VH and VL chains are cloned into multiple copies of a phage display vector or bacteriophage that express individual scFv to form phage display libraries expressing the generated scFv
15 constructs.

It will be understood that the *Macaca mulatta* immunoglobulin polypeptides (*e.g.*, scFv) of the present invention provided herein can be humanized. Humanized immunoglobulin polypeptides originate from non-human species having one or more complementarity determining regions (CDRs) from the non-human species and a
20 framework region from a human immunoglobulin molecule (see, *e.g.* US 5,585,089). These include immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.* molecules that contain an antigen binding site that specifically binds an antigen. The immunoglobulin molecules of the invention can be of any class (*e.g.*, IgG, IgE, IgM, IgD and IgA) or subclass of immunoglobulin molecule.

25 In some embodiments, rhesus monkeys (*Macaca mulatta*) can be used as vaccine recipients against any antigen to generate desired antibody (Ab) responses and then monoclonal antibodies (mAbs) can be isolated therefrom using standard molecular cloning approaches. Rhesus monkeys are much more closely related to humans than other species used routinely to generate mAbs (*e.g.*, mice, rabbits, goats, donkeys, and the like). The
30 rhesus monkey Ab repertoire is also closely related to that of humans, and thus rhesus monkey mAbs better reflect human Ab responses. In the experimental primate model, immunizations can be performed using either antigens and/or adjuvants that cannot be used in humans due to ethical considerations. Such vaccinations in RMs can also include

breaking tolerance by immunizing RMs with an oncogene in an appropriate adjuvant. The resulting mAbs – stemming from a closely related species – may have favorable profiles when used either prophylactically or therapeutically in humans.

5 IV. Recombinant Expression Vectors and Host Cells

Methods known in the art for isolating cells producing antibody of interest include hybridoma technology (Kohler, G. & Milstein C, 1975, Nature 256:495-497), the selected lymphocyte antibody method (Babcook et al, 1996, Proc. Natl. Acad. Sci, 93, 7843-7848; WO 92/02551; de Wildt et al, 1997, J. Immunol. Methods, 207:61-65 and in Lagerkvist, et al, 1995, BioTechniques 18(5):862-869) and isolation of antibodies from bacterially expressed libraries, *e.g.* phage display libraries. These methods rely on the identification of a pool of cells producing an antibody of interest followed by the isolation of individual antibody producing cells which are then clonally expanded and, if desired, the subsequent identification of the sequence of their V_H and V_L chain genes. Alternatively, where more than one cell is present in a sample positive for the antibody of interest, these cells can be cultured together and used as a source of, *e.g.* template cDNA produced from mRNA by reverse transcription, for the isolation of the nucleic acid encoding an antibody.

In another aspect, the present invention also provides phage expression vectors. The term “filamentous phage” or “filamentous bacteriophage” refers to a viral particle capable of displaying a heterogenous polypeptide on its surface. Although one skilled in the art will appreciate that a variety of bacteriophage can be employed in the present invention, in preferred embodiments the vector is, or is derived from, a filamentous bacteriophage, such as, for example, f1, fd, Pfl, M13, etc. The filamentous phage can contain a selectable marker such as, but not limited to, tetracycline (*e.g.*, “fd-tet”). Various filamentous phage display systems are well known to those of skill in the art (see, *e.g.*, Zacher et al. (1980) Gene 9: 127-140, Smith et al. (1985) Science 228: 1315-1317 (1985); and Parmley and Smith (1988) Gene 73: 305-318).

In another aspect, the present invention also provides transformed cells and progeny thereof into which a nucleic acid molecule encoding an antibody, antibody libraries, scFv libraries, antibody fragment, V_H or V_L libraries has been introduced by means of recombinant DNA techniques *in vitro*, *ex vivo* or *in vivo*. The present invention further provides expression vectors useful for transforming such cells. The transformed cells can be propagated and the introduced nucleic acid transcribed, or encoded protein expressed. It is understood that a progeny cell can not be identical to the parental cell, since there can be

mutations that occur during replication. Transformed cells include but are not limited to prokaryotic and eukaryotic cells such as bacteria, fungi, plant, insect, and animal (*e.g.*, mammalian, including monkey) cells. The cells can be present in culture, in a cell, tissue or organ *ex vivo* or present in a subject.

- 5 In one embodiment, the term “transformed” refers to a genetic change in a cell following incorporation of nucleic acid (*e.g.*, a transgene) exogenous to the cell. Thus, a “transformed cell” is a cell into which, or a progeny of which a nucleic acid molecule has been introduced by means of recombinant DNA techniques. Cell transformation to produce host cells can be carried out as described herein or using techniques known in the art.
- 10 Accordingly, methods of producing cells containing the nucleic acids and cells expressing the immunoglobulins of the invention are also provided.

- Typically, cell transformation employs a vector. The term “vector,” refers to, *e.g.*, a plasmid, virus, such as a viral vector, or other vehicle known in the art that can be manipulated by insertion or incorporation of a nucleic acid, for genetic manipulation (*i.e.*,
- 15 “cloning vectors”), or can be used to transcribe or translate the inserted polynucleotide (*i.e.*, “expression vectors”). Such vectors are useful for introducing nucleic acids, including a nucleic acid that encodes a polypeptide provided herein operably linked with an expression control element, and expressing the encoded protein *in vitro* (*e.g.*, in solution or in solid phase), in cells or *in vivo*.

- 20 A great variety of vector and/or expression systems can be used for cloning the single chain variable fragment into a DNA construct. Such systems include, among others, chromosomal, episomal and virus-derived systems, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova
- 25 viruses, such as SV40, vaccinia, viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

- In another embodiment, the expression systems can contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to
- 30 maintain, propagate or express polynucleotides to produce a polypeptide in a host can be used. The appropriate nucleotide sequence can be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., *Molecular Cloning, A Laboratory Manual*.

In one embodiment, provided herein is an expression system as described herein, comprising the scFv obtained by the methods described herein. In another embodiment, the step of randomly linking the flexible linker-comprising V_H and V_L amplicons used to generate the libraries described herein is done using splicing by overlap extension (SOE).

5 Introduction of nucleic acid encoding polypeptides provided herein can also be carried out by conventional methods known in the art such as osmotic shock (*e.g.*, calcium phosphate), electroporation, microinjection, cell fusion, etc. Introduction of nucleic acid and polypeptide *in vitro*, *ex vivo* and *in vivo* can also be accomplished using other techniques. For example, a polymeric substance, such as polyesters, polyamine acids,
10 hydrogel, polyvinyl pyrrolidone, ethylene-vinylacetate, methylcellulose, carboxymethylcellulose, protamine sulfate, or lactide/glycolide copolymers, polylactide/glycolide copolymers, or ethylenevinylacetate copolymers. A nucleic acid can be entrapped in microcapsules prepared by coacervation techniques or by interfacial polymerization, for example, by the use of hydroxymethylcellulose or gelatin-
15 microcapsules, or poly (methylmethacrolate) microcapsules, respectively, or in a colloid drug delivery system. Colloidal dispersion systems include macromolecule complexes, nano-capsules, microspheres, beads, and lipid- based systems, including oil-in-water emulsions, micelles, mixed micelles, and liposomes.

The use of liposomes for introducing various compositions into cells, including
20 nucleic acids, is known to those skilled in the art (see, *e.g.*, U.S. Pat. Nos. 4,844,904, 5,000,959, 4,863,740, and 4,975,282). A carrier comprising a natural polymer, or a derivative or a hydrolysate of a natural polymer, described in WO 94/20078 and U.S. Pat. No. 6,096,291, is suitable for mucosal delivery of molecules, such as polypeptides and polynucleotides. Piperazine based amphiphilic cationic lipids useful for gene therapy also are
25 known (see, *e.g.*, U.S. Pat. No. 5,861,397). Cationic lipid systems also are known (see, *e.g.*, U.S. Pat. No. 5,459,127). Accordingly and in one embodiment, viral and non-viral vector means of delivery into cells or tissue, *in vitro*, *in vivo* and *ex vivo* are included. These methods can be employed in the event that a scFv is identified that can inhibit a specific intracellular pathway and as such would be required to be delivered into the cell(s) of
30 interest. In another embodiment, the present invention comprises methods of use of a polynucleotide, vector, polypeptide and/or fragment thereof as herein described and/or compositions comprising the same in treating, inhibiting or preventing a pathologic condition or disease.

In another embodiment, DNA constructs comprise regulatory elements necessary for expression of nucleotides. Such elements include, for example, a promoter, an initiation codon, a stop codon, and a polyadenylation signal. In addition, enhancers are often required for expression of a sequence that encodes an immunogenic target protein. As is
5 known in the art, these elements are preferably operably linked to the sequence that encodes the desired protein. In another embodiment, regulatory elements are selected that are operable in the species to which they are to be administered. Initiation codons and stop codons are included in one embodiment, as part of a nucleotide sequence that encodes the scFv protein. In one embodiment, the initiation and termination codons must be in frame
10 with the coding sequence.

In one embodiment, viral, eukaryotic and prokaryotic promoters are known in the art and are included for use in the methods and compositions provided herein. In another embodiment these promoter sequences regulate expression of the encoded polynucleotide sequences, and in some embodiments of the present invention, are operably linked to
15 polynucleotides encoding the polypeptides of this invention. In additional embodiments of the present invention, these promoters are either constitutive or inducible, and provide a means of high and low levels of expression of the polypeptides of this invention, and in some embodiments, for regulated expression of multiple polypeptides of the invention, which in some embodiments are expressed as a fusion protein.

The promoters will typically control expression, optionally with an operator
20 sequence and can include ribosome binding site sequences for example, for initiating and completing transcription and translation. According to additional embodiments, the vector can also contain expression control sequences, enhancers that can regulate the transcriptional activity of the promoter, appropriate restriction sites to facilitate cloning of
25 inserts adjacent to the promoter and other necessary information processing sites, such as RNA splice sites, polyadenylation sites and transcription termination sequences as well as any other sequence which can facilitate the expression of the inserted nucleic acid.

IV. Uses and Methods of the Invention

30 The *Macaca mulatta* immunoglobulin molecules (*i.e.*, oligonucleotide molecules having the oligonucleotide sequence or immunoglobulin nucleic acid, polypeptide, polypeptide homolog, and/or fragment thereof molecules having the corresponding sequence of said molecule), described herein can be used in one or more of the following

methods: a) screening assays; and b) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, and monitoring clinical trials).

The isolated nucleic acid molecules of the invention can be used, for example, to detect and isolate cognate antigens.

5

A. Screening Assays

In one aspect, the invention relates to a method for preventing in a subject, a disease or condition associated with an unwanted or less than desirable immune response. Subjects at risk for a disease that would benefit from treatment with the claimed agents or methods
10 can be identified, for example, by any or a combination of diagnostic or prognostic assays known in the art and described herein.

In one embodiment, the libraries described herein, are used in the methods provided. Accordingly, in one embodiment, provided herein is a method of identifying antigen-specific *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv),
15 comprising the step of panning the libraries described herein against the antigen, identifying *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv) that specifically bind the antigen; and isolating the identified antigen-specific *Macaca mulatta* immunoglobulin polypeptides (*e.g.*, scFv). In one embodiment, the method further comprises the step of increasing affinity between the isolated antigen-specific *Macaca*
20 *mulatta* immunoglobulin polypeptides and the antigen (*e.g.*, as performed by site-specific mutagenesis and the like so as to generate diversity).

In one embodiment, molecular techniques are used to recapitulate the immune repertoire of rhesus monkey subjects with diseases such as, but not limited to, viral infections and cancers and screen the resulting combinatorial antibody libraries generated
25 using the methods described herein, for *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv) that specifically target antigens or cells expressing antigens that characterize these diseases, for example, cell-surface HIV antigens. In another embodiment, these *Macaca mulatta* immunoglobulin polypeptides are isolated, amplified and screened *in vitro* for their ability to specifically recognize, neutralize, and/or kill cells
30 expressing such antigens. In another embodiment, *Macaca mulatta* immunoglobulin polypeptides are screened for their ability to inhibit the replication of cell-free virus or the transcytosis of cell-free across epithelial cells.

In one embodiment, such antibody libraries are screened using one or more cell-based or *in vitro* assays. For such assays, *Macaca mulatta* immunoglobulin polypeptides of

the present invention (*e.g.*, scFv), purified or unpurified, are typically added exogenously such that cells are exposed to individual variants or groups of variants belonging to a library. These assays are typically, but not always, based on the biology of the ability of the immunoglobulin polypeptide to bind to antigen and mediate some biochemical event, for example effector functions like cellular lysis, phagocytosis, ligand/receptor binding inhibition, inhibition of growth and/or proliferation, apoptosis, etc. Such assays often involve monitoring the response of cells to antibody, for example cell survival, cell death, cellular phagocytosis, cell lysis, change in cellular morphology, or transcriptional activation such as cellular expression of a natural gene or reporter gene.

For some assays additional cells or components, that is in addition to the target cells, can need to be added, for example serum complement, or effector cells such as peripheral blood monocytes (PBMCs), NK cells, macrophages, and the like. Such additional cells can be from any organism, *e.g.*, humans, mice, rats, rabbits, monkeys, etc. Crosslinked or monomeric antibodies can cause apoptosis of certain cell lines expressing the antibody's target antigen, or they can mediate attack on target cells by immune cells which have been added to the assay. Methods for monitoring cell death or viability are known in the art, and include the use of dyes, fluorophores, immunochemical, cytochemical, and radioactive reagents. For example, caspase assays or annexin-fluorconjugates can enable apoptosis to be measured, and uptake or release of radioactive substrates (*e.g.* Chromium-51 release assays) or the metabolic reduction of fluorescent dyes such as alamar blue can enable cell growth, proliferation, or activation to be monitored. In one embodiment, the DELFIA[®] EuTDA-based cytotoxicity assay (Perkin Elmer, MA) is used. Alternatively, dead or damaged target cells can be monitored by measuring the release of one or more natural intracellular proteins, for example lactate dehydrogenase. Transcriptional activation can also serve as a method for assaying function in cell-based assays. In this case, response can be monitored by assaying for natural genes or proteins which can be upregulated or down-regulated, for example the release of certain interleukins can be measured, or alternatively readout can be via a luciferase or GFP- reporter construct. Cell-based assays can also involve the measure of morphological changes of cells as a response to the presence of an antibody. Cell types for such assays can be prokaryotic or eukaryotic, and a variety of cell lines that are known in the art can be employed. Alternatively, cell-based screens are performed using cells that have been transformed or transfected with nucleic acids encoding the antibodies.

In some embodiments, the screening of populations of *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv), such as the altered variable region populations produced by the methods of the invention, involve immobilization of the populations of altered variable regions to filters or other solid substrate. This is particularly advantageous because large numbers of different species can be efficiently screened for antigen binding. Such filter lifts will allow for the identification of altered variable regions that exhibit substantially the same or greater binding affinity. Alternatively, if the populations of altered variable regions are expressed on the surface of a cell or bacteriophage, for example, panning on immobilized antigen can be used to efficiently screen for the relative binding affinity of species within the population.

Another affinity method for screening populations of altered variable regions polypeptides is a capture lift assay that is useful for identifying a binding molecule having selective affinity for a ligand (Watkins *et. al.*, (1997)). This method employs the selective immobilization of altered variable regions to a solid support and then screening of the selectively immobilized altered variable regions for selective binding interactions against the cognate antigen or binding partner. Selective immobilization functions to increase the sensitivity of the binding interaction being measured since initial immobilization of a population of altered variable regions onto a solid support reduces non-specific binding interactions with irrelevant molecules or contaminants which can be present in the reaction.

Another method for screening populations or for measuring the affinity of individual *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv) is through surface plasmon resonance (SPR). This method is based on the phenomenon which occurs when surface plasmon waves are excited at a metal/liquid interface. Light is directed at, and reflected from, the side of the surface not in contact with sample, and SPR causes a reduction in the reflected light intensity at a specific combination of angle and wavelength. Biomolecular binding events cause changes in the refractive index at the surface layer, which are detected as changes in the SPR signal. The binding event can be either binding association or disassociation between a receptor-ligand pair. The changes in refractive index can be measured essentially instantaneously and therefore allows for determination of the individual components of an affinity constant. More specifically, the method enables accurate measurements of association rates (k_{on}) and disassociation rates (k_{off}). Methods for measuring the affinity, including association and disassociation rates using surface plasmon resonance are well known in the arts and can be

found described in, for example, Jonsson and Malmquist, *Advances in Biosensors*, 2:291-336 (1992) and Wu et al. *Proc. Natl. Acad. Sci. USA*, 95:6037-6042 (1998). Moreover, one apparatus well known in the art for measuring binding interactions is a BIAcore 2000 instrument which is commercially available through Pharmacia Biosensor, (Uppsala, Sweden).

In one embodiment, the methods described herein are used to remove *Macaca mulatta* immunoglobulin polypeptides of the present invention (e.g., scFv) which do not exhibit the desired affinity from the library, to arrive at “optimized” libraries of the present invention, or assemble the immunoglobulin polypeptides based only on the desired characteristics using molecular biology techniques available in the art and as described herein.

In one embodiment, a final immunoglobulin polypeptide, or fragment such as an scFv is generated by the process described herein and is then affinity-purified or isolated after expression. Proteins can be isolated or affinity-purified in a variety of ways known to those skilled in the art. Standard purification methods include chromatographic techniques, electrophoretic, immunological, precipitation, dialysis, filtration, concentration, and chromatofocusing techniques. As is well known in the art, a variety of natural proteins bind antibodies, for example bacterial proteins A, G, and L, and these proteins can find use in the present invention for purification. Purification can often be enabled by a particular fusion partner. For example, proteins can be purified using glutathione resin if a GST fusion is employed, Ni⁺² affinity chromatography if a His-tag is employed or immobilized anti-flag antibody if a flag-tag is used. For general guidance in suitable purification techniques, see *Protein Purification: Principles and Practice*, 3rd Ed., Scopes, Springer- Verlag, N. Y., 1994.

The biophysical properties of immunoglobulin polypeptides, for example stability and solubility, can be screened using a variety of methods known in the art. Protein stability can be determined by measuring the thermodynamic equilibrium between folded and unfolded states. For example, antibodies of the present invention can be unfolded using chemical denaturant, heat, or pH, and this transition can be monitored using methods including but not limited to circular dichroism spectroscopy, fluorescence spectroscopy, absorbance spectroscopy, NMR spectroscopy, calorimetry, and proteolysis. As will be appreciated by those skilled in the art, the kinetic parameters of the folding and unfolding transitions can also be monitored using these and other techniques. The solubility and overall structural integrity of an antibody can be quantitatively or qualitatively determined

using a wide range of methods that are known in the art. Methods which can find use in the present invention for characterizing the biophysical properties of antibodies and antibody fragments include gel electrophoresis, isoelectric focusing, capillary electrophoresis, chromatography such as size exclusion chromatography, ion-exchange chromatography, and reversed-phase high performance liquid chromatography, peptide mapping, oligosaccharide mapping, mass spectrometry, ultraviolet absorbance spectroscopy, fluorescence spectroscopy, circular dichroism spectroscopy, isothermal titration calorimetry, differential scanning calorimetry, analytical ultra- centrifugation, dynamic light scattering, proteolysis, and cross-linking, turbidity measurement, filter retardation assays, immunological assays, fluorescent dye binding assays, protein-staining assays, microscopy, and detection of aggregates via ELISA or other binding assay. Structural analysis employing X-ray crystallographic techniques and NMR spectroscopy can also find use. In one embodiment, stability and/or solubility can be measured by determining the amount of protein solution after some defined period of time. In this assay, the protein can or can not be exposed to some extreme condition, for example elevated temperature, low pH, or the presence of denaturant. Because function typically requires a stable, soluble, and/or well-folded/structured protein, the aforementioned functional and binding assays also provide ways to perform such a measurement. For example, a solution comprising an antibody could be assayed for its ability to bind target antigen, then exposed to elevated temperature for one or more defined periods of time, then assayed for antigen binding again. Because unfolded and aggregated protein is not expected to be capable of binding antigen, the amount of activity remaining provides a measure of the antibody's stability and solubility.

B. Detection Assays

Portions or fragments of the nucleic acid molecules identified herein can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with immunological function and/or (ii) allow identification of an individual or sample from a minute

For example, once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of a rearranged immunoglobulin gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the *Macaca mulatta* immunoglobulin polypeptide sequences

nucleotide sequences, described herein, can be used to map the location of, for example, rearranged *Macaca mulatta* immunoglobulin genes on a chromosome. The mapping of the *Macaca mulatta* immunoglobulin polypeptide sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

5 Briefly, *Macaca mulatta* immunoglobulin genes can be mapped to chromosomes by various methods such as PCR using the oligonucleotides of the invention, somatic cell hybrid techniques, and the like (see, for example, D'Eustachio, P. et al. (1983) Science 220:919-924). Other mapping strategies which can similarly be used to map a *Macaca mulatta* immunoglobulin polypeptide sequences to its chromosome include in situ
10 hybridization (described in Fan, Y. et al. (1990) Proc. Natl. Acad. Sci. USA 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence in situ hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one
15 step (see Verma et al., Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988). Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding
20 sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data (such data are found, for example, in McKusick, V., Mendelian Inheritance in Man, available on-
25 line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. et al. (1987) Nature 325:783-787.

The *Macaca mulatta* immunoglobulin molecules of the present invention can also
30 be used to identify individuals and/or samples from minute biological samples so as to, for example, track the immunoglobulin production of rhesus monkeys or other hosts vaccinated with particular antigens.

The *Macaca mulatta* immunoglobulin molecules described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an in situ hybridization technique, to identify a specific tissue, *e.g.*, lymphocytes. This can be very useful in cases where a forensic pathologist is presented
5 with a tissue of unknown origin. Panels of such *Macaca mulatta* immunoglobulin polypeptide sequences probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, *Macaca mulatta* immunoglobulin primers or probes can be used to screen tissue culture for contamination (*i.e.*, screen for the
10 presence of a mixture of different types of cells in a culture).

In another embodiment, the nucleic acid sequences of the invention can be used as probes to detect the presence of *Macaca mulatta* immunoglobulin sequences (for example in screening for particular such sequences in a biological sample), the biological sample to be analyzed, such as serum or splenic preparations, can be treated, if desired, to extract the
15 nucleic acids contained therein. The resulting nucleic acid from the sample can be subjected to gel electrophoresis or other size separation techniques; alternatively, the nucleic acid sample can be dot blotted without size separation. In order to form hybrid duplexes with the targeting sequence of the probe, the targeted region of the nucleic acid must be in single stranded form. Where the sequence is naturally present in single stranded
20 form, denaturation will not be required. However, where the sequence is present in double stranded form, the sequence will be denatured. Denaturation can be carried out by various techniques known in the art. Subsequent to denaturation, the analyte nucleic acid and probe are incubated under conditions that promote stable hybrid formation of the target sequence in the probe with the putative targeted sequence in the analyte, and the resulting duplexes
25 containing the probe(s) are detected.

Detection of the resulting duplex, if any, is usually accomplished by the use of labeled probes; alternatively, the probe can be labeled, but can be detectable by specific binding with a ligand which is labeled, either directly or indirectly. Suitable labels, and methods for labeling probes and ligands are known in the art, and include, for example,
30 radioactive labels which can be incorporated by known methods (*e.g.*, nick translation or kinasing), radioactive isotopes, biotin, fluorescent groups, chemiluminescent groups (*e.g.*, dioxetanes, particularly triggered dioxetanes), digoxigenin, enzymes, antibodies, luminescent agents, precipitating agents, dyes, and the like. Such combinations can be

useful for RACE, where the source cDNA can be homopolynucleotide tailed at the 3'-end using terminal transferase enzyme, providing a location for strand synthesis using a complementary primer.

5 C. Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring of clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically.

Accordingly, one aspect of the present invention relates to diagnostic assays for detecting
10 the presence of specific antigens using *Macaca mulatta* immunoglobulin molecules in the context of a biological sample (*e.g.*, blood, serum, cells, or tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted antigen expression and/or activity. The invention also provides for prognostic (or predictive) assays for determining whether an
15 individual is at risk of developing a disorder associated with an antigen detectable using *Macaca mulatta* immunoglobulin sequences.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of antigens using *Macaca mulatta* sequences in clinical trials.

20 These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

In one embodiment, provided herein is a method of diagnosing the presence of an antigen, wherein the method comprises in another embodiment, the step of contacting a biological sample with a *Macaca mulatta* immunoglobulin polypeptide of the present
25 invention (*e.g.*, scFv) that binds to the antigen such that the presence of the antigen is detected in the biological sample.

In one embodiment, the method is an *in vitro* diagnosis method. For example, the polypeptide may recognize a target surface antigen of an infected cell or a tumor or abnormal tissue growth.

30 In another embodiment, the method is an *in vivo* diagnosis method. *In vivo* techniques for detection of a target antigen include introducing into a subject a labeled anti-*Macaca mulatta* immunoglobulin polypeptide. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. In one embodiment, the biological sample contains

polypeptide molecules from the test subject. Any antigen of interest can be used for immunisation of an animal, including mammals, primates, monkeys, and humans. Such antigens include any substance that can be recognised by an antibody, including proteins, glycoproteins and carbohydrates. Antigens also include biologically active proteins, such as hormones, cytokines, and their cell surface receptors, bacterial or parasitic cell membranes or purified components thereof, and viral antigens. Antigens can also include cells. Such cells can express endogenous proteins as antigens, for example, presented on the cell surface. Such cells can also transiently or stably express antigens. Antigens of particular interest are those involved in diseases and which show aberrant expression or aberrant activity associated with a disease. Such diseases include cancers, inflammatory disorders and immune disorders.

In another embodiment, the method of diagnosing the presence of an antigen comprises locating said bound *Macaca mulatta* immunoglobulin polypeptide of the present invention (*e.g.*, scFv) in said tissue sample. In another embodiment, the diagnosing the presence of an antigen comprises determining whether said location of said bound antibody indicates the presence of an antigen in said biological sample.

In another embodiment, provided herein is a method of diagnosing a viral infection in a subject, comprising the step of obtaining a biological sample from the subject. In another embodiment, the method of diagnosing a viral infection in a subject comprises the step of contacting the biological sample with a *Macaca mulatta* immunoglobulin polypeptide of the present invention (*e.g.*, scFv). In another embodiment, the method of diagnosing a viral infection in a subject comprises the step of analyzing the level of a viral antigen in the biological sample. In another embodiment, the method of diagnosing a viral infection in a subject comprises the step of and comparing the expression of the biological sample to a standard, whereby and in another embodiment, if the standard is taken from a healthy subject or pool of subjects and the level of the viral antigen is different than the standard by a predetermined threshold, the subject has, or is at risk of developing a disease associated with the viral infection. Otherwise and in another embodiment, if the standard is taken from a subject or pool of subjects correctly diagnosed with a viral infection and the level of the viral antigen is different than the standard by more than a predetermined threshold, the subject does not have or is at low risk of developing a disease associated with the viral infection. Alternatively, the methods just described can be adapted for diagnosing

any condition in which an antigen is diagnostic of the condition, such as a bacterial infection, tumor, etc.

Proteins from cells can be isolated using techniques that are well known to those of skill in the art. The protein isolation methods employed can, for example, be such as those
5 described in Harlow and Lane (Harlow and Lane, 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

In one embodiment, *Macaca mulatta* immunoglobulin polypeptide of the present invention (*e.g.*, scFv) can be used in methods such as Western blots or immunofluorescence techniques to detect target antigen. In such uses, it is generally preferable to immobilize
10 either the immunoglobulin polypeptides or antigen on a solid support. Suitable solid phase supports or carriers include any support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite.

One skilled in the art will know many other suitable carriers for binding the immunoglobulin polypeptide or antigen, and will be able to adapt such support for use with the present invention. For example, protein isolated from cells can be run on a polyacrylamide gel electrophoresis and immobilized onto a solid phase support such as nitrocellulose. The support can then be washed with suitable buffers followed by treatment
20 with the detectably labeled antibody. The solid phase support can then be washed with the buffer a second time to remove unbound antibody. The amount of bound label on the solid support can then be detected by conventional means. Means of detecting proteins using electrophoretic techniques are well known to those of skill in the art (see generally, R. Scopes (1982) *Protein Purification*, Springer-Verlag, N.Y.; Deutscher, (1990) *Methods in*
25 *Enzymology* Vol. 182: *Guide to Protein Purification*, Academic Press, Inc., N.Y.).

In another embodiment, Western blot (immunoblot) analysis is used to detect and quantify the presence of a polypeptide in the sample. This technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter,
30 a nylon filter, or derivatized nylon filter), and incubating the sample with the *Macaca mulatta* immunoglobulin polypeptide of the present invention (*e.g.*, scFv) that specifically bind a target antigen. The anti-antigen immunoglobulin polypeptides specifically bind to the antigen on the solid support. These anti-antigen immunoglobulin polypeptides may be

directly labeled or alternatively may be subsequently detected using labeled immunoglobulin polypeptides (*e.g.*, labeled sheep anti-human antibodies) that specifically bind to the antigen.

In a more preferred embodiment, the antigen is detected using an immunoassay. As used herein, an immunoassay is an assay that utilizes an immunoglobulin polypeptide to specifically bind to the analyte. The immunoassay is thus characterized by detection of specific binding of an antigen to an immunoglobulin polypeptide as opposed to the use of other physical or chemical properties to isolate, target, and quantify the analyte.

The antigen is detected and/or quantified using any of a number of well recognized immunological binding assays (see, *e.g.*, U.S. Pat. Nos. 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also Asai (1993) *Methods in Cell Biology* Volume 37: *Antibodies in Cell Biology*, Academic Press, Inc. New York; Stites & Terr (1991) *Basic and Clinical Immunology* 7th Edition.

Immunological binding assays (or immunoassays) typically utilize a "capture agent" to specifically bind to and often immobilize the analyte (polypeptide or subsequence). The capture agent is a moiety that specifically binds to the analyte. In a preferred embodiment, the capture agent is an immunoglobulin polypeptide that specifically binds a target antigen.

Immunoassays also often utilize a labeling agent to specifically bind to and label the binding complex formed by the capture agent and the analyte. The labeling agent may itself be one of the moieties comprising the antibody/analyte complex. Thus, the labeling agent may be a labeled polypeptide or a labeled anti-antibody. Alternatively, the labeling agent may be a third moiety, such as another antibody, that specifically binds to the antibody/polypeptide complex.

In one preferred embodiment, the labeling agent is a second antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second can be modified with a detectable moiety, *e.g.* as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin.

Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins are normal constituents of the cell walls of streptococcal bacteria. They exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (see, generally Kronval, *et al.* (1973) *J. Immunol.*, 111: 1401-1406, and Akerstrom (1985)

J. Immunol., 135: 2589-2542).

As indicated above, immunoassays for the detection and/or quantification of a polypeptide can take a wide variety of formats well known to those of skill in the art.

Preferred immunoassays for detecting a polypeptide are either competitive or
5 noncompetitive. Noncompetitive immunoassays are assays in which the amount of captured analyte is directly measured. In one preferred "sandwich" assay, for example, the capture agent (anti-peptide antibodies) can be bound directly to a solid substrate where they are immobilized. These immobilized antibodies then capture polypeptide present in the test sample. The polypeptide thus immobilized is then bound by a labeling agent, such as a
10 second human antibody bearing a label.

In competitive assays, the amount of analyte (polypeptide) present in the sample is measured indirectly by measuring the amount of an added (exogenous) analyte (polypeptide) displaced (or competed away) from a capture agent (anti-peptide antibody) by the analyte present in the sample. In one competitive assay, a known amount of, in this
15 case, a polypeptide is added to the sample and the sample is then contacted with a capture agent. The amount of polypeptide bound to the antibody is inversely proportional to the concentration of polypeptide present in the sample.

In one particularly preferred embodiment, the antibody is immobilized on a solid substrate. The amount of polypeptide bound to the antibody may be determined either by
20 measuring the amount of polypeptide present in a polypeptide/antibody complex, or alternatively by measuring the amount of remaining uncomplexed polypeptide. The amount of polypeptide may be detected by providing a labeled polypeptide.

The assays of this invention are scored (as positive or negative or quantity of polypeptide) according to standard methods well known to those of skill in the art. The
25 particular method of scoring will depend on the assay format and choice of label. For example, a Western Blot assay can be scored by visualizing the colored product produced by the enzymatic label. A clearly visible colored band or spot at the correct molecular weight is scored as a positive result, while the absence of a clearly visible spot or band is scored as a negative. The intensity of the band or spot can provide a quantitative measure
30 of polypeptide.

In one embodiment, a "predetermined threshold" refers to a level, range, or measurement empirically determined. In another embodiment, the threshold depends upon the particular population of subjects. In yet another embodiment, an apparently healthy

population will have a different “normal” range or level of the standard than will a population of subjects which have had a prior infection or other condition. In some embodiments, the predetermined thresholds can be delineated from the “normal” range according to significant up- or down-modulation of the standard according to levels of
5 significance defined herein.

As an alternative to making determinations based on the absolute expression level of the marker, determinations may be based on the normalized expression level of the marker. Expression levels are normalized by correcting the absolute expression level of a marker by comparing its expression to the expression of a gene that is not a marker, *e.g.*, a
10 housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene, or epithelial cell-specific genes. This normalization allows the comparison of the expression level in one sample, *e.g.*, a subject sample, to another sample, *e.g.*, a non-cancerous sample, or between samples from different sources.

15 Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a marker, the level of expression of the marker is determined for 10 or more samples of normal versus cancer cell isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The mean expression level of each of the genes assayed in the larger number of samples is
20 determined and this is used as a baseline expression level for the marker. The expression level of the marker determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that marker. This provides a relative expression level.

Preferably, the samples used in the baseline determination will be from disease cells
25 or normal cells of the same tissue type. The choice of the cell source is dependent on the use of the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether the marker assayed is specific to the tissue from which the cell was derived (*versus* normal cells). In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values
30 based on accumulated data. Expression data from normal cells provides a means for grading the severity of the disease.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted expression of a target antigen using *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv) provided herein. Aberrant antigen expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. As used herein, the term “unwanted” includes an unwanted phenomenon involved in a biological response such as immune cell activation. For example, the term unwanted includes a target antigen having expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation in a target antigen, such as a viral infection like HIV, an autoimmune disorder, an immunodeficiency disorder, an immune system cancer, etc., using *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv) provided herein. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted target antigen expression or activity in which a test sample is obtained from a subject and a target antigen is detected using said *Macaca mulatta* immunoglobulin polypeptides, wherein the presence of the target antigen is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted antigen expression or activity. As used herein, a “test sample” refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, cerebrospinal fluid or serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, polypeptide, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted *Macaca mulatta* immunoglobulin polypeptide sequences expression or activity. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted target antigen expression or activity in which a test sample is obtained and the target antigen expression or activity is detected using *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*,

wherein the abundance of target antigen is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted *Macaca mulatta* immunoglobulin polypeptide sequences expression or activity). In some embodiments, the *Macaca mulatta* immunoglobulin polypeptides of the present invention not only detect the antigen, but can neutralize and/or destroy the antigen.

The methods described herein can be performed, for example, by utilizing pre-packaged diagnostic kits described herein, which can be conveniently used, *e.g.*, in clinical settings to diagnose subjects exhibiting symptoms or family history of a disease or illness involving a target antigen of interest.

3. Therapies

In one embodiment, provided herein is a method of delivering a biologically active agent to cells displaying a target antigen, for example to treat a bacterial infection, a viral infection, autoimmunity, a tumor, etc., comprising contacting said cells with a *Macaca mulatta* immunoglobulin polypeptide of the present invention (*e.g.*, scFv), wherein in another embodiment, said *Macaca mulatta* immunoglobulin polypeptide of the present invention (*e.g.*, scFv) is operably linked to said agent.

In one embodiment, provided herein is a method of delivering a biologically active agent and a single chain fragment variable antibody isolated from the nucleic acid library encoding a *Macaca mulatta* immunoglobulin polypeptide of the present invention (*e.g.*, scFv), whereby the biologically active agent and the single chain fragment variable antibody are delivered concomitantly but individually.

In one embodiment, the *Macaca mulatta* immunoglobulin polypeptides of the present invention (*e.g.*, scFv) provided herein are “biologically active”, meaning that they are able to binding to the target antigen, for example, to inhibiting binding of the target to antigens and/or prophylaxis or therapy of antigen-mediated diseases. The term “biologically active”, when used in reference to any of the biologically active agents described herein also refers to the agent's ability to modulate the immune response in a manner that can lead to a preventive, diagnostic, or therapeutic effect as will be understood by a skilled artisan. In another embodiment, the biologically active agent provided herein is a radioisotope, a toxin, a cytokine, a chemokine or any other molecule that can modulate the immune response in a manner that can lead to a preventive, diagnostic, or therapeutic effect as will be understood by a skilled artisan.

In one embodiment, the term “treatment” in the compositions and methods provided herein refers to therapeutic treatment. In another embodiment it refers to prophylactic, or suppressive measures for a disease or disorder. Thus, for example, successful administration of *Macaca mulatta* immunoglobulin molecules of the present invention (*e.g.*,
5 scFv or nucleic acids encoding them) prior to onset of the disease results in treatment of the disease. “Treatment” also encompasses administration of an isolated such immunoglobulin molecule after the appearance of the disease in order to eradicate the disease. Successful administration of an agent, such as an scFv polypeptide provided herein, after onset and after clinical symptoms have developed, with possible abatement of clinical symptoms and
10 perhaps amelioration of the disease, comprises treatment of the disease. In another embodiment, a skilled artisan would understand that treatment does not necessarily result in the complete absence or removal of symptoms. Treatment also embraces palliative effects: that is, those that reduce the likelihood of a subsequent medical condition. In another embodiment, those “in need of treatment” include animals already having the disease or
15 disorder, as well as those prone to having the disease or disorder, including those in which the disease or disorder is to be prevented. In another embodiment, a subject or animal is successfully “treated” for a condition such as infection or cancer if, after receiving a therapeutic amount of an modified molecule provided herein, the subject shows observable and/or measurable reduction in or absence of one or more of the following: reduction in the
20 number of diseased cells or absence of diseased cells related to the condition; reduction in the tumor size; inhibition (*i.e.*, slow to some extent and preferably stop) of cancer cell infiltration into peripheral organs including the spread of cancer into soft tissue and bone; inhibition (*i.e.*, slow to some extent and preferably stop) of tumor metastasis; inhibition, to some extent, of tumor growth; and/or relief to some extent, one or more of the symptoms
25 associated with the specific cancer; reduced morbidity and mortality, and improvement in quality of life.

In another embodiment, provided herein is a method of preventing formation of a condition characterized by aberrant expression or activity of a target antigen in a subject, comprising the step of administering an effective amount of a *Macaca mulatta*
30 immunoglobulin polypeptide of the present invention (*e.g.*, scFv) provided herein. In another embodiment, provided herein is a method of treating, preventing, or ameliorating the symptoms associated with such a condition in a subject comprising the step of

administering to said subject an effective amount of a *Macaca mulatta* immunoglobulin polypeptide of the present invention (e.g., scFv) provided herein.

Accordingly, *Macaca mulatta* immunoglobulin polypeptides of the present invention (e.g., scFv) can find use in a wide range of products. In one embodiment, the immunoglobulin polypeptides are a therapeutic, a diagnostic, or a research reagent. In one embodiment, the immunoglobulin polypeptides are therapeutic, can be used for agricultural or industrial uses, and/or can find use in an antibody composition that is monoclonal or polyclonal. They can further be agonists, antagonists, neutralizing, inhibitory, or stimulatory. In one embodiment, they are used to kill target cells that bear the target antigen, for example virally infected cells. In an alternate embodiment, the immunoglobulin polypeptide is used to block, antagonize, or agonize the target antigen. In an alternate embodiment, the immunoglobulin polypeptides are used to block, antagonize, or agonize the target antigen and kill the target cells that bear the target antigen.

The *Macaca mulatta* immunoglobulin polypeptides of the present invention (e.g., scFv), including subsequences, modified forms, multimers and nucleic acids encoding them, can be incorporated into pharmaceutical compositions. Such pharmaceutical compositions are useful for administration to a subject *in vivo* or *ex vivo*, and for providing therapy for a physiological disorder or condition treatable with an antibody as provided herein.

In one embodiment the compositions of this invention comprise a polypeptide of this invention, alone or in some embodiments, in combination with a second pharmaceutically active or therapeutic agent. In one embodiment, the term “pharmaceutically active agent” refers to any medicament which satisfies the indicated purpose. In some embodiments, the term “agent” of this invention is a decongestant, antibiotic, bronchodilator, anti-inflammatory steroid, leukotriene antagonist or histamine receptor antagonist, and the like.

In one embodiment, the route of administration can be parenteral, or a combination thereof. In another embodiment, the route can be intra-ocular, conjunctival, topical, transdermal, intradermal, subcutaneous, intraperitoneal, intravenous, intra-arterial, vaginal, rectal, intratumoral, transmucosal, intramuscular, intravascular, intraventricular, intracranial, inhalation (aerosol), nasal aspiration (spray), intranasal (drops), sublingual, oral, aerosol, a gel formulation for vaginal or intrarectal application, or suppository or a combination thereof. In one embodiment, the dosage regimen will be determined by skilled

clinicians, based on factors such as exact nature of the condition being treated, the severity of the condition, the age and general physical condition of the patient, body weight, and response of the individual patient.

For intranasal administration or application by inhalation, solutions or suspensions
5 of the compounds mixed and aerosolized or nebulized in the presence of the appropriate carrier suitable. Such an aerosol can comprise any agent described herein.

For parenteral application, particularly suitable are injectable, sterile solutions, preferably oily or aqueous solutions, as well as suspensions, emulsions, or implants, including suppositories and enemas. Ampoules are convenient unit dosages. Such a
10 suppository can comprise any agent described herein.

Sustained or directed release compositions can be formulated, *e.g.*, liposomes or those wherein the active compound is protected with differentially degradable coatings, *e.g.*, by microencapsulation, multiple coatings, etc. Such compositions can be formulated for immediate or slow release. It is also possible to freeze-dry the new compounds and use
15 the lyophilisates obtained, for example, for the preparation of products for injection.

For liquid formulations, pharmaceutically acceptable carriers can be aqueous or non-aqueous solutions, suspensions, emulsions or oils. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions,
20 including saline and buffered media. Examples of oils are those of petroleum, animal, vegetable, or synthetic origin, for example, peanut oil, soybean oil, mineral oil, olive oil, sunflower oil, and fish- liver oil.

In one embodiment, a composition of or used in the methods of this invention can be administered alone or within a composition. In another embodiment, compositions of
25 this invention admixed with conventional excipients, *i.e.*, pharmaceutically acceptable organic or inorganic carrier substances suitable for parenteral, enteral (*e.g.*, oral) or topical application which do not deleteriously react with the active compounds can be used. In one embodiment, suitable pharmaceutically acceptable carriers include but are not limited to water, salt solutions, alcohols, gum arabic, vegetable oils, benzyl alcohols, polyethylene
30 glycols, gelatine, carbohydrates such as lactose, amylose or starch, magnesium stearate, talc, silicic acid, viscous paraffin, white paraffin, glycerol, alginates, hyaluronic acid, collagen, perfume oil, fatty acid monoglycerides and diglycerides, pentaerythritol fatty acid esters, hydroxy methylcellulose, polyvinyl pyrrolidone, etc. In another embodiment, the

pharmaceutical preparations can be sterilized and if desired mixed with auxiliary agents, *e.g.*, lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, coloring, flavoring and/or aromatic substances and the like which do not deleteriously react with the active compounds. In another embodiment, they can
5 also be combined where desired with other active agents, *e.g.*, vitamins.

Pharmaceutical compositions include “pharmaceutically acceptable” and “physiologically acceptable” carriers, diluents or excipients. In one embodiment, the terms “pharmaceutically acceptable” and “physiologically acceptable” refers to any formulation which is safe, and provides the appropriate delivery for the desired route of administration
10 of an effective amount of at least one compound for use in the present invention. This term refers to the use of buffered formulations as well, wherein the pH is maintained at a particular desired value, ranging from pH 4.0 to pH 9.0, in accordance with the stability of the compounds and route of administration. The terms include solvents (aqueous or nonaqueous), solutions, emulsions, dispersion media, coatings, isotonic and absorption
15 promoting or delaying agents, compatible with pharmaceutical administration. Such formulations can be contained in a liquid; emulsion, suspension, syrup or elixir, or solid form; tablet (coated or uncoated), capsule (hard or soft), powder, granule, crystal, or microbead. Supplementary active compounds (*e.g.*, preservatives, antibacterial, antiviral and antifungal agents) can also be incorporated into the compositions.

20 Pharmaceutical compositions can be formulated to be compatible with a particular local or systemic route of administration. Thus, pharmaceutical compositions include carriers, diluents, or excipients suitable for administration by particular routes.

Specific non-limiting examples of routes of administration for compositions of the invention are inhalation or intranasal delivery. Additional routes include parenteral, *e.g.*,
25 intravenous, intradermal, subcutaneous, oral, transdermal (topical), transmucosal, and rectal administration.

Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents;
30 antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity

such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide.

A skilled artisan will understand how to properly dose the desired compositions of the present invention. The compositions provided herein to be administered to a
5 subject contains, in one embodiment, a quantity of the active agent in a therapeutically effective amount for relief of the particular disease or condition being treated. An amount adequate to accomplish this is defined as a therapeutically effective dose. Amounts effective for this use will depend upon the severity of the disease and the general state of the subject's own immune system. Dosing schedules will also vary with the disease state
10 and status of the subject, and will typically range from a single bolus dosage or continuous infusion to multiple administrations per day (*e.g.*, every 4-6 hours), or as indicated by the treating physician and the subject's condition. It should be noted, however, that the present invention is not limited to any particular dose. Generally, however, for immunoglobulin polypeptides of the present invention, the preferred dosage is 0.1 mg/kg to 100 mg/kg of
15 body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to
20 enhance uptake and tissue penetration (*e.g.*, into the epithelium). A method for lipidation of antibodies is described by Cruikshank *et al.* (1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent
25 applications cited throughout this application, as well as the Figures, are incorporated herein by reference.

EXAMPLES

Example 1: Primer Design

30 To design primers specific for rhesus monkey (RM) Ig genes (Table 1, 2, and 3), RM Ig V gene sequences obtained in previous experiments, as well as sequences available through public databases, IMGT, the international ImMunoGeneTics information system® (IMGT/LIGM-DB) and NCBI, The National Center for Biotechnology Information, were

analyzed. Sequence analysis and germline gene identification were performed using IMGT-V-Quest software. Sequence alignment was performed using the DNASTar Lasergene software package (DNASTar Inc). In addition, using known sequences of rare human V genes, the RM genome was BLASTed to identify potential counterparts. In some
 5 embodiments, the primers specific for rhesus monkey (RM) Ig genes can be selected from those listed in Table 3 (*e.g.*, single primers, primer pairs, groups of primer pairs, etc.). To validate newly designed primers, V_H and V_L genes were amplified using cDNA derived from mRNA isolated from the mixture of bone marrow B cells of 10 RMs. The resulting PCR products were cloned into the sequencing plasmid vector, pCR4-TOPO (Invitrogen),
 10 and sequenced. Nucleotide sequences of 300 randomly chosen clones were analyzed using IMGT/V-QUEST and translated into the amino acid sequences shown in Table 4. In addition, representative nucleic acid and amino acid sequences encoding Ig V_H, V_L lamda, V_L kappa, as well as scFv and linkers, are shown in Tables 5 and 6 (n, sequence number; F, V-gene subgroup family). Primers listed in Tables 1, 2 and 3 are suitable to amplify any
 15 *Macaca mulatta* V/D/J gene combination.

Example 2: RT-PCR of Ig V_H and V_L genes from single B cell

Single B-cell RT-PCR was used to validate the ability of developed primers amplify V_H and V_L genes from single B cells. The cDNA synthesis and Ig amplification were
 20 performed as previously described in Tiller *et al.* (2008) *J. Immunol. Meth.* 329:112-124, with the following modifications. The frozen plates with single memory B cells were thawed, and reverse transcription was performed by adding 3 µl of random hexamer primers (Applied Biosystems) at 50 µM, 1 µl of 10 mM dNTP mix (Invitrogen), 0.0625 µl of Igepal CA-630 (Sigma), 40 units of RNaseOUT™ (Invitrogen), 1.25 µl of 0.1 DTT
 25 (Invitrogen) and 0.25 µl of SuperScript III (Invitrogen) into each well. Reaction conditions for reverse transcription were as follows: 42°C for 10 min, 25°C for 10 min, 50°C for 60 min and 94°C for 5 min. The cDNA plates were stored at -20°C until further use. The IgH, Igλ and Igκ V genes were amplified independently by semi-nested PCR starting from 3 µl of cDNA as a template. All PCRs were performed in 96-well PCR plates in a total volume
 30 of 50 µl containing water, 5 µl of 10× buffer, 1 µl of dNTP mix, each at 10 mM, 1 µl of MgCl₂ at 25 mM (Qiagen), by 1 µl of primer mix specific to FR1 region (Table 1 and/or 3) and primer specific to the constant part of heavy or light chain (Table 2 and/or 3) at 10 µM, and 0.4 µl of HotStart Taq DNA polymerase (Qiagen). The PCR thermocycler program

was: 95°C for 15 min; 50 cycles (95°C for 30 sec, 50°C for Igκ and Igλ or 54°C for IgH for 30 sec and 72°C for 1 min); 72°C for 10 min. For PCR2, 3 μl of PCR1 product were used as a template. Primers specific to FR1 region were as in PCR1 and for PCR2 we used Cγ_GSP2, Cκ_GSP2 and Cλ_GSP2 or V_H_forward, Vκ_forward and Vλ_forward or γ-PCR2, λ-PCR2, and κ-PCR2 (Table 1, 2, and 3). Fifty cycles were used with the same parameters as those for the first round, with annealing temperature of 52°C for Igκ and Igλ and 57°C for IgH for 30 sec. After PCR2, the DNA fragments were isolated and subjected to direct sequencing with the reverse primer used for PCR2. Sequences were analyzed with IMGT/V-QUEST and are shown herein.

Example 3: Design of primers specific to leader sequences of VH genes.

Variable genes of antibodies which have undergone somatic hypermutation of CDRs during maturation can also contain mutations in conserved framework regions (Jung *et al.* (2001) *J. Mol. Biol.* 309:701 and Scheied *et al.* (2011) *Science* 333:1633). In antibody genes with a high frequency of mutations, framework region 1 (FR1) may no longer be recognized by specific oligonucleotides due to acquired mutations, which leads to loss of amplification of such immunoglobulin genes. To overcome this obstacle, primers specific to the leader sequence (peptide region) of rhesus monkey VH genes have been designed. The leader sequence is located at the 5'-end of any VH gene and is very conserved. Since leader sequences for rhesus monkey VH genes have not been published or are not available through public resources, the following algorithm to retrieve RM leader sequences from the rhesus monkey genomic databases at National Center for Biotechnology Information (NCBI) or at Baylor College of Medicine has been used. Sequences of human VH genes were taken to survey RM genomic databases in order to find orthologous rhesus VH genes. Once RM VH genes were identified (human nomenclature of Ig V genes is used to name RM Ig genes), the sequence located 5'-terminal to the newly found VH gene was compared with known human leader sequences in order to localize RM leader sequences. When the leader sequence was determined, introns were cut from assembled sequences. Using sequences thus obtained, three sets of primer specific to rhesus monkey leader sequences of VH genes were designed (Tables 8-10). RM germline VH genes used for leader primers design are listed in Table 12.

Example 4: PCR amplification of VH genes using different combinations of primers.

Primers listed in Tables 1-3 and 8-10 can be used to amplify rhesus monkey VH and VL genes in different combinations in direct, nested and semi-nested PCR. All reverse

5 primers can be used to obtain rhesus monkey VH and VL genes by RACE-PCR. To amplify VH genes, forward primers given in Table 1 VH1(FR1)back – VH7(FR1)back can be combined with any reverse primer for heavy chain (V_H-forward [Table 1]; C_γ_GSP1 or C_γ_GSP2 or C_γ_γCT [Table 2]; γ-PCR1 or γ-PCR2 [Table 3]) in regular or semi-nested PCR. To amplify VH genes, forward primers given in Table 3 “VH1” – “VH7” can be

10 combined with any reverse primer for heavy chain (V_H-forward [Table 1]; C_γ_GSP1 or C_γ_GSP2 or C_γ_γCT [Table 2]; γ-PCR1 or γ-PCR2 [Table 3]) in direct or semi-nested PCR. To amplify VH genes, in some embodiments, primers that bind to leader sequences can be used. To amplify VH genes, forward primers given in Table 8 can be combined with any reverse primer for heavy chain (V_H-forward [Table 1]; C_γ_GSP1 or C_γ_GSP2 or

15 C_γ_γCT [Table 2]; γ-PCR1 or γ-PCR2 [Table 3]) in direct or semi-nested PCR. To amplify VH genes, forward primers given in Table 9 can be combined with any reverse primer for heavy chain (V_H-forward [Table 1]; C_γ_GSP1 or C_γ_GSP2 or C_γ_γCT [Table 2]; γ-PCR1 or γ-PCR2 [Table 3]) in direct or semi-nested PCR. To amplify VH genes, forward primers given in Table 10 can be combined with any reverse primer for heavy chain (V_H-forward

20 [Table 1]; C_γ_GSP1 or C_γ_GSP2 or C_γ_γCT [Table 2]; γ-PCR1 or γ-PCR2 [Table 3]) in direct or semi-nested PCR. To amplify VH genes, forward primers given in Table 9 can be used as inner primers for nested PCR while primers listed in Tables 8 or 10 can be used as outer primers and both sets can be combined with appropriate reverse primer (V_H-forward [Table 1]; C_γ_GSP1 or C_γ_GSP2 or C_γ_γCT [Table 2]; γ-PCR1 or γ-PCR2 [Table 3]).

25 Primers listed in Table 1 (VH1(FR1)back – VH7(FR1)back) or primers listed in Table 3 (VH1-VH7) can be used as inner primers for nested PCR with primers listed in Tables 8, 9 or 10 used as outer primers. RT-PCR was used to validate the ability of newly developed primers specific to leader sequences to amplify V_H and V_L genes. Total or mRNA was isolated from rhesus monkey PBMC or purified B cells. The cDNA synthesis and Ig VH

30 genes amplification were performed as described in Example 2. Figures 3 and 4 illustrate VH gene amplification with primers specific to the leader sequence listed in Tables 8, 9 and 10. Examples of VH genes and their allelic variants obtained with help of forward primers listed in Tables 8, 9 and 10 and reverse primers listed in Table 3 are given in Table 11. The

human nomenclature is used to name RM Ig VH and VL genes. Primers listed in Tables 8, 9 and 10 are suitable to amplify any *Macaca mulatta* heavy chain V/D/J gene combination.

Incorporation by Reference

5 All publications, patents, and patent applications mentioned herein are hereby incorporated by reference in their entirety as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

Also incorporated by reference in their entirety are any polynucleotide and
10 polypeptide sequences which reference an accession number correlating to an entry in a public database, such as those maintained by The Institute for Genomic Research (TIGR) on the world wide web at tigr.org and/or the National Center for Biotechnology Information (NCBI) on the world wide web at ncbi.nlm.nih.gov.

15 Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

Table 1

Primer name (location)	Primer Sequence	Amplification
V_κ1 (FR1) back	GACATYGWMMTGWCCCAGTCTCC	PCR1
V_κ2 (FR1) back	SAWRYWGTGMTGACHCAGWCTCC	PCR1
V_κ3 (FR1) back	SAHATHRHRHTSACKCAGTCTCC	PCR1
V_κ4/7 (FR1) back	GACATYGTGMTGACCCAG	PCR1
V_κ5 (FR1) back	GAKCTWCAGATGACMCAG	PCR1
V_κ5 (FR1) back	GAAACGAYACTCACRCAG	PCR1
V_κ6 (FR1) back	GAWATTGTGMTGACWCAGTC	PCR1
V_κ7 (FR1) back	GACATTGTGCTGACCCAGTCTCC	PCR1
V_κ forward	GGTGCAGCCACAGCTCGTTTGAT	PCR1
V _κ 1 Nco I (FR1) back	CCAGCCGGCCATGGAGGACATYGWMMTGWCCC AGTCTCC	Nested PCR
V _κ 2 Nco I (FR1) back	CCAGCCGGCCATGGAGSAWRYWGTGMTGACHC AGWCTCC	Nested PCR
V _κ 3 Nco I (FR1) back	CCAGCCGGCCATGGAGSAHATHRHRHTSACKC AGTCTCC	Nested PCR
V _κ 4/7 Nco I (FR1) back	CCAGCCGGCCATGGAGGACATYGTGMTGACCC AG	Nested PCR
V _κ 5 Nco I (FR1) back	CCAGCCGGCCATGGAGGAKCTWCAGATGACMC AG	Nested PCR
V _κ 5 Nco I (FR1) back	CCAGCCGGCCATGGAGGAAACGAYACTCACRC AG	Nested PCR
V _κ 6 Nco I (FR1) back	CCAGCCGGCCATGGAGGAWATTGTGMTGACWC AGTC	Nested PCR
V _κ 7 Nco I (FR1) back	CCAGCCGGCCATGGAGGACATTGTGCTGACCC AGTCTCC	Nested PCR
V _κ forward	CCACCGGAGCCGCCGCCGCCGACCCACCACC ACCAGAACCACCACCACCCGGTGCAGCCACAG CTCGTTTGAT	Nested PCR
V_λ1 (FR1) back	CAGTCTGTRCTGAATCAGCC	PCR1
V_λ2 (FR1) back	CAGKCTGCCCCYGAYTCAGYTC	PCR1
V_λ3A (FR1) back	TCCTCTGGGCTGACTCAG	PCR1
V_λ3B (FR1) back	TCCTMTGAGCTGACACAGCCDCC	PCR1
V_λ4 (FR1) back	CAGCYTGTGCTGACTCARTCGCC	PCR1
V_λ5 (FR1) back	MAGSCTRTGCTGACTCAGCCRRC	PCR1
V_λ6/7/8 (FR1) back	CAGACTGTGGTGACYCAGGAGYC	PCR1
V_λ9 (FR1) back	CAGCYTGTGCTGACTCARCCACC	PCR1
V_λ10 (FR1) back	CAGGCAGGGCTGACTCAGCCACC	PCR1
C_λ forward	GGCAGCCTTGGGCTGACCTAGGAC	PCR1
V _λ 1 NCOI (FR1) back	CCAGCCGGCCATGGAGCAGTCTGTRCTGAATC AGCC	Nested PCR
V _λ 2 NCOI (FR1) back	CCAGCCGGCCATGGAGCAGKCTGCCCYGAYTC	Nested PCR

	AGYTC	
V _λ 3A NCOI (FR1) back	CCAGCCGGCCATGGAGTCCTCTGGGCTGACTC AG	Nested PCR
V _λ 3B NCOI (FR1) back	CCAGCCGGCCATGGAGTCCTMTGAGCTGACAC AGCCDCC	Nested PCR
V _λ 4 NCOI (FR1) back	CCAGCCGGCCATGGAGCAGCYTGTGCTGACTC ARTCGCC	Nested PCR
V _λ 5 NCOI (FR1) back	CCAGCCGGCCATGGAGMAGSCTRTGCTGACTC AGCCRRC	Nested PCR
V _λ 6/7/8 NCOI (FR1) back	CCAGCCGGCCATGGAGCAGACTGTGGTGACYC AGGAGYC	Nested PCR
V _λ 9 NCOI (FR1) back	CCAGCCGGCCATGGAGCAGCYTGTGCTGACTC ARCCACC	Nested PCR
V _λ 10 NCOI (FR1) back	CCAGCCGGCCATGGAGCAGGCAGGGCTGACTC AGCCACC	Nested PCR
V _λ _forward	CCACCGGAGCCGCCGCCGCCGACCCACCACC ACCAGAACCACCACCACCGGCAGCCTTGGGCT GACCTAGGACG	Nested PCR
V _H 1 (FR1) back	SAGGWSCAGCTGSTRSARTCYGG	PCR1
V _H 2 (FR1) back	CAGGTGACCTTGAAGGAGTCTGG	PCR1
V _H 3/H5 (FR1) back	SAGGTGSAGYTGKTGSAGTCTGG	PCR1
V _H 4/H6 (FR1) back	SAGGTRCARCTSCAGSAGTCDGG	PCR1
V _H 7(FR1) back	CAGGTGCAGCTGGTGCAATCTGG .	PCR1
V _H _forward	GCCCTTGGTGGAGGCTGAGGA	PCR1
V _H 1 (FR1) back	GGGTCTGGGCGGCGGCGGCTCCGGTGGTGGTGG ATCCSAGGWSCAGCTGSTRSARTCYGG	Nested PCR
V _H 2 (FR1) back	GGGTCTGGGCGGCGGCGGCTCCGGTGGTGGTGG ATCCCAGGTGACCTTGAAGGAGTCTGG	Nested PCR
V _H 3/H5 (FR1) back	GGGTCTGGGCGGCGGCGGCTCCGGTGGTGGTGG ATCCSAGGTGSAGYTGKTGSAGTCTGG	Nested PCR
V _H 4/H6 (FR1) back	GGGTCTGGGCGGCGGCGGCTCCGGTGGTGGTGG ATCCSAGGTRCARCTSCAGSAGTCDGG	Nested PCR
V _H 7(FR1) back	GGGTCTGGGCGGCGGCGGCTCCGGTGGTGGTGG ATCCCAGGTGCAGCTGGTGCAATCTGG	
V _H _forward-NotI	TGTTCTGCGGCCGCGCCCTTGGTGGAGGCTGA GGA	Nested PCR
scFv_Back (Sfi I-NcoI)	GCGGCCAGCCGGCCATGGAG	SOE-PCR (splicing by overlap extension PCR)
scFv_FW (Not I)	TGTTCTGCGGCCGCGCCCTTGG	SOE-PCR

Table 1. Set of subgroup gene-specific primers used to amplify the VL and VH domains of rhesus monkey IgG. PCR1 primers are used to amplify the Ig variable domain V-J and V-D-J for the light and heavy chains, respectively. The nested PCR primers are used to add the cloning site

5 Sfi I and Nco I at the N-terminus of the light chains and the forward linker sequence to the C-terminus of the variable domain, which overlaps with the nested PCR V_H back-primer to generate a linker of twenty amino acids. The scFv_Back and scFv_FW are used to amplify the full-length scFvs after overlapping of the linker tails during the nested PCR. The direction of the primer sequences is 5' to 3'. Forward primers are complementary to the sense strand, and back-primers are complementary to the antisense strand. The symbols for degenerate nucleotides are as follows: Y = C or T; R = A or G; W = A or T; S = C or G; K = T or G; M = A or C.

Table 2

Primer name	sequences
C λ λ -CT	TTTTCTGGTTGACGGAGGATA
C λ GSP1	gcttgragctcctcagagg
C λ GSP2	GGGYGGGAACAGAGTGACC
C γ γ CT	TTTACCCGGAGACAGGGAGAGGC
C γ GSP1	TGTGCAGAGCCTCATGCATCACGG
C γ GSP2	GTAGTCCTTGACCAGGCAG
C κ GSP1	CCTGATGGGTGACTTCGCAGGCG
C κ GSP2	GAGGCAGTTCCAGATTTCAA
C κ CT	GCACTCTCCCCTGTTGAAGCTC
C κ GSP1	CCTGATGGGTGACTTCGCAGGCG
C κ GSP2	GAGGCAGTTCCAGATTTCAA
RM C γ 1h	GCTGCCACCACCACATGTTTTTATC
RM C γ 2h	CACGTGGAACGACATGGGAGCCC
RM C γ 3h	GTCACCACATGGGGGTGTGAACTC
RM C γ 4h	GGGCATGGGGGTGTGAACTC

Table 2. Set of gene-specific primers used to amplify rhesus monkey Ig constant regions. CT are forward primers that hybridize to the 3'end of the light or heavy chains. RM_C γ 1-4h are primers specific to the each isotype hinge constant region. GSP1 primers hybridize to the 5' end of the corresponding constant region, and GSP2 primers hybridize within the constant region at a relatively conserved position.

Table 3. Set of subgroup-specific primers used to amplify rhesus monkey VH and VL genes from PBMC and single B cells

Forward primer	5' – 3' sequence
VH-1	SAGGWSCAGCTGGTRCAATCCGG
VH-2	CAGGTGACCTTGAAGGAGTCTGG
VH-3/5/7	SAGGTGCAGYTGGTGSAGTCTGG
VH-4/6	CAGGTGCARCTGCAGGAGTCRGG
VH-5	GAGGTGCAGCTGGTGCAGTCTGG
VH-6	CAGGTACAGCTGCAGCAGTCAGG
VH-7	CAGGTGCAGCTGGTGCAATCTGG
VL-1	CAGTCTGTRCTGACVCAGCCDCC
VL-2	CAGKCTGCCCYGAYTCAGYCTCC
VL-3A	TCCTCTGGGCTGACTCAG
VL-3B	TCCTMTGAGCTGACACAGCCDCC
VL-4	CAGCYTGTGCTGACTCARTCGCC
VL-5	MAGSCTRTGCTGACTCAGCCORRC
VL-6	AATTTTATGCTGACTCAGCCC
VL-8/7	CAGACTGTGGTGACYCAGGAGYC
VL-9	CAGCYTGTGCTGACTCARCCACC
VL-10	CAGGCAGGGCTGACTCAGCCACC
Vκ-1	GACATYCAGATGWCCCAGTCTCC
Vκ-2	GATAYTGTGATGACCCAGACTCC
Vκ-3	SAAATWGTRWTGACKCAGTCTCC
Vκ-4	GACATYGTGMTGACCCAGTCTCC
Vκ-5	GAAACGACACTCACGCAGTCTCC
Vκ-6	GAWRTTGTGMTGAOWCAGTCTCC
Vκ-7	GACATTGTGCTGACCCAGTCTCC
Reverse primer	5' – 3' sequence
γ-PCR1	GGACAGCCKGGAAGGTGTGC
γ-PCR2	GCCTGAGTTCCACGACACGGTCAC
λ-PCR1	CCGCGTACTTGTTGTTGCTCTGT
λ-PCR2	CAGAGGAGGGCGGGAASAGA
κ-PCR1	GAGGCAGTTCCAGATTTCAA
κ-PCR2	GGTGCAGCCACAGCTCGTTTGAT

N = A+G+C+T; V = A+C+G; D = A+T+G; B = T+C+G; H = A+T+C; W = A+T; S = C+G;
 K = T+G; M = A+C; Y = C+T; R = A+G;

5

Table 4. Examples of VH and VL sequences (translated) derived using primers listed in Tables 1-3

[illegible]

Table 5. Examples of VH and VL sequences amplified with Table 1-2 primers and scFv assembled with primers listed in Table 1

n	F	Sequences
1	VH1	GAGGTGCAGCTGGTGGAGTCTGGGGCTGGCGTGAAGAAGCCTGGGTCTCAGTGAAGGTGTCTGTGAGCTTCCGGATACAC CTTCACTGGCTACTCCGTACACTGGGTGCGACAGGCCCTCGACAAGGACTTGAGTGGATGGGATGGATCAACCTTTACAATG GCAACACAAAATATTACAGAAGTTCAGGGCAGAGTCACCATGACCAGAGACACATCGACTAACACAGCCTACATGGAGCTG AGCGGCCTGACATCTGACGACACGGCCGTATATTACTGTGCGAGAAGGACTGTGGTGGTATTTACTTCTATCGAACTTTATGA CTTCTGGGGCCAGGGGGTCTGGTCACCGTCTCCTCA EVQLVESGAGVKKPGSSVKVSCQTSYGFYFTGYSVHWVRQAPRQGLEWMGWINPYNGNTKYSQKFQGRVTMTDRDSTNTAYMEL SGLTSDDTAVYYCARRTVVVFSTIELYDFWGGQVLTIVSS
2	VH2	CAGGTGACCTTGAAGGAGTCTGGGCTGTGCTCGTGAAGCCACACAGACCCTCAGCTGACCTGCACCTTCTCTGACTTCTC TCTCAGTTCTACTGGAATAGGTGTGGGCTGGATCCGTGAGCCCCAGGGAAGGCCCTGGAATGGCTTGACGCGATTGATTGGG ATGATAATAAATACTATAATACATCGCTGAAGAGCAGCCTCACCATCTCCAAGGACACCTCCAAAAACCAGGTGGTTCTAACA ATGACCAACATGGACCCTGTGGACACAGCCACATATTACTGTGACGGGTATGGCGGGGTATGACTACTGGGGCCAGGGAGT CCTGGTCACCGTCTCCTCA QVTLKESGVLVLPKPTQTLTCTFSDFLSSTGIGVWIRQPPGKALEWLARIDWDNDKYYNTSLKSSLTISKDTSKNQVVL MTNMDPVDATYYCARVWRGYDYWGQGVLTVSS
3	VH3	CAGGTGCAGCTGGTGCAGTCTGGGGCGGCTTGGTCCAGCCTGGGGGGTCCCTGAGACTCTCCTGTGTAGTCTCTGGATTAC CTTCAGTAGCTATGGCATGTACTGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCTCAGCTATTAATAGTGGTGGT GTAACACATACTACGCACTCTGTGAAGGGCCGATTACCATGTCCAGAGACAACTCAAAGAACACGCTCTCCCTGCAAATG AACAGCCTGAGAGCTGAGGACACGGCCGTGTATTATTGTGCGAAACGGCCGCTGGATCTAGGTACAGTCACGGTTTAGATT CTGGGGCCAGGGGTCTGTCGTCACCGTCTCCTCA QVQLVQSGGGLVQPGGSLRLSCVVSGETFSSYGMVWRQAPGKLEWISA INSGGNTYYADSVKGRFTMSRDNSKNTLSLQ NSLRAEDTAVYYCAKRPRGSRYSHGLDSWGQGVVTVSS
4	VH4	CAGGTGCAACTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCGGAGACCCTGTCCCTCACCTGTGTTGTCTCTGGTGGCTC CATCACTGGCTTCTACTACTGGAGTGGATTGCGACAGTCCCAGGGAAGGGACTGGAGTGGATTGGCTACATCGCTGACA GTGGGCGCACCAAGTCCAGTCCGTCCCTCCAGAGTCGAGCCATTATTTCAAGAGACACGTCCAAGAATGAGTTCAACCTGCAG ATGAGTCTGTGACCGTCGCGNACACGGCCATGTATTATTGTGCGAGAGTGAGATTACAGTTGGGTAAAGGACTACTGGGG CCAGGACTCCCGGTACCGTCTCCTCA QVQLQESGPGLVKPSSETLSLTCTVSSGSIITTFYYWSWIRQSPGKLEWIGYIADSGRTKSSPSLQSRALISRDTSKNEFNLQ MRSVTVAXTAMYYCARVRFSTWVKDYWGQGLPVTSS
5	VH5	GAGGTGCAGCTGGTGCAGTCTGGAGCAGAGGTGAAAAGGCCCGGGGAGTCTCTGAAGATCTCCTGTAAGACTTCTGGATACAC CTTTACCAGCTACTGGATCACCTGGGTGCGCCAGATGCCGGGAAAGGCCCTGGAGTGGATGGGGGCGATTGATCCTAGTGATT CTGAAACCAGATACAACCGTCTCTCCAAGGCCAGGTACCATCTCAGCCGACAAGTCCATCAGTACCACCTACCTGCAGTGG AGCGGGCTGAAGGCCTCGGACACGGCCACGTATTACTGTGCGAAAGGTTGGGATACAGTAACTACCTACGTTTGGATTCTCTG GGGCAAGGGGTCGTCGTCACCGTCTCCTCA EVQLVQSGAEVVKRPGESLKI SCKTSGYFTFSYWITWVRQMPGKLEWMAIDPSDSETRYNPFSQGVITISADKSISTTYLQW SGLKASDTATYYCAKGDVTTTYGLDSWGQGVVTVSS
6	VH7	CAGGTGCAGCTGGTGCAGTCTGGGCTGAGGTGAAGCAGCCGGAGCCTCAGTGAAGGTCTCCTGCAAGGCTTCTGGATACAC CTTCACTGGCTATGGTATCAACTGGCTGCGACAGGTCCATGGACAACGGCTTGAGTGGATGGGATGGATCAATACCGACAGTG GGAACACAGATATGCCAGGGCTTCAGAGAACGATTACCTTCTCCATGGACACCTCTATCACCACGGCATATCTGGAGATC AGCAGCCTAAAGCCTGAGGACACGGCCGTTTATTACTGTACTAGAGGCATTGCCCAACAAGGTTTGATTCTCTGGGGCCAAAG GGTCGTCGTCACCGTCTCCTCA QVQLVQSGAEVVKRPGASVKVSKASGYFTGYGINWLRQVHGQRLEWMGWINTDSGKQTYAQGFRRERFTFSMDTSITAYLEI SSLKPEDTAVYYCTRGIAHGLDSWGQGVVTVSS

n	F	Sequences
7	VL1	CAGTCTGTGCTGACGAGCCGCCCTCAGCATCTGGGGCCCCGGGAGAGGTCACCATCTCCTGCACTGGGAGCAGCTCCAA CATCGGGCGGGTTTTTATGTATCTTGGTACCAGCAGTTCACAGGAACAGCCCCAACTCCTCATCTATGACAATAATAAGC GACCTCAGGGGTTTCTGACCGATTCTCTGGTCCAAGTCTGGGACCTCAGCCTCCCTGACCATCACTGGGCTCCAGACTGAG GATGAGGCTGATTATTACTGCCTGTCTATGACAGCAGCCTGAGTGCTTACATCTTCGGTGCTGGGACCCGGCTACCGTCTCT A QSVLTQPPSASGAPGQRVTISCTGSSSNIGAGFYVSWYQQFPGTAPKLLIYDNNKRFSGVSDRFSGSKSGTSASLTITGLQTE DEADYCLSYDSSLSAYIFGAGTRLTVL
8	VL2	CAGGCTGCCCTGATTCACTCTCCCTCAGTGTCTGGGTCTCTGGACAGTCGGTCACCATCTCCTGCACTGGAACCCACAGTGA CATCGGTGCTTACAACTATGTCTCTGGTACCAACACCGTCCAGGCACAGCCCCAACTCATACTTTCTGCGGTCACTGAGC GGCCCTCAGGGGTTTCTGATCGCTTCTCTGGTCCAAGTCTGGCAGCTCGGCTCCCTGACCATCTCTGGACTCCAGGCTGAG GACGAGCTGATTATTACTGCTACTCATATGACGAGCAGTAAACTTTCTACATCTTCGGTGCGGGGACCCGGCTACCGTCTCT A QAALIQSPSVSGSPGQSVTISCTGTHSDIGAYNYVSWYQHRPGTAPKLLISAVSQRFSGVSDRFSGSKSGSSASLTISGLQAE DEADYYCYSYAGSKTFYIFGAGTRLTVL

9	VL3	TCCTCTGAGCTGACACAGCCGCCCTCGGTGTACAGCGCCTCAGGACAGACGGCCAGGATCACCTGTGGGGAGACAATATTGG AAGTAAAGTTGTGCACTGGTACCAGGAGAAGCCACCACAGGCCCTGTGGTGGTCATGTATGCTGATGACAGACGGCCCTCAG GCATCCCTGAGCGATTCTCTGGCTCCAACCTCAGGGAACACCGCCACCCTGACCATCAGAAGGGTCGAGGCCGGGGATGAGGCT GACTATTACTGTGAGGTGTGGGGCAGTAGTAGTAAATACGCTGTATTCTGGCGGAGGGACCCGGCTGACCGTCCTA SSELTQPPSVSAASGQTARI TCGGDNIGSKVVHWYQEKPQAPVVMYADDRRPSGIPERFSGSNSGNTATLTI RRV EAGDEA DYYCQVWSSSKYAVFVGGTRLTVL
10	VL4	CAGCCTGTGCTGACTCAATCGCCGTCTGCTTCTGCCTCGTTGGGAACCTCTGTCAAACCTCACCTGCACCTGACCAGTGAGCA CAGCAACTACTATATTTTCTGGTTTCAACAGAGACCGGGGAGATCTCCCCGATATTTAATGCAGGTTAAGAGTGATGGCGCCC AGACCAAGGGGATGGGATCCCCGATCGCTTCTGGGCTCCAGTTCTGGGGCTGACCGCTACCTCACCATCTCCAACCTCCAG TCTGACGACGAGGCTGAATATTTCTGTGGAGAGTGCCACACGATTGATGCCAGTACGGTTGCATCTTCGGTACTGGGACCCG GCTCACCCTCCTA QPVLTPQSPASASLGTSVKLTCTLTSEHSNYYIFWFQQRPGRSPRYLMQVKS DGAQTKGDGIPDRFLGSSSGADRYLTISNLQ SDDEAEYFCGECHTIDAQYGCIFGTGTRLTVL
11	VL5	CAGGCTATGCTGACTCAGCCGGCCTCCCTCTCAGCATCTCCTGGAGCATCAGCCAGTCTCATGACACCTTCAGCGGTGGCAT CAATGTTGTGGCTACCATATTTCTGGTACCAGCAGAAGCCAGGGAGTCTCCCCGGTATCTTCTGAGGTACAAATCAGACT CAGATAAGGGCCAGGGCTCTGGAGTCCCCAGCGCTTCTCTGGATCCAAAAATGCTTCAGCGAACACAGGGATTTTACGCATC TCTGGCTCCAGTCTGAGGATGAGGCTGACTATTACTGTGCCATTGGGCACAGCAGCGGGTGGGTATTCTGGCGGAGGGACCCG GCTGACCGTCCTA QAMLTQPPASLSASPGASASLTCTFSGGINVAGYHIFWYQQKPGSPPRYLLRYKSDSDKGQSGVPSRFSGSKNASANTGILRI SGLQSEDEADYYCAIGHSSGWVFGGTRLTVLG
12	VL7	CAGACTGTGGTGAAGTCAAGGACCTCAATGACTGTGTGTCCTCCAGGAGGGACAGTCACTCTCACCTGTGGCTCCAGCACTGGAGC AGTCACCACTGGTAACTATCCAACTGGTTCCAGCAGAAGCCTGGACAAGTCCCGAGGGGACTGATCTACAATACAAACAGCA AACACTCTGGACCCCTGCCCGTCTCAGGCTCCCTCGCTGGGGGCAAAGCTGCCCTGACACTGTGAGGTGCACAGCCCGAG GACGAGGCTGACTTTTACTGCTTGTCTTACTATAGTGGTGTACCTCTTCGGTGTGGGACTCGGCTCACCCTCCTA QTVVTQEP SMTVSPGGTTLTTCGSSTGAVTSGNYPNWFQQKPGQVPRGLIYNTNSKHSWTPARFSGSLAGGKAALTL SGAQPE DEADFYCLLYSGAYLFGAGTRLTVL
13	VL8	CAGACTGTGGTGAAGTCAAGGACCTCAATGACTGTGTGTCCTCCAGGAGAACAGTCACTCACCTGTGGCTGAGCTCTGGCTC AGTCTCTCCCAATACCTACCCAGCTGGTTCAGCAGACCCAGGCCAGGCTCCACGCACGCTCATCTACAGTACAAACAGTC GCCCTTCTGGGGTCCCTGATCGCTTCTCTGGTTCATCCTTGGGAACAAAGCTGCCCTAACCATCACGGGGGCTCAGGCAGAC GATGAATCTGATTATTATTGTATGCTCTATAGAAGTACTGGCATTTACATTTTCGGTCTGGGACCCGGCTCACCCTCCTA QTVVTQEP SLSVSPGGTTLTTCGLSSGSVSPNTYPSWFQQTPGQAPRTL IYSTNSRPSGVPDRFSGSILGNKAALTI TGAQAD DESDYYCMLYRSTGIYIFGPGTRLTVL
14	VL10	CAGCTTGTGCTGACTCAACCACCTCAGTGTCCGAGGGCTTGAGACAGACGGCCAGCTCACCTGTAGTGGAACAGCAACAA TGTTGGCCACCAGGGAGCATCTTGGCTGCAGCAGCACCAGGGCCACCCTCCCAAGCTGCTGTCTACAGGGGTAAACACCGCC CCTCGGGGATCTCAGAGAGATTCTCTGCATCCAGGTGAGGAAACACAGCCTCCCTGACCATCACCGGACTCCAACTGAGGAT GAGGCTGACTATTACTGTTCACATGGGACAGCAGCCTCGGCGCTTACATCTTCGGTGTGGGACCCGACTCACCCTCCTA QLVLTQXPXSVSEGLRQTATLTCSGNSNNVGHQASWLQHQHPPKLLSYRGNRPSGISERFSASRSNGNTASLTITGLQPED EADYYCSTWDSSLGAYIFGAGTRLTVL
15	VL11	CAGCTTGTGCTGACTCAGTCGCCCTCCCTGTCTGCATCCCCGGGAGCATCGGCCAGACTCCCTGCACCTGACCAGTGACCTCAGTGTGGTACTAAAAACA TGACTGGTACCAGCAGAAGCCAGGGAGCGCTCCAGATTATCTGTACTACTACTCCGACTCAGACAAACAGGTGGGACCTGGGGTCCCAATCGAGTCTC TGGCTCAAAGGAGACCTCAAGTAACACAGCGTTTTTGTCTATCTCTGGCTCCAGCCTGAGGACGAGGCCGATTACTGTGCTGATGACAGTAGTTCT AATGTGTTTCGGAAGTGGCGCCAAGTTGACCGTCTC QLVLTQSP SLSASPGASARLPCTLTSDLSVGTKNMYWYQQKPGSAPRFLYYYS DSKQVPGVFNVRVSGSKETSSNTAFLLI SGLQPEDEADYYCQLYDSSSNVFGSGARLTVL

n	F	Sequences
16	VK1	GACATTAGATGTCCAGTCTCCATCCTCTGTCTGCATCTGTAGGAGACAGAGTCAACATCACTTGTGCGGGCAGTCAAGG CATTAGCAGTTATTTAGCGTGGTATCAGCAGAAACAGGGAAAGCCCCGAAGCCCTTGATCTATTACGCATCCAATTTGAAAA GTGGGGTCCCATCAGGTTCCAGCGGCAGTGGATTGGGACAGATTTTATTCTCACCATCACCAGCTGCAGACTGAAGATTTT GCAACTTATTACTGTCAACAATATCAGGATAACCTTCCACTTTCGGCGGAGGGACCAAGGTGGAGATCAAACGA DIQMSQSPSSLSASVGDVITICRASQGISSYLAWYQQKPGKAPKPLIYYASNLESGVPSRFSGSGFGTDFILITITSLQTEDF ATYYCQQYQDNLPTFGGGTKVEIKR
17	VK2	GATATTGTGATGACCCAGACTCCAGTCTCCCTGGCCATCACCCCTGGAGAGCCGGCCTCCATCTCTGCAAGGTCTAGTCAGAG CCTCTACATAGTAATGGAAACACCTTGTGCACTGGTACCTGCAGAAGCCAGGCCAGTCTCCGAAGCTCCTCATTTATGGGG GTTCCAACAGAGTCTCTGGAGTCCAGAGAGGTTTCACTGGCAGTGGGTCTGACACTGATTTACACTGAAATCAGCAGAGTG GAGGCTGAGGATGTTGGGATTTATTACTGCATGGAAGCTTTAGACTGGCCGTATATTTTGGCCAGGGGACCAAGTGGAGAT CAAACGA DIVMTQTPVSLAITPGEPAISCRSSQSLHNSNGNTLLHWYLQKPGQSPKLLIYGGSNRVSGVPERFSGSGSDTDFTLKISR EAEVDGIYCYMEALDWPYIFGQGTKEIKR
18	VK3	GAAATAGTAATGACTCAGTCTCCAGCCACCCTGTCTTTGTCTCCAGGAGAAAGAGCCACCCTCTCTCTGC AGGGCCAGTCAGAGTGTGAGCAGAGCTTAGCTGGTACCAGCAGAAACCTGGGCAGGCTCCAGGCTCCTCATCTATGGTGC ATCCAGCAGGGCCACTGGCATCCAGACAGGTTTCACTGGCAGTGGGTCTGGGACAGAGTTCACTCTCACCATCAGCAGCTGG AGCCTGAAGATGTTGCAGTTTATTATTGTGACGAGAATAGTAACTGGCCGCTCACTTTCGGCGGAGGGACCAAGGTGGAGATC AAA EIVMTQSPATLSLSPGERATLSCRASQSVSSSLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTEFTLTITSSLEPEDV AVYYCQNSNWPLTFGGGTKEIK

19	VK4	GACATTGTGCTGACCCAGTCTCCAGACTCCCTGGCTGTGTCTCTGGGCGAGAGGGTCACCATCAACTGCAAGTCCAGCCGGAC TCTTTTATATACTTCCAACAATAAGAACTACTTAGCCTGGGATCAGCAGAAACCAGGACAGGCTCCTAAGTTGCTCATTATT GGGCGTCTACCCGGGAATCCGGGGTCCCTGACCGATTCACTGGCAGTGGGTCTGGGACAGATTTCACTCTCACCATCAGTGAC CTGCAGGCTGAAGATTGGGAGTGTATTATTGTGAGCACTATTATACCCTCTCTCCTTCGCGCGAGGGACCAAGTGGA GATCAAAAC
		DIVLTQSPDLSLAVSLGERVTINCKSSRTLLYTSNNKNYLAWDQQKPGQAPKLLIYWASTRESGVPRFSGSGSGTDFTLTISD LQAEDLGYYCQHYHTPLTFGGGTKVEIK
20	VK6	GGAAGTTTGCTGACTCAGTCTCCAGCCTTTCCGGTCTGTGGCTCTAAAGGAGAAAGTCACCATCACCTGCCAGGCCAGTCAGAG CATTGGTAGGAGTTTACACTGGTACCAACAGAAACCGGATCAGTCTCCAAAGCTCATCATCAAATATGCTTCCAGTCCATCT CAGGGGTCCCTCAAGGTTCACTGGCAGTGGATCTGGGACAGATTTCACTCTCAATATAGCTGGAGGCTGAAGATGTT GCGACGTATTACTGTGAGCAGAGTAGTAGTTTCTCCTCCTTCGCGCGAGGGACCAAGGTGGAGATCAAA
		GSLLTQSPAFRSVALKEKVTITCQASQSIGRSLHWYQQKPDQSPKLIKYASQSIGVPSRFSGSGSGTDFTLTINSLEAEDV ATYYCQSSSELTTFGGGTKVEIK
21	VK7	GATATCGAGCTCACCCAGTCTCCAGCCTCTTTGGCCGTGTCTCCAGGGCAGAGGGCCACCATCACCTGCCGAGCCAGTGAGAC TCTCAGTTTCTATGGAATAAACACCTTCAGTGGTATCAACAGAAACCAGGACAGCCTCCTAAACTCTTGATTACCAAGCAT CCAATAAAGACGTTGGGGTCCCAGCCAGGTTCCGCGGAGTGGGTCTGGGACCGATTTCACCTCACAATTAATCCTGTGGAG GCTGACGATGCTGCAACTTATTACTGTCTAGAGAGTAAGAATTTACCGTACATTTTGGCCAGGGGACCAAGTGAGATCAA ACAAGGTGTGGCTGCACCAAGGGCGAATTCGTTAACTGCAGGACAGTCCCTTAGGAGGTATTTAGT
		DIELTQSPASLAVSPGQRATITCRASETLSTFYGINNLHWYQQKPGQPPKLLIYQASNKDVGPVRFAGSGSGTDFTLTINPVE ADDAATYYCLESKNLPYIFGQGTKVEIK

	scFv	GACATCCAGATGTCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGA GTCACCATCACTTGTCTGGGCAAGTCAGGGCATTAGCAGTTATTTAGCCTGGTAT CAGCAGAAACCAGGGAAAGCCCCCTAAGCTCCTGATCTATAAGGCGTCCAGTTTG CAAAGTGGGGTCCCATCAAGGTTCAAGCGGCAGTGGATCTGGGACAGATTTCACT CTCACCATCAGCAGCCCGCAGCCTGAAGATTTTGCAGTTTATTACTGTCAGCAT GGTTATGGTATCCTGTACAGTTTGGCCAGGGGACCAAAGTGAGATCAAACGA GCTGTGGCTGCACCGGGTGGTGGTGGTTCTGGTGGTGGTGGGTCTGGGCGGCGGC GGCTCCGGTGGTGGTGGATCCCAGGTACAACCTGCAGCAGTCAGGCCAGGACTG GTGAGGCCTTCGGAGACCCTGTCCCTCACCTGCGGTGTGTCTGGTGGCTCCCTC ACCGGGTACTACTGGGGTGGATCCGCCAGGCCCCAGGGAAGAGACTGGAGTGG ATTGGGAATATCGGTGGTGGAAAGTGGGACTACCGACTACAATCCCTCCCTCAAG AGTCGAGTCACGGTTTCACTAGACACGTCCAAGAAGCATATCTCCCTGGATGTG ACCTATGTGACCGCCGCGGACACGGGCGTCTATTACTGTGCGAGAGACTTGGAG GGCACTGCTTTTGATTACTGGGGCCAAGGGCTCGGGGTCTATCGTCTCCTCAGCC
		DIQMSQSPSSLSASVGDRTITCRASQGISSYLAWYQQKPGKAPKLLIYKASSLQSGVPSRFSG SGSGTDFTLTISSPQPEDFAVYYCQHGYGILYFSGQGTKVEIKRAVAAPGGGSGGGSGGGGS GGGGSQVQLQQSGPGLVRPSETLSLTCTGVSGGSLTGYWGWIRQAPGKRLIEWIGNIGGSGTDD YNPSLKSRTVSVSDTSKHHISLDVTVTAADTGVIYCARDLEGTAFDYWGQGLGIVSSA
	Linker	GGTGGTGGTGGTCTGTGGTGGTGGTGGGTCGGGCGGCGGCGGCTCCGGTGGTGGT GGATCC
		GGGGSGGGSGGGGSGGGGS

Table 6. Examples of VH and VL sequences amplified from rhesus monkey single B cells with primers listed in Table 3

IGHV sequences obtained with primers specific to FR1 region of V_H gene (VH-1 (FR1) back - VH-7 (FR1) back, listed in Table 1) and primers specific to gamma heavy chain constant gene (V_H-forward listed in Table 1 and C_γ_GSP2 listed in Table 2):

Clone: TRPBMCC3-1-gamma (IGHV3-66, IGHJ4*02)

GCCTGAGTTCCACGACACGGTCACGGGTTACAGGAAGTAGTCCTTGACCAGGCAGCCCAGGGCCG
CTGTGCTCTCGGAGGTGCTCCTGGAGGAGGGCGCCAGGGGGAAGACCGATGGGCCCTTGGTGA
GGCTGAGGAGACGGTGACCAGAACTCCCCGGCCCCAGACATCCAATGAGTCCCAATAGCAACCACT
CCCAGGACAGTTCAATCCATTTGTGGCCTTTTACACAGTGACAGACAGCCGTGCTCCTCAACTCTCAG
GCTGTCCATTTGAAGATACAGTGTGTTCTTGGCGTTCTCTTTGGCGATGGTGAATCGGCCCTTCACG
GAGTCTATGTACCATGTGCTACCCCCAGCAGTATATGTGGGTTCTACTGAAGGTGAATCCAGAGCCT
GTAGAAGAGAGTCTCAGGGAGCCCCCAGGCTGTACCAAGCCTCACCCAGATTGCACCAACTGCACC
TGTGGACAGCCTGGAAGGTGTGCGGACAGCCAGATTGCACCAGCTGCACCTG

Clone: 6-1 (5/7/10) (IGHV3-11*03, IGHJ4*02)

GAGGTGCAGCTGGTGCAGTCTGGGGGAGGCTTGGTCCAGCCTGGCGGGTCCCTGAGACTCTCCTG
TGCAGCCTCCGATTTCAGCTTCAGTGGCTACGAAATGCACTGGGTCCGCCAGGCTCCAGGGAAGTG
GCTGGAGTCCGTCTCAGATATCGGTGGTGATAGTAGTTACACACACTACGCAGACTCTGTGCAGGG
CCGATTACCATTTCCAGAGACAACGCCAGAACTCCCTGTCTCTGGAATGAACAGCCTGAGAGCC
GAGGACACGGCCGTGTACTACTGTGCGAGACGGGGATGGAGTGACTACTTTGACTACTGGGGCCA
GGGAGTCCTGGTCACCGTCTCCTCAGCCTCCACCAAGGGGCCCATCGGTCTCCCCCTGGCGCCCTC
CTCCAGGAGCACCTCCGAGAGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCTGAAC
CCGTGACCGTGTCTGTGAACTCAGGC

Clone: γO1-2 (IGHV1-24*01, IGHJ4*02)

ATGAANGTGTGCACGCCGCTGGTCAGGGAGCCTGAGTTCCACGACACGGTCACGGGTTACAGGAA
GTAGTCCTTGACCAGGCAGCCCAGGGCCGCTGTGCTCTCGGAGGTGCTCCTGGAGGAGGGCGCCA
GGGGGAAGACCGATGGGCCCTTGGTGGAGGCTGAGGAGACGGTGACCAGGGCGCCCTGGCCCCA
GAAGTCGAAGTAATCGCTAGTTCGGGCTATGGTATGGGACGCACAGAAATAGACGGCCGTGTCCTG
AGATGTCAGGCTGCTCAGTTCATGTAGGCTGTGTCTGTGGATGGGTCTCGGTTCATGGTGAGTCT
GCCCTGGAATTTGCTGCNCGAAGAATTTACCATATAGAGGATCCAAACCTCCCATCCACTCAAGC
CCTTTTGCAGGGGCTGTGCGACCCAGTGTATGGATAATTCGCTGAAGGTGTATCCGGAAACCTTGC
ANGAGACCTTCACTGAGGCCCCAGGCTTCTTACCTCAGCCCCAGACTGCACCAGCTGCACCTC

Clone: BMHV23 (IGHV2-70*12, IGHJ3)

CAGGTGACCTTGAAGGAGTCTGGTCCCGCTCTGGTGAAACCCACACAGACCCTCACGCTGACCTGC
ACCTTTTCTGGGTTCTCACTCAGTACGAGGGGTATGGGTGTGGACTGGATCCGTCAGACCTCACGG
AAGACCCTGGAGTGGCTTGACACATTGATTGGGATGATGATGAGCGCTACAATACATCTCTGAAGA
CCAGGCTCACCATCTCCAAGGACACCTCCAAAAACCAGGTGGTGCTAAGAATGACCAATGTGGACC
CTATGGACACAGGCACATATTACTGTGCACGGGTCGCCGCTGTTAGTGGTTGCTATATGTGTGCTTT
TGATTTCTGGGGCCAAGGGCTCAGGGTCACCGTCTCCTCA

Clone: BMH43 (IGHV4-39, IGHJ4)

CAGGTGCAACTGCAGGAGTCAGGCCCAGGACTGGTGAAGCCTTCGGAGACCCTGTCCCTCACTTGC
GCTGTCTCTGGTGGGTCCATCAGCGATAACTACTGGAAGTGGATCCGCCAACCCCCAGGGAAGGGG
CTGGAGTGGATTGGATATATCGGTGGTAGTAGTGGGAGCACCCACTACAATTCCTCCCTCAAGGATA
GAGTCACCATTTGTCGGACACGTCCAGAAATGAGGTCTCTTTGAGTCTGAACTCCGTGACCGCCGC
GGACTCGGCCGTATATTATTGTGCGAGAAATTTGGACAAGACGGTGGTGTCTGGTCAGGGCTACTG
GGGCCAGGGAGTCTGGTCAACCGTCTCCTCT

IGLV sequences obtained with primers specific to FR1 region of V λ gene (V λ -1 (FR1) back - V λ -10 (FR1) back, listed in Table 1) and primers specific to lambda light chain constant gene (V λ _forward listed in Table 1 and C λ _GSP2 listed in Table 2):

Clone: lambda_B3-1_T7---08/03/2010 (IGLV2-11, IGLJ1)

5 CAGAGGAGGGCGGGAACAGAGTGACCGTGGGGGAAGCCTTGGGCTGACCTAGGACGGTGAGCCG
GGTCCCAGCACCGAGGATGTAAAAAGTGTTGCTGCCTGCATATGAGATGCAGTGATAATCAGCCTCA
TCCTCAGCCTGGAGCCCAGAGATGGTCAGGGAGGCCGTGTTGCCAGACTTGGAGCCAGAGAAGCG
AGCAGAGACCCCTGAGGGCCGCTTACTGACCTCATAAATCATGAGTTTGGGGGCTTTGCCTGGATG
10 CTGTTGGTACCAAGACACAGCATTATAATACCCAATATCACTGCTGGTTCCAGTGACAGGAGATGGTG
ACCGACTGTCCAGGAGACCCGGACACAGAGG

Clone: TRPBMC-B4-1-lambda (IGLV10-54, IGLJ3)

CAGAGGAGGGCGGGAAGAGAGTGACCGAGGGGGCAGCCTTGGGCTGACCCAGGACGGTCAGCCG
GGTCCCTCCGCCGAATACCCAACCACTGAGGCTGCTGTCCCAGGCTGAGCAGTAATAGTCAGCCTC
15 ATCCTCAGGTTGGAGTCGGGTGATGGTCAGGGAGGCTGTGTTTCTGACCTGGATGCAGAGAATCT
CTCTGAGATCCCCGAGGGCCGGTTGTTATTCCTGTAGGACAGCAGTTTGGGAGGGTGCCCTGGTA
CTGCTGCAGCCAAGTTGCTCCCTGGTGGCAACATTGTTGCTGTTTCCAGTGACAGGTGAGCGTGCC
CGTCTGTCTCAAGTCCTTGGACACTGAGGGTGGCTGAGTCAGCACAAAGCTG

Clone: λ G2-1 (5/26/10) (IGLV1-47, IGLJ1)

GCAAGCCTTGGGNTGACCTAAGGACGGTGAGCCGGGTCCCAGCAGCGAANATGTAATTGCGCAGG
20 CTGTTATCCCAAGTTGCACAGTAATAATCAGCCTCATCCTCAGATCGGAGACCANNGATGGCCAGGG
AGGCTGAGGTGCCAGACTTANAGNCNGANAATCGGTCAGGGACCCCTGAGGGTCGCTGATTAGTAT
AATAGATGAGGAGTTTGGGGGCCGTTTCTGGGAGCTGCTGGTACCAGTACACATACTACTTCCAAT
GTTGGAGCTACTTCCNGAGCAAGAGATGGCGACACTCTGCCCCGGGAGCCCCAGACGCTGAGGGTG
GCTGAGTCAGCCCTGCCTGA

25 **Clone: λ O2-1 (5/26/10)** (IGLV5-52, IGLJ3)

CAGCTTGTGCTGACTCAGTCGCCTTCCATTCCGCATCTCCTGGAGCGCCAGCCAGACTCATCTGCA
CGCTGAGCAGTGGCTTCAGTGTTGGTGACTTCTGGATACGGTGGTACCAACAACAGCCAGGGACCC
CTCCCCGGTATCTCCTGTACTTCCGCTCAGCCTCAGATAAGCACCGAGGCTCTGGAGTTCCAGCC
30 GCTTCTCTGGATCCAATGATGCATCAGCCAATGCAGTGATTCTTCATATCTCTGGGCTCCAGCCTGA
GGATGAGGCTGACTATTACTGTGGTGCGTGGCATGGCAACTCTAAGACTTGGATTTTCGGCGGGGG
GACCCGGCTGACCGTTCTTAGGTCAGCCCAAGGCTTGC

Clone: 17-1 (5/10/10) (IGLV8-61, IGLJ1)

CAGCCTTGGGCTGACCCGAGGACGGTGAGCCGGGTCCCAGCACCGAAGATGTATGCCACTACCCA
TGTAGAGCGTACAGTAATAATCAGATTCATCGTCTGCCTGAGCCCCCGTGATGGTGAGGGCAGCTTT
35 GTTCCCAAGGATGGAGCCAGAGAAGCGATCAGGGACCCCAAGGGCGAGTGTTTGTGCTGTAGAT
GAGCGTGCGTGGAGCCTGGCCTGGGGTCTGCTGGTACCAGCTGGGGTAGTTACTGGTAGAGACTG
AGCCAGAGCTCAGGCCACAGGTGAGTGTGACTGTTCTCCAGGGGACACTGACAACGATGACTCCT
GAGTACCACAGTCTG

Clone: C3-8_T7---07/03/2010 (IGLV11-55(ORF), IGLJ6)

40 cagcttgctgactcagtcgcccctccctgtctgcatccccgggagcatcgccagactcccctgcaccctgaccagtgcactcagtggtgactaa
aagcatgtactggtaccagcagaagccagggagcgctcccagattatcctgtactactactccgactcagacaaacaggtgggacctggggctc
ccaatcagctctgtgctccaaggagacctaagtaacacagcggttttctcatctctgggctccagcctgaggacgagggccgattattactgtcagc
tgtatgacagtagttctaattgtgttcggaagtggcaccagttgacgctcctcggtcagcccaaggctcccccttggtcactctctcccgcctcctctg

IGKV sequences obtained with primers specific to FR1 region of V_κ gene (V_κ-1 (FR1) back - V_κ-7 (FR1) back, listed in Table 1) and primers specific to kappa light chain constant gene (V_κ_forward listed in Table 1 and C_κ_GSP2 listed in Table 2):

Clone: kG2-2 (5/26/10) (IGKV7-3*01(P), IGKJ1)

5 GATACTGTGATGACCCAGACTCCAGCCTCTTTGGCCGTGTCTCCAGGGCAGAGGGCCACCATCACC
TGCAGAGCCAGTGAGAGTGTGAGTTTCTTTGGAATAAACCTCATTCACTGGTATCAGCAGAAACCAG
GACAACCTCCTGAACTCCTGATTTACCAAGCATCCAATAAAGACACTGGGGTCCCAGCCAGGTTTCAG
CGGCAGTGGGTCTGGGACCGATTTACCCCTCACAATTAATCCTGTGGAAGCTGACGATGCTGCAGA
10 TTATTACTGTCTGCAGAGTAAGAATTCTCCGACATTCGGCCAAGGGACCAAGGTGGAAATCAAACGA
GCTGTGGCTGCACC

Clone: kO2-3 (5/26/10) (IGKV2-40, IGKJ2)

GTGCAGCCNCAGCTCGTTTGATCTCCACTTTGGTCCCCTGGCCAAAACCTGTACGGAGGCTGTGTAC
CTTGCATGCAATAATAAATTCCTATATCCTCAGCCTCCACCCTACTGATTTTCAGTGTGAAATCAGCG
15 CCTGACCCACTGCCACTGAACCTGTCTGGGACTCCAGAGGCCCGGTATGAAGCAAAATACATCAGG
AGCTGTGGAGACTGGCCCGGCTTCTGCAGAAACCAATAAAAAATAGGTTTTTCCATTACTATGCAGGA
GGCTCTGACTAGACCTGCAGGAGATGGAGGCCGGCTCTCCAGGGGTGACGGGCAGGGAGAGTGG
AGACTGGGACATCTGGATGTC

Clone: 22-1 (5/10/10) (IGKV1-17, IGKJ2)

GGTGCAGCCACAGCTCGTTTGATCTCCACTTCGGTCCCCTGGCCAAAACCTGTGAGGGTGACCATAAT
20 ACTGTAGACAGTAATAAGTTGCAAAATCTTCAGGCTGCAGGCTGGTGTATGGTGAGAGTGAAATCTGT
CCCAGATCCACTGCCGCTGAATCTTGATGGGACCCAGCTTCCAAAGTGGATGCAGCATAGATCAG
GCGCTTAGGAGGTTTTCTGGTTTCTGCTGATGCCAATTTAAAAAAGTGGAATGCCCTGACTTGCC
CTGCAAGTGATGGTGACTGTGTCTCCTGCAGATGCAGACAGGGAAGATGGAGACTGGGTCATCTGA
ATGTC

Clone: TRPBMCA-3-1-Kappa (IGKV1D-16, IGKJ4)

CCAGACTCCATCCTCCCTGTCTGCATCTGTAGGAGACAAAGTCACCATCACTTGCCGAGCCAGTCAG
25 GATATTAGCAGCTGGTTAGCCTGGTATCAGCAGAAACCAGGGAAAGCCCCCTAAGCCCCCTGATCTATA
AGGCATCCAGTTTACAACTGGGGTCCCTTCAAGGTTTCAGCGGCAGTGGATCTGGGACAGATTTTCAC
TCTCACCATCAGCAGCCTGCAGCCTGAAGATTTTGCAACTTATTACTGTCAACAATATACTAGTTACC
30 CGCTCACTTTTCGGCGGAGGGACCGAGGTGGAGATCAAACGAGCTGTGGCTGCACC

Clone: BMK32 (IGKV3D-15, IGKJ4)

GAAATAGTAATGACTCAGTCTCCAGCCACCCTGTCTTTGTCTCCAGGAGAAAGAGCCACCCTCTCCT
GCAGGGCCAGTCAGAGTGTGAGCAGCAGCTTAGCCTGGTACCAGCAGAAACCTGGGCAGGCTCCC
35 AGGCTCCTCATCTATGGTGCATCCAGCAGGGCCACTGGCATCCAGACAGGTTTCAGTGGCAGTGGG
TCTGGGACAGAGTTCACTCTCACCATCAGCAGCCTGGAGCCTGAAGATGTTGCAGTTTATTATTGTC
AGCAGAATAGTAACTGGCCGCTCACTTTTCGGCGGAGGGACCAAGGTGGAGATCAAA

VH and VL gene names are given using Ig gene nomenclature for *Homo sapiens*.

Table 7. Sequences of Rhesus monkey VH and VL genes obtained with primers listed in Table 3

	V lambda
24.3D6	NNNNNNCCGGCAGCCACCTGAATCCTCGTGTCCAAGTCTCTTGGACAGTCGGTCAC CATCTCCTGCACTGGAACCAGCAGTGACATTGGCGGTTATAATGACGTCTCCTGGTAC CAGCAGCACCCAGGCACAGCCCCCAGACTCCTGATTTATGATGTCAGTAAGCGGCC TCAGGGGTCTCTGATCGCTTCTCTGGCTCCAAGTCTGGCAACACGGCCTCCCTGACC ATCTCTGGGCTCCAGGCTGAGGACGAGGCTGATTACTGCTGCTCATATAGGAGT GGAAGCACTTTCTTATTTCGGAGGAGGGACCCGGCTGACCGTCTAGGTCAGCCCAAG GCTGCCCCCTCGGTCACTCTGTTCCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAG
28A1	CAGACTGTGGTGA CT CAGGAGCCTGCATTGTCTGTGGCCTTGGGACATACAGTCAGA ATGACCTGCCAAGGAGACAGCCTCAAACCCTATTCTACAACTGATGCCATCAGAAGC CAGGCCTTGTCCCTGTACTGGTCGTCAGTGGTTACAATTATCGCCTCTCAGGGATCCC AGGCCGACTCTCTGTCTCCTGGTCAGGACCCACAGGTTCTTGACCCTCATTGCGAC TCAGGGTAGAAGGATGAGGGCTGANCTATTAATGTCAACTCTNTCGCGANCNGGTAG GGCGGTTNCCGCGAGTGTGTTGCGGAAGGTGGCANCCNGAGTGACNTCCNCGTNC GCCNNAGTTC
28D6	cagtctgccctgactcagcctccctcaatgtctgggtctcctggacagtcggtcaccatctcctgtctggaaccagcagtgac attggaattataaccgtgtctcctggtagcaacaacacccaggcaaaagccccaactcatgattatgagtcaccagcagcgg ccctcaggggtctctgatcgcttctctggctccaagtcgtggaacacggcctcctgaccatctctgggtccaggtgaggatg aggctgattattactgcagctcatatgcaggcaggaatacttggatttcggcgaggaggacccgtctgacctccngtngcna ggtcc
28D10	GACTNNGGCNACTAGGAGGTGNAGCCGGGTCCCAGCCACCGAAGATGTAATGATTAC CACTGCTGTCTGCTGATTGACAGTAATAGTCAGCCTCATCTTCTGCCTGGGCCCCAC TGATGGTCAAGGTAAGTGTGTCCTGAACTGGAGCTAGAGAATCGCTCAGGGATCC CTGAGGGCCTCTCACTGTCTTTATATATACACCAGCACAGGGGCCTGGCCTGGCTTCT GCTGGAACCACTGAGCATATTTTTTTGCCAGTATTTCTCCAGAGCAGGTGATCCTGGC CGTCTGTCTGGGGACACTGACACTGAGGGNGNCTGAGTCAGCACANGCTGA
32A7	tcctatgtgtgactcagtcgcccctcggtgtcagcgccctcaggacagacggccaggatcacctgtgggggagacaacattgg aactaaatatgtgactggtaccaacagaagccaacgcaggccctgcgcaggatcattatactgataggaaacggccctca gggatccctgagcgaattctctggctccaactcagggaacaccgccaccctgaccatcagcggggtcagggccgggagtgag gctgactattactgtcaggtgtgggacagtagtagtgatcatgtgttattcggaggaggacccgggtgacctcctaggtcagc ccaaggctgccccctcggtcactctgttcccgccctcctctgaggagctcaagccaacaaggccacactggtgtgtctcataa gtgactctaccgggagccgtgacagtggcctggaa
32C10	cagactgtggtgacctcaggagcctcgtgtgggtgtcccctggaggaaacagtcacactcacgtgtggcctgaactctggctca gtctctgccactaactacgcaggttgatcagcagacccaggccagcctccactcacgtgactacaacacaaagagtcg cccttatgggtctccgatcgcttctctggctccatccttggaaacagagctgtcctcaccatctcgggggtcagacagacgat gaatctgattattactgtgtcctctacctaagactggcagttggatttcggcgaggaggacccgggtcacctgcttaggtcagcc caaggctgccccctcggtcactctgttcccgccctcctctgaggagctcaagccaacaaggccacactggtgtgtctcataagt gacttctaccgggagcgtgacagtggtcctggaangcagatagcagcccggtcaaggcgaggagtggaacacaccac acccttccaaacaagcaacaacaagtagcggccagcagctatctganctgacctgaacagtggaaantccacaga agctacagctgccagggtcacncatgaaggnancagtggaanaana
33B2	cagtctgtgtgactcagccaccctcagtttctggggcccccgggcagacgggtcaccatctcctgcaccgggagcagctcaa catcggggcggtatggtatgttctggtatcaacaattccagggaacggccccaactcctcatcttggaaaataaagcga ccctcagggatttctgaccgattctctggctccaagtcgtgttctcagcctccctgacctcactggactccaatctgaggatgag gctgattattactgtcagcatgggatacctatctcggtgtgttttcgaggaggaggacccgggtcacctgcttaggtcagccca aggctgccccctcggtcactctgttcccgccctcctctgaggagctcaagccaacaaggccacactggtgtgtctcataagtga cttctaccgggagccgtgacagtggcctggaaggc
33B3	tcctatgtgtgactcagccaccctcagtttctggggcccccgggcagacgggtcaccatctcctgcaccgggagcagctcaa catcggggcggtatggtatgttctggtatcaacaattccagggaacggccccaactcctcatcttggaaaataaagcga ccctcagggatttctgaccgattctctggctccaagtcgtgttctcagcctccctgacctcactggactccaatctgaggatgag gctgattattactgtcagcatgggatacctatctcggtgtgttttcgaggaggaggacccgggtcacctgcttaggtcagccca aggctgccccctcggtcactctgttcccgccctcctctgaggagctcaagccaacaaggccacactggtgtgtctcataagtga

	cttctacccgggagccgtgacagtggcctggaaggca
33B6	cagactgtggtgacccaggagccctcactgactgtgtcccccaggagggacagtcactctcacctgtgcttcagttactggagc agtcaccagtgtgacattccacactggttccagcagaaacctggacaagccccaagacactgatttatgatgcaaccaaca cgctctcctggacccctgcccgggtctcagggtccctgtgtgggggtaaagtgtccctgacactttcgggtgagccagccgagga cgaggcagactattactgctgctatactacagtgggtgacgttattcggaggagggacccggctcaccgtcctaggtcagcc caaggctgccccctcggtcactctgttcccgccctcctctgaggagcttaagccaacaaggccacactggtgtgtctcataagt gacttctacccgggagccgtgacagtggcctggaaggcagatagcagccccgtcaaggcgggagtgagaccaccacac cctccaaacaaagcaac
33C6	tcctatgtgctgactcagccaccctcagcatctggggccccgggagaggggtaccatctcctgactgggggaagtccaa catcggggcaaatcggtatgtatcctggtaccaacaactcccaggaaaagcccccaactcctcatcttgaagatgacaagc gacattcaggggtctctggccgattctctggctcaagtctggttccctcagcctccctgacctaccgggctccgatctgaggat gaggctgattattactgctctgctgggatgacactcttattatggtgttcgggggagggacccggctgaccgtcctaggtcagcc caaggctgccccctcggtcactctgttcccgccctcctctgaggagcttaagccaacaaggccacactggtgtgtctcataagt gacttctacccgggagccgtgacagtggcctggaag
33C10	tcctatgtgctgactcagccaccctcgggtgctagtgctccctggacagacggccaagattacctgctctggaatgttctggcaaa atattatgctcattggtaccaacagaagtcaggccaggcgctgtattggtgattataaagacagtgtgagggccttcagggtatc cctgaacgattctctggctccagctcagggaccacagtcaccctgacctcagcggggcccaggctgaggatgagactgact attattgttactctgggatgacaaaaaagtgttattcggaagagggacccggctgaccgtcctagggtcagcccaaggctgccc cctcggtcactctgtcccgccctcctctgaggagcttcaagccaacaaggccacactggtgtgtctcataagtacttctacccg ggagccgtgacagtggcctggaaggcagatagcagccc
33C12	cagctgtgtgctgactcagccaccctcagcgtctggggccccgggaggggggtaccatctcctgactgggagcacctcca acatcggggcggggttattatgtgtctggtaccggcagttcccaggaaacagccccgagctcctcatatgaaagtataagc gacctcaggggttctgaccgactctgtgctcagactgtgttccgggccccctgaccattagtgactccaatctgaagat gaggctgattattactgctcaggatgggaaagcaccctgagaaccgtgttattcggaggagggacccggctgaccgtcctagg tcagcccaaggctgccccctcggtcactctgttcccgccctcctctgaggagcttcaagccaacaaggccacactggtgtgtctc ataagtacttctacccgggagccgtgacagtggcc
33E1	cagctgtgtgctgactcagccaccctcagtgctcaagctcttggacagtcgggtaccatctcctgctctggaaccagcagtga ttggtggttataatggcgtctcctggtaccagcagcactcagggacagccccagactcctgatttatgatgtcagtaagcgcc ctcaggggtctctgatcgtctctggtccaagtctggcaacacggcctccctgacctctctgggtccaggctgaggacgag gctgattatcactgctgctcatataggggtggaagcacttcatctctgctactgggacccggctcacctnaggttagccaagtc ca
33E3	tcctatgtgctgactcagccaccctcgggtgctcagcgccctcaggacagacggccaggatcacctgtgggggagggcacattg gaagtaaataatgtatactggtaccagcagaagccagcgagccctgtgcagggtcatctatgctgataacaacggccctc agggatccctgagcgattctctggtccaattcagggaacaccggccaccctgacctcagcggggtcgaggccggggatga ggctgactattactgtcaggtgtgggacagtagtagtgatcatgtgttattcggaggagggacccggctcaccgtcctagggtcag cccaaggctgccccctcggtcactctgttcccgccctcctctgaggagcttcaagccaacaaggccacactggtgtgtctcata agtacttctacccgggagccgtgacagtggcctggaaggcagatagc
33F4	tcctatgtgctgactcagccaccctcgggtgctcagcgccctcaggacagacggccaggatcacctgtgggggggacaacattg ggagtaggtatgtacactggttccagcagaagccgtcagggccctgtgcagggtcatctattctgatagcaaacgccccctcag ggatccctgagcgattctctggtccaactcaggaacaccggccaccctgacctcagcggggtcgaggccggagatgagg ctgactattactgtcagctgtgggacagtagtagtgatcatcgatattcggcgaggagggacccggctcaccgtcctagggtcagc ccaaggctgccccctcggtcactctgttcccgccctcctctgaggagcttcaagccaacaaggccacactggtgtgtctcataa gtgacttctacccgggagccgtgacagtggcctggaaggcagata
33F12	cagctgtgtgctgactcagccaccctcgggtgctcagcgccctcaggacagacggccaggatcacctgtgggggagaccacattg gaagtaaataatgtatactggtaccagcagaagccagcgagccctgtgcagggtcatctatgctgataacaacggccctc agggatccctgagcgattctctggtccaattcagggaacaccggccaccctgacctcagcggggtcgaggccggagatgagg ggctgactattactgtcaggtgtgggacagtagtagtgatcatgtgttattcggaggagggacccggctcaccgtcctagggtcag cccaaggctgccccctcggtcactctgttcccgccctcctctgaggagcttcaagccaacaaggccacactggtgtgtctcata agtacttctacccgggagccgtgacagtggcctggaaggcagatagcagccccgtcaaggcgggagtgagaccacca caccctcaaaacaagcaacaacaag
33G4	cagtttgtgctgactcagccaccctcagtgctcagtgctcccccaggacagacggccaggatcacctgntntggagaaatcctggc aaaaaaatgtgctcagtggttccagcagaagccagggccagggccctgtgtgtggtatataaagacagtgtgagggccctca gggatccctgagcgattctctagctccaggtcagggacaacagttaccttgacctcagtggtggccagggcag aagatgaggctgactattactgtcaatcagcagacagcagtggtatcatccggtattcggcgaggagggacccggctgacctcc tagtngccnaggtcc

33H7	tcctatgtgtgactcagccaccctcagtgctcggtgtccccaggacagacagccagcatcacctgctctggagataaatttggg aatgcataatgcttactgtgtaccagcagaagccaggccaggtccctgtactggtcatctataaaaaatagcaaccggccctcagg gatccctgagcgattctctggctccaactcggggaacacggccaccctgaccatcagcggggtcgaggctggggatgaggc cgactattactgtgtacaagctatggcagtgaggagcagctggcagctacatcttcggtgtctgggaccggctcaccgtctagg tcagcccaaggctgccccctcggtcactctgttcccgccctctctgaggagcttcaagccaacaaggccacactggtgtgtctc ataagtgtacttaccgggagccgtgacagtggtggaaggcagatagcagcccgtaaggcgggagtgaggaccac cacaccctccaaacaaa
24.3A5	gcaggctctgaatcactcgccgacctcgtgtctcgggccccagggcagggggtcaccatctctgctctggaagcagctccaa tatcgggagaaattatgtatctgtgtaccagcaggtccagggaacggccccagactcctcatctatcaagataataaccgac cctcaggggttctgaccgattctctggctccaagtctgtgtacccagcctccttgccatcactggactccagactggagatga ggctgatttactgctgagcctgggataccagcctgagtgctcggtcttcggcgaggaggaccggctgaccgtcctaggctc agcccaaggctgccccctcggtcactctgttcccgccctcctctgaggagcttcaagccaacaaggccacactggtgtgtctca taagtgtacttaccgggagccgtgacagtggtcgtggaaggcagatagcagccccgtaaggcgggagtgaggaccacca cacctccaaacaaagcaacaacaagtagcggccagcagctatctgagcctgacgctgagcagtggaagtcacacag aagctacagctgccaggtcacgcataagggagcaccgtggagaagacagtgccccctacagaatgttcatagctagct ggccagacatgataagatacattgatgagttggacaaaccacaactagaatgcagtgaaaaaatgctttattgtgaaatttg tgatgtattgtcttattgttaaccattataagctgcaataaacaagtttaacaacaacaatgcatcttattgtttcaggttcaggg ggagggtgtgggaggttttttaagcaagtaaaacctctacaaatgtggtatggaatttctaaataacagcatagcaaaa ctttaacctccaaatcaagcctctacttgaatccttttctgaaggaagaataaggcataagcatcaggggctgttgccaatgtgca ttagctgtttgcgcccactctctttcaggagtttaaaattatgtattttcccaagggttgaaacacttctcttttaggtttaaaggc actacctccccatttcttttataaaaattcaaaataatttaaaacattcatgccaagaaataaagggtttttataggcaaatcca aagcctcaggg
28A1	CAGACTGTGGTGACTCAGGAGCCTGCATTGTCTGTGGCCTTGGGACATACAGTCAGA ATGACCTGCCAAGGAGACAGCCTCAAACCCTATTCTACAACTGATGCCATCAGAAGC CAGGCCTTGTCCCTGTACTGGTCGTCACTGGTTACAATTATCGCCTCTCAGGGATCCC AGGCCGACTCTCTGTCTCCTGGTCAGGACCCACAGGTTCTTGACCCCTCATTGCGAC TCAGGGTAGAAGGATGAGGGCTGANCTATTAATGTCAACTCTNTCGCGANCNGGTAG GGCGGTTNCCGCGAGTGTGTTGCGGAAGGTGGCANCCNGAGTGACCNCCNCGTNC GCCNNAGTTC
28D6	cagctgtcccctgactcagcctccctcaatgtctgggtctcctggacagtcggtcaccatctcctgctctggaaccagcagtgac attggaattataaccgtgtcctgtgtaccaacaacaccagggaagcccccactcatgatttatgagtcactgagcgg ccctcaggggtctctgatcgtctctgtgctcaagctgtggaacacggcctcctgaccatctctgggtccaggctgaggatg aggctgatttactgcagctcatatgcaggcaggaatactttggtatttcggcgagggaaccgctgacctcngngtngcna ggtcc
V kappa	
23.2F6	TGACATTGTGCTGACCCAGTCTCCCATCAGGCTGACATTGTGATGACCCAGTCTCCCT CCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACTATCACTTGCCGGGCAAGTGAGA GCGTTAACAACATATGTAAATTGGTATCAGCAGAGACCAGGGAAAGCCCCCTAAGCTCCT TATCTATAGGACATCCACTTTACAAAGTGGGGTCCCGTCAAGGTTACCCGGCAGTGGA TCTGGGGCAGATTACACTTTACCATCAGCATCCTGCAGCCTGAGGATGTTGCAACTT ATTACTGTCAGCATACTTATACTGCCCCATTCACTTCGC
24.1B7	caagtgtggtgactgtgtctctacagatgcagacagggaggatggagacagggctcatcaaatgtcaagggtggttcgg gcnthtgcattatttggcctggtctccgagccaggggaaggctccgctgcgctgttctattaccatccactttgaaagt gggtcacatcgaggttcagcggacgtgatttaggggttgattccgctctgctctgcagctctggagtgtgattgatggacaaatgt ccctgtgtatccctgatggactacaaagatgcaagatacagttctacttctcttaccagcaangtgnagattttanagctttgtc ctactactcacagtggttgggtgggtcaaaatgttctacttcatcttttactgactcaatccctcctcaatgatgatttgggtctggcg tcaccgaaacgacatttgggacctaacttttgcacacctcaatataagtggaaaacggaaaactacagattttaccgctac ctttaataaactgcttttagagtcgaattgtccaagaatcaatttattgatttctcctcctgaccgaccgatagggtccacaagaa taaatttacgcaggtatccgccaagagtcgaatttcaacaacctcgtgttagaagtcggccccctccgattgattaaagaacc cccaggtttaccctacgttaaatgaaagatttaaaaacacggggttagagtaaccgcaagttgctatgacctgggggca aaaatttctcgctaattaggctccctcccaatnaaaacacgggttgagcaatataacgnaccgggccccagggcgga cacgtttccccaagaaaaaccccaaaaaaaacacgggttgagcaatataacgnaccgggccccagggcgga gtgattaaatccctaaaaaagggtcccttcccccttttaaaantcttttgggaagggtggttgggaaaaacaaaaa naaaaancccnnaattttctgaanacttttccgcggggttttagaggagacgagancggnnaaaatctttcttannnn

- 89 -

24.3	CAGGTGCAGCTGCAGGAGTCGGGCCAGGACTGGTAAAGCCTTCGGAGACCCTGTCCCTCACCTGCGCTGTGCGC TGGTGGCTCCATCAGCAGGAACACTACTGGAGCTGGATCCGCCAGCCCCAGGGAAGGGACTAGAGTGGATTGGAC GCATCTCTGGAAGTGGTGGGACCACCGACTACAATCCCTCCCTCAAGAGTCGAGTCTCCATTTCAACAGACAGG TCCAGGAACCAGTTCTCCCTGAAACTGTCTTCTGTGACCGCCGCGGACACGGCCGTGATTACTGTGCGAGAGA CGCCCCAGTTTGGAGGGGTGACTACTGGGGCCAGGGAGTCTGGTACACCGTCTCCTCAGCTAGCTTGGTGCTAG CACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCT GCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCTGTGGAACCTCAGGCGCCCTGACCAGCGCGTGCAC ACCTTCCCGGCTGTCTACAGTCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTT GGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCC AATCTGTGACAAAACCTCACACATGCCACCGTGCCAGCACCTGAACCTCTGGGGGGACCGTCAGTCTTTCTC TTCCCCCAAAACCAAGGACACCTCATGATCTCCCGGACCCCTGAGGTCCCATGCGGTGGTGGTGGACCTGAG CCCCAAGAACCCTGAAGTCAAGTTCACCTTGGTACCTGGACGGCGTGAAAGTTCATAATGCCCAACAAACCCC CGGGGAGGACCATTACAACCCCCCGTTCCCGGTGGTTACAGGTCTTACCGTTCCGGAACAGAAATGGGTTAA AGGCCAAGGAATTAATTAAGGAAAGTTTCCCAAAAAAGCCCTTCCCCACCCCCATTGGAAAAAACCCCTCTCC CAAACCAAGGGGGACCCCCCAAAAAACAAAGGGGGTAAACCCCGGCGCCCCCTCCCGGGGGGGGAGAATACC AAAAAACCCGGTTCCCCCTGAACCCCCGCGGGGAAAAAAGGGTTTTTATTTCCCAAAAAGAAAACCTCCCCGGGGGA
24.3	CTGGTGCAGCTGGTGGAGTCTGGCCCAGCAGTGGTGAAGCCCTCGGAGACCCTGTCCCTCACCTGCGCTGTGTC TGGTGGTCCATCAGCAGTGGTCACTGGTGGACCTGGATCCGCCAGTCCCGAGGGAAGGGACTGGAATGGATTG GGGGTGTCTATGGTCCCGGTGAGAGCACCCATAACAACCCCTCCCTCAAGAGTCGAGTCAACATTTCAAAAGAC ACGTCCAAAACAGTTCTCCCTGAAGCTGAACCTCTGTGACCGCCGCGGACACCGCCGTGTTTTACTGTGCAAG AGCCACATTCTGTAATGATCGTGGCTGCTCATGGGGCTACAACCTGGTTCGATGTCTGGGGCCCGGAGTCTCTGG TCACCGTCTCCTCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCCTCCTCCAAGAGCACCTCTGGG GGCACAGCGGCCCTGGGCTGCCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGTCTGTGGAACCTCAGGCGC CCTGACCAGCGCGTGCACACCTTCCCGGCTGTCTACAGTCTCAGGACTCTACTCCCTCAGCAGCGTGGTGA CCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCAGCAACACCAAGGTG GACAAGAAAGTTGAGCCCAAATCTGTGACAAAACCTCACACATGCCACCGTGCCAGCACCTGAACCTCTGGG GGGGAGCTCAGTCTTCTCTTCCCCCAAAACCCAAAGGACACCCCTCATGATCTCCCGGACCCCTGAAGTCACAT GCCGTGTGGGTGGACCGTGACCCACAAAACCCCTGGAGGTCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT TAATGCAAAACAAACCCCGGAGGAGCCTTAAACCGCCCTTACCGTTTGGTTCAGCGTCTCTCCGTCCTGGC CAGAAATGGGTAAATGCAAGGATTCACATTTCAAGGTTCCAAAAAACCCCTTCCACCCCATCGGAAAAAACCCCT CTCCAAAGCCAAAGGGCACCCCCAAAAACCCAGGGGTATACCCTGCCCTCCTCCGGAGGAAAAAACACAAAAC CAAATTAACCCCTAACTGCCCGGTATAAGGTTTTAATCTCCACAAATTCCTGGGAAAGTGGAACAATTTGGGA ACCCAGAAAACATACACCCCTCCGGTGAATTCAGAGGTCTTCT
24.3	GCAGCTGCAGGAGTCAGGCACACCTTCCAGGCTGTCCAGCCCGGGGGGTCCCTGAGACTCTCCTGTTTCACTCTC TGGATTGCGCTTCAGCGAATATGGCATGCAGTGGGTCCGCCAGGCACCCAGGGAAGGGGCTGCAGTGGATCGCGT CTATTAATAAGGTGCGCGATAAGACAAGGTTGCGAGACTCCGTGAAGGGCCGACTCACCTCACCAGAGACAAC TCAGATTACACTCTCTCCCTGCAAATGAACAGCCTGAAAATTTAGGACACCGGCCCTGTATTATTGTGCGACGGG GAGGATGGACTACTGGGGCCGGGGAGTTCTGGTCAACCGTCTCCTCAGCTAGCACCAAGGGCCCATCGGTCTTCC CCCTGGCGCCCTCCTCCAGGAGCACCTCCGAGAGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCT GAACCCGTGACCGTCTCCTCAGCTAGCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCCTCCTCCAAGAGCAC CTCTGGGGGCACAGCGGCCCTGGGCTGCCCTGGTCAAGGACTACTTCCCGAACCCGTGACGGTGTCTGTGGAAC CAGGCGCCCTGACCAGCGCGGTGCACACCTTCCCGGCTGTCTACAGTCTCAGGACTCTACTCCCTCAGCAGC GTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCAGCAACA CCAAGGTGGACAAGAAAGTTGAGCCCAAATCTGTGACAAAACCTCACACATGCCACCGTGCCAGCACCTGAA CTCCTGGGGGGGACCGTCACTCTTCTTCCCCCAAAACCCCAAGACACCCCTCATGATCTCCCGGACCCCTG AAGTCCCATGCTTGGGTGGTGGACCTTGAGCCACAAAACCCCTGAGGTTTCACTTACCTGGTTACCTGGACGG GCTGGAAGGTGCAAAATTGCCAAAACAAAGCCCGGGGAGGAAGCTTACAACACCACCTACCGGGGTGGTAAAGGT TCTTACCCCTCCTGGAACCAAAATGGCTGAATTTGGAAGGAATACATTTGTCAGGGTCTCCACAAAGCCCTCCCC GCCCCCTTCGAAAAACACTTTCCCAAGCCAAAGGGGACCCCCAAAAAACAGGGGTAAACCCCGCCCTCCTCC GGGGAGAAATAACCAAAACAGTCCCCCGTCACTCCGGGGAAAGAGCTTCTTCCCCCAACCC
28A1	GAGGTGCAGCTGGTGAATCTGGAGCAGAGGTGAAAAGGCCCGGAGAGTCTCTGAG GATCTCCTGTAAGACTTCTGGATACGATTTTTCCACCTCCTGGATCAGCTGGGTGCGC CAGATGCCCGGGAAAGGCCTGGAGTGGATGGGAATGATCTATCCTGATGATTCTGAT ACTAAATACAGGCCGTCTTGAAGGCCAGGTCAACATGTGAGTCGACCAAGTCCATCA GCACCGCCTTCTGCAATGGGGCAGCCTGAAGGCCTCGGACACAGCCACCTATTACT GTGCAAAAGGGGGATCCTTTTTACGACAATGATTACGGTACTATTTTTCCGGCCCTG GTTGATGTCTGGGGCCCGGAGTCCCGGTACCGTCTCTTACGCTCCACCAAGG GCCCATCGGTCTTCCCCCTGGCGCCNTCCTCCAGGAGCACCTTCCGAGAGCACAGC GGCCNTGGGCTGCCTGT

28D6	ggtgcnaagctggtgcaatcngggcgggcttggtccagcctggcgggccctgagactctcctgtgcagcctctggattcatct tcagtaactcctggatggcctggttccgcaggctccaggaaggggctagagtgggtagcccgtattaaaaggaagccga tggtgagaaggtgggctatgccgcatctgtgaacggcagattcaccatctcaagagacgattcaaaagacacagtgacctgc aaatgaatagcctgaaaagcgaggacacggcgtctatttctgcaccacagataagaggtggaacgacgcgttcgatgtctg gggcccgggagtcctggtcaccgtctcctcagcctccaccaagggccatcggtcttccccctggcgccctcctccaggagca cctccgagagcacagcgccntgggctgctgtcnangac
28D10	aggtgcaagctggtgcaatctggggaggcgttggtccagcctggcgggccctgagactctcctgtgcagcctctggattcacc ttgatgattatgccatgcactgggtccgccaagctccaggaagggctggagtgggtatctaggattagttggaatagtggtac catatactacgcagactctgtgaaggccgattcaccatctccagagacaacgccaagaactccctgtttctgcaaatggaca ggctgagagctgaggacacggcgtgtattactgtactagtaggatagcggcagcttttgactactggggccaggagtcctg gtcaccgtctcctcagcctccaccaagggccatcggtcttccccctggcgccctacctccaggagcacctccgagagcacag cgccccctgggctgctgtcnaagata
32A7	ggtgcaagctggtgcantctggggcgggcttgccaaagcctgggggatccctgagactctcctgtgcagcctctggattcacc tcagttcctattctatgcactgggtccgccaggctccaggaaggggactggagtgggtctcagctattagtagtggtgtaacac atactacgcagactccgtgaaggccgattcaccatctccagagacaactcaaaagacacgctctccctggaaatgaagag cctgagacctgaggacacggcgtttattactgtgcaaaaaccagaggtggaacaacggggagcgactacggttttgattctc ggggccaaggggtcgtcgtcaccgtctcctcagcctccaccaagggccatcggtcttccccctggcgccctcctccaggagc acctccgagagcacagcgccctgggctgctgtcnagacnntcc
32C10	aggtgcaagctggtgcantctggggaggcgttggtccagcctggcgggggtccctgagactctcctgtgcagcctctggcttatgt tcagtacacatgatatgcattgggtccgccaacctccaggaaggggtctggagtgggtctcagttatcagtgaagaggtgata ataagtactacgcggactctgtgaaggccgattcaccatctccagagacaacgccaagaactcttggttctgcaaatgaac agcctgagggccgaggacacggcgtctatttgtgttagtaagactgcagtgactcttctgttctccgcttggacaactgg ggcctgggctcctggtcaccgtctcctcagcctccaccaagggccatcggtcttccccctggcgccctcctccaggagcacc tccgagagcacagcgccctgggctgctgtcnaggacnttc
33B2	gaggtgcagctggtgcagctctggagttagaggtgaaaaggccggggatttctgaggtatccctgtaagacctctggagaca ctattaccagcctctggatcaactgggtgcccagatgcccgggaaaggcctggaatggatgggaatgatctacctgaagat tctgatgtcacatttagtccgtcttccaaggccagggtcaccatgtcagtcgacaaatccatcaataccgctacctgcgggtgga acagcctgaaggcctcggaacccgccacgtattattgtgctcacctacctgtgatgatagtggtgtctgcaaatcttgactac tggggccaggagtgccagggtcaccgtctcctcagctagcaccaagggccatcggtcttccccctggcacctcctccaaga gcacctctgg gggcacagcgccctgggctgctgtgtaaggactactccccgaaccgggtgacgggtgc
33B3	gaggtgcagctggtgcagctctggggaggccttggtccagcctggcggggtccctgagactctcctgtgcagcctctggattac tcagtactctgcttgacctgggtccgcccaggccttgaggagggcctggagtgggttggtctatattagaagcaagtcatacaa tctcgagacagcataccgctcggtgaaaggcagattcaccctcctcaagagatgattcaagaacatggatttctgcaaat gagcagcctgaaaaccgaagacacggcgtctatttctgttctgggggtgggtgcacactgactcctggggccaggagtcct ggtcaccgtctcctcagctagcaccaagggccatcggtcttccccctggcacctcctccaagagcacctctggggccacag cgccctgggctgctgtgtaaggactactccccgaaccgggtgacgggtgctgtggaactcaggcgccctgaccagcggtg gcacaccttcccggtgtctacagctcctcaggactctactccctcagcagcgtggtgacgtgccctccagcagcttgggcac ccagacctacatctgcaacgtgaatcacaagcccagcaaac
33B6	nggtgcaggctggtgcagctctggggaggccttggtacagcctgggggtccctgagactctcctgtgcagcctctggattcacc ttcagtaactactacatacactgggtccgcccaggctcaagggaagggctagagtgggttaggttaataagaagtaaaactga cagttccacgacagagatgtgctgggtggtggaaggcagattcaccatcgcaagagatgattcaagaacacgctgtatctgc aaatgagcagcctgaaaaccgaggacacaggcgtatatttcttctcattcattagtcactggggccaggagtcctgtgca cgtctcctcagcctccaccaagggccatcggtcttccccctggcgccctcctccaggagcacctccgagagcacagcggc cctgggctgctgtcnagac
33C6	caggtgcagctggtgcaatccggaccagaggtgaagaggcccgagactctctgacgatccctgtaagacctctggattctc ctctaactaccttgatcaactgggtgcccagatgcccgggaaaggcctggagtggatgggaatgatctatcttggtgactct gacaccacatatagtccgtccttccggggccaagtcaccatgtcagtcgacaaatccatcaagaccgctacctccagtgggg caacctgaaggcctcggaacccgcaacttattactgcgctctcggactgtaatgtagtggtgtctgcaaatcttgattactg ggggcaggagtgccagggtcaccgtctcctcagctagcaccaagggccatcggtcttccccctggcacctcctccaagagc acctctgggg
33C10	AGGTGCANGCTGGTGCANTCTGGAGGAGGCTTAATTACGCCGGGGGGGTCCCTGAG ACTCTCCTGTGCAGCCTCTGGAATCACCTTCAGGAAGTATGGCATACTGGGTCCGC

	CAGGTTCCAGGGAAGGGGCTGGAGTGGGTGGCACGTATATCGTCTAATGGAAATAAC AAATACTACGCAGACTCTGTGAAGGACCGATTACCATCTCCAGAGACAATTCTAAGA ACATGCTATATCTTCAAATGAATAACCTGAAATTGGGGGACACGGCCATTTATTACTGT ACGAGAGATTTCGACAGAGCACATTTTTGACTCCTGGGGCCAGGGAGTCCTGGTCACC GTCTCCTCAGCCTCCACCAAGGGCCCATCGGTCTTCCCCCTGGCGCCCTCCTCCAGG AGCACCTCCGAGAGCACAGGCGGCCCTGGGCTGCCTGTCAAGGAC
33C12	agggtcaggctgggtgcantctgnggctgaggtgaagaagcctggggcctcagtgaagcttacctgaaggcttctggattctcc ttcgccaactttgccatcagctggatccgacaggcccttgacaaggacttgagtgatgggagggatcgctccctctattggtg tgacaaactacgcacagcagttccgggacagactcacgataaacgcgacacgtccacaggcacagcctccctggaactg atcagcctcagatctcaggacacggccatttattactgtgcgagagaaggggtgacctccaactcattcgatgtctggggccgg ggagttctggtcaccgtctcctcagcctccaccaagggcccatcggtctccccctggcgccctcctccaggagcacctccgag agcacaggcggccctgggctgctgtcaaggaca
33E1	nggtgcangctgggtgcaatctgcccaggactgggaagccttcggagaccctgtccctcacctgcgctgtctctgggtgctcc atcagcactaactactggacctggatccgccaagcccccagggaaggactggagtgattggacgtctctctggtattggtg gaacatcgactacaacccctccctcaagaggcagtcaccatttcaacagacacgtccaagaaccagttctccctgaagctg aactctgtgaccgcccggacacggccgtgtattactgtgcgagagatccgcatgtgagacacagtgatacagttatcttgac tactggggccaggagtcctggtcaccgtctcctcagcctccaccaagggcccatcggtctccccctggcgccctcctccag gagcacctccgagaggcaccagcggccctgggctgctgtcaaggac
33E3	tcctatgtgctgactcagccacccctgggtgacggcctcaggacagacggccaggatcacctgtgggggagaccacattg gaagtaaataatgtatactggatccagcagaagccagcgagccctgtgcaggtcatctatgtctgataacaaacggccctc agggatccctgagcgttctctggctccaattcagggaacaccgcccaccctgaccatcagcggggtcgaggcggggatga ggctgactattactgtcaggtgtgggacagtagtagtgatcatgtgttattcggaggaggagaccgggtcacctgctaggctcag ccaaggtgccccctcggtcactgttcccgcctcctctgaggagcttcaagccaacaagggcacactggtgtgtctcata agtgacttctacccgggagccgtgacagtgccctggaaggcagatagcagcccgtaaggcgggagtgagaccacca caccctccaaacaagcaacaacaagtac
33F4	cagggtcagctggtgcaatctggggcagacttggtccagcctggcggtccctgagactctcctgtgcagcctctggattcacct tcagtc agtcctggatgagctgggtccgccagggtccagggaaggggctagagtggttggccgtattaaaaggaaagctgatggtga gacaacagattacgccgatctctgaaaggcagattcacatctcaagagacgattcaaaggacacactgtacctgcagatg aatagcctgaaaaccgaggacacggccgtatattactgcaccgcgacaccgcattctacggaatactggggccaggagtc ctggtcacctgtcctcagcctccaccaagggcccatcggtctccccctggcgccctcctccaggagcacctccgagagcac agcggccctggctgctgtcnangacnttnc
33F12	agggtcangctgnggcagtcngggccaggctggtgacgccttcggagaccctgtccctcacctgcactgtctcggttggtc catcagcggctactattggagttggattcgtaattcccaggatgggactggagtgattggacacgtcgggtggaactactgg gaccaccttctacacccctccctcaagagtcgggtctccatttcggcagtcctccgaagaaccagtttccctgacgttgacct ctgtgaccgcccggacacggccgtctatttctgtgcgagaatcgagcctggctattggggccaggagtcctcgtcacctgt cctcggcctccaccaagggcccatcggtctccccctggcgccctcctccaggagcacctccgagagcacagcggccctggg ctgctgtcaagaca
33G4	cnggtgcagctggtgcaatctggggcaggcttggtccagcctggcggtccctgagactctcctgtgcagcctctggattcatctt cagtaactcctggatgagctgggtccgccagggtccagggaaggggctagagtggttggccgattaaaaggaaagctga agggtgagacaacagattacgccgatctgtgaaaggcagattcacatctcaagagacgattcaaagaacacagcttaccta caaatgaacagcctgaaaagcaggacacggccgtgtattactgcaccacagccaaaaggacacagtggtcctcacta ctggggccaggagtcctggtcaccgtctcctcagcctccaccaagggcccatcggtctccccctggcgccctcctccagga gcacctccgagagcacagcggccctgggctg

Table 8. Rhesus monkey IgHV Primers for Leader Region (Set I)

Group	Note	Sequence	~Tm °C	Product Size
Mamu IGHV1	Ta 60	1- ATGGACTTGACCTGGAAGATCCTCCTCT -28	82.51 (69)	519

Group	Note	Sequence	~Tm °C	Product Size
Mamu IGHV2	Ta 61	3- GGACACGCTTTGCTCCACGCTCCTGCTGCTG -33	91.6 (84)	518

Group	Note	Sequence	~Tm °C	Product Size
Mamu IGHV3	Ta 60	11- GGAGTTTGTGCTGAGCTGGGTTTTC -39	83.4 (67)	

5

Group	Note	Sequence	~Tm °C	Product Size
Mamu IGHV4	Ta 61	4- AAGCACCTGTGGTTCTTCCTCCTCCTGG -31	84.7 (73)	514

Group	Note	Sequence	~Tm °C	Product Size
Mamu IGHV5	Ta 61	5- ATGGGGTCAACTGCCATCCTCACC -28	84.7 (70)	518

Group	Note	Sequence	~Tm °C	Product Size
Mamu IGHV6	Ta 61	1- ATGTCTGTCTCCTTCCTCATCGTCCTC -31	86.1 (68)	522

Group	Note	Sequence	~Tm °C	Prod uct Size
Mamu IGHV7	Ta 60	1- ATGGACCTCACCTGGAGCATCCTTTTC -31	84.5 (71)	520

10

Table 9. Rhesus monkey IgHV Primers for Leader Region (Set II)

Group	Note	Sequence	~Tm °C	Primer Size
Mamu IGHV1	Primer #1	73-GAGAAGCGAATTGAGTCCAGTCAAGGGGACG-103	85.1 (77.5)	30
Mamu IGHV1	Only for one gene	72-GGAGGCAGAAACCAGGCCAGTCAAGTGAGAC-102	85.2 (77.6)	30
Mamu IGHV1		30-GGTGACAGCAGCCACAGGTAAGGGGTCCCCAGG-62	92.3 (84.7)	32
Mamu IGHV1	Only for one gene	30-GGTGGCAGCAGCAACATGTAAGGGGCTCCCAGG-62	92.8 (85.2)	32

Group	Note	Sequence	~Tm °C	Primer Size
Mamu IGHV2	Primer #2	14-GCTCCACGCTCCTGCTGCTGACCRTECC-41	86 (78.4)	27
Mamu IGHV2		47-GTGAGTGCTGTGRTCARGSACTMMYTCAGGG-77	69.8 (62.2)	30
Mamu Consensus for IGHV2		127-CCACAGGGGTCTTSTCCCAGGTSACCTTG-155	79.5 (71.9)	28

Group	Note	Sequence	~Tm °C	Primer Size
Mamu IGHV3	Fw starts at 207	189-GCAGGTGYSCAGKGTGAGGTGCAGYTGGTGGAG-221	81.4 (73.8)	32
Mamu IGHV3	Primer #3; Skipped 1 genes ordered	106-ATGAGTGAGAGAAACAGTGGATATWTGTGGCA-137	77.5 (69.9)	31

Group	Note	Sequence	~Tm °C	Primer Size
Mamu IGHV4	Fw starts at 141	33-GGCAGCTCCCAGATGTGAGTGTCTCGGGGATG-64	90 (82.4)	31
Mamu IGHV4	Primer #4	82-GAGGCTGTCTCTGATCCCAGGGCTCACTGTGGG-114	90.2 (82.6)	32

5

Group	Note	Sequence	~Tm °C	Primer Size
Mamu IGHV5	Primer #5; Fw1 starts from 141	3-GGGGTCAACTGCCATCCTCACCCTCCTACTGG-34	88.4 (80.8)	31
Mamu		120-CCCCACAGGAGTCTGTGCCGAGGTGCAGC-149	91.6	29

IGHV5			(84)	
-------	--	--	------	--

Group	Note	Sequence	$\sim T_m$ °C	Primer Size
Mamu IGHV 7	Primer #7; Fw1 starts from 144	60-AGGTCCCAGGGCGGAGGCAGAAACCAG-86	86.8 (79.2)	26

Group	Note	Sequence	$\sim T_m$ °C	Primer Size
Mamu IGHV 6	Primer #6; Fw1 starts from 144	70-GTATTCACAGCAGCATTACAGAGTGAGGGG-100	81.7 (74.1)	30

Table 10. Rhesus monkey IgHV Primers for Leader Region (Set III)

Group	Note	Sequence	~Tm °C	Primer Size
Mamu LVH1.1	all	1- ATGGACTTGACCTGGAAGATC-21	62 (54.3)	21
LVH1.2	Only one	1- ATGGACCTCACCTGGAGCATC -21	67.5 (60)	21

Group	Note	Sequence	~Tm °C	Primer Size
Mamu LVH2	all	5- ACACGCTTTGCTCCACGCTCC -25	74 (66.3)	21

Group	Note	Sequence	~Tm °C	Primer Size
Mamu LVH3.1	Two genes	1-ATGAGGAAAAGTCAGTGTGAGCC-23	67 (59.4)	23
Mamu LVH3.2	One gene	1- ATGAGTGAGAGAAACAGTGGATATT -25	63.5 (56)	25
Mamu LVH3.3	Three genes	1-ATGGAATTTAGGCTGAGCTGG- 21	65.8 (58)	21
Mamu LVH3.4	Three genes	1-ATGGAGTTGGGGCTGAGCTG- 20	68.8 (61.2)	20
Mamu LVH3.5	Three genes	1-ATGGAGTTTGTACTGAGCTGGT- 22	62.5 (55)	22
Mamu LVH3.6	Eight genes	1-ATGGAGTTTGGCCTGAGCTG-20	66.5 (59)	20

Group	Note	Sequence	~Tm °C	Primer Size
Mamu LVH4	All	1- ATGAAGCACCTGTGGTTCTTC-21	63.2 (55.6)	21

Group	Note	Sequence	~Tm °C	Primer Size
Mamu LVH5	All	3- GGGGTCAACTGCCATCCTCAC -23	70.3 (62.7)	21

5

Group	Note	Sequence	~Tm °C	Primer Size
Mamu LVH6	All	1- ATGTCTGTCTCCTTCCTCATCG-22	64.8 (57.2)	22

Group	Note	Sequence	$\sim T_m$ $^{\circ}\text{C}$	Primer Size
Mamu LVH7	All	1- ATGGACCTCACCTGGAGCATC-21	67.5 (60)	21

delivered by the synthesizer).

Table 11. Rhesus monkey VH genes amplified with leader region primers from Tables 8-10 wherein individual subfamily members shown for each VH gene family are numbered according to the corresponding human subfamily name

5

	Set I (Table 7)	Set II (Table 8)	Set III (Table 9)
Primer for subfamily	IGHV1	IGHV1 1-2 1-46	IGHV 1.1 1-46-01 1-8 1-58-01 1-2 1-69-03 1-68-01 IGHV 1.2 1-f-01
	IGHV2 2-70-10 2-70-04 2-5-04 2-5-10 2-70-12 2-5-08	IGHV2 2-70-11 2-5 2-70-10 2-70-04	IGHV2 2-5-02 2-10-01 2-70
	IGHV3 3-66-02 3-64-01 3-74-03 3-48 3-43-02 3-h 3-74-02 3-72 3-16-02 3-23	IGHV3 3-66-02 3-23-04	IGHV3.1 NA IGHV3.2 3-66-02 3-61 3-53-01 3-4 3-15-06 3-b 3-74 3-28 3-h 3-72 3-23 IGHV3.3 3-72 3-66 3-15 3-64 3-13 3-47 3-74 IGHV3.4 3-h 3-21 3-48 3-9 3-30 3-23 3-NL1 3-15 3-68 3-73 3-20 3-72 3-33 IGHV3.5 3-73 3-70 IGHV3.6 3-49 3-74 3-66 3-43 3-72 3-h 3-20 3-23 3-NL1 3-73
	IGHV4 4-30-2-04 4-4-07 4-61-05 4-59-10 4-39-07	IGHV4 4-30-4-06	IGHV4 4-39-07 4-4-06 4-4-07 4-61 4-34-07
	IGHV5 5-51-01 5-a-04 5-a-01	IGHV5 NA	IGHV5 5-51-05 5-51-02 5-51-01 5-a-04 5-a-01
	IGHV6 6-1-10	IGHV6 6-1-10	IGHV6 6-1-01
	IGHV7 7-43-01 7-81	IGHV7 7-81 2-70	IGHV7 7-81-01 7-34-1-02 7-4-1-03
	Mix 2-70 2-5	Mix 3-66-02 4-30 3-23-04 2-5 2-70-11 5-61 2-70-10 4-39 2-70-04 4-4	Mix 2-70-3 5-51 7-81-01 3-48 3-19-01 3-74 3-15-04 3-11 4-59-04

Table 12. Rhesus monkey immunoglobulin heavy chain germline V genes

Gene family	Cognate Human gene name	Rhesus monkey genome contig	Sequence
IGHV1	L22582	gnl Mmul_051212 Chr7:169628815-169629993	atggacttgacctggaagatcctcctcttggtgacagcagccacaggtgaaggggtccccaggtctcatt cttgagaagcgaattgagtcagtcaggtcaaggggacgttcactcctgtcttctctccacaggtgccc ctccCAGGTGCAGCTGGTGCAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAAC CACATCCTGAGAGTGTGAGAAACCCAGGAAGGAGGCAGCTGTGC
IGHV1	M29809	gnl-Mmul_051212-Chr7-169628837-169630326	atggacttgacctggaagatcctcctcttggtgacagcagccacaggtgaaggggtccccaggtctcatt cttgagaagcgaattgagtcagtcaggtcaaggggacgttcactcctgtcttctctccacaggtgccc ctccCAGGTGCAGCTGGTGCAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAAC CACATCCTGAGAGTGTGAGAAACCCAGGAAGGAGGCAGCTGTGCTGGCAT
	M99637	RCD gnl-Mmul_051212-Chr7-169628838-169630125	atggacttgacctggaagatcctcctcttggtgacagcagccacaggtgaaggggtccccaggtctcatt cttgagaagcgaattgagtcagtcaggtcaaggggacgttcactcctgtcttctctccacaggtgccc ctccCAGGTGCAGCTGGTGCAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAAC CACATCCTGAGAGTGTGAGAAA
	M99641	RCD gnl-Mmul_051212-Chr7-169767710-169769045	atggacctcacctggagcatcctttcttggtggcagcagcaacatgtaaggggtccccaggtcccagg gcggaggcagaaaccaggccagtcagtgagacttcacgaatcctgtgtcctctccacaggtgccc ctccCAGGTGCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGCAGCCTGGGGCC TCAGTGAAGGTCTCCTGCAAGGCTTCTGGTTACAGCTTACCACCTATGGTAT GAATTGGGTGCGACAGGCCCTGGACAGGGCCTTGAGTGGATGGGATGGAT GAACACCTACACTGGGAACCCAACGTATGCCAGGGCTTACAGAACGGTTT GTCTTCTCCATGGACACCTCTGTACGACAGTGTACCTGCAGATCAGCAGCCT AAAGGGTGAGGACACGGCCGTGTATTACTGTGCGAGATACACAGCTGTGGAA ACCCACATCCTGAGAGTGTGAGAAGCCCTGAGGGAAGAGGTAGCTGTGCTGA GCTGAGGCAGTGGTGCAGCAGTTTTCTAAACTTCCGTATGATCTCATTTTGCA TTGACTTCCACTTTTATATTAGCCAA
	M99642	RCD gnl-Mmul_051212-Chr7-169628813-169629993	atggacttgacctggaagatcctcctcttggtgacagcagccacaggtgaaggggtccccaggtctcatt cttgagaagcgaattgagtcagtcaggtcaaggggacgttcactcctgtcttctctccacaggtgccc ctccCAGGTGCAGCTGGTGCAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAAC

			CACATCCTGAGAGTGTGAGAAACCCAGGAAGGAGGCAGCTGTGCTG
	Z18904	gnl Mmul_0 51212 Chr7: 169628884- 169630129	atggacttgacctggaagatcctcctcttggtgacagcagccacaggttaaggggtccccaggtctcatt cttgagaagcgaattgagtcaggtcaaggggacgttcactcctgtctctccacaggtgcccc ctccCAGGTGCAGCTGGTGAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACT
	M99639	gnl Mmul_0 51212 Chr7: 169628838- 169630125	atggacttgacctggaagatcctcctcttggtgacagcagccacaggttaaggggtccccaggtctcatt cttgagaagcgaattgagtcaggtcaaggggacgttcactcctgtctctccacaggtgcccc ctccCAGGTGCAGCTGGTGAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAC CACATCCTGAGAGTGTGAGAAA
	X92210	gnl Mmul_0 51212 Chr7: 169571494- 169573015	atgtctgggcctgagatcatcaccataaccacatccctccccagagaagccctgagagcacag ctcctcattagggaccggacctggaggatcctctcttggtggcagcactacaggttaaggggttccca agtccagtgatgagaaggggtgagttcagtcagtgaggctctcattcctgcgtcctccccaca gggtccactccCAGGTGAAGCCGGTGAGTCTGGAGCTGAGATGAAGAAGCCT GGGTCTCAGTGAAGGTCTCCTGGAAGGCTTCCAGATACACCTTTACCAGCTA CATTATGCACTGGGTGCGACAGAACCCTGGACAAGGGCTTGAGTGGACGGG ATTTTAGTTATCTGAGAGATGTTTCATATAACATTTATTCTGTAAGCAAATTC AGGGATAGTCGAATGAATCAAATTAAGAAATCTGATACAGAACTTCTCTGA ATTGATCTTTGTAAACATCAATCTCTGAATCAATGTTGTAATATTTTGAAC CAACACAAAGTTCACATTTTAACTCTATTTTGTCTCTATTTAA
	X07448	gnl Mmul_0 51212 Chr7: 169628821- 169630384	atggacttgacctggaagatcctcctcttggtgacagcagccacaggttaaggggtccccaggtctcatt cttgagaagcgaattgagtcaggtcaaggggacgttcactcctgtctctccacaggtgcccc ctccCAGGTGCAGCTGGTGAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAC CACATCCTGAGAGTGTGAGAAACCCAGGAAGGAGGCAG
	X62109	gnl- Mmul_0512 12-Chr7- 169628831- 169630223	atggacttgacctggaagatcctcctcttggtgacagcagccacaggttaaggggtccccaggtctcatt cttgagaagcgaattgagtcaggtcaaggggacgttcactcctgtctctccacaggtgcccc ctccCAGGTGCAGCTGGTGAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAC CACATCCTGAGAGTGTGAGAAACCCAGG
	X92209	gnl- Mmul_0512 12-Chr7- 169628824- 169630312	atggacttgacctggaagatcctcctcttggtgacagcagccacaggttaaggggtccccaggtctcatt cttgagaagcgaattgagtcaggtcaaggggacgttcactcctgtctctccacaggtgcccc ctccCAGGTGCAGCTGGTGAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTCCAGGGCAGAGTCAC

			CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAAC CACATCCTGAGAGTGTGAGAAACCCAGGAAGGAGG
	X92343	gnl- Mmul_0512 12-Chr7- 169628835- 169630190	atggacttgacctggaagatcctccttggtagacagcagccacaggttaaggggtccccaggtctcatt cttgagaagcgaattgagtcagtcaggggacgttcactcctgtctctccacaggtgccc ctccCAGGTGCAGCTGGTGAGTCTGGGGCTGAGATTAAGCAGCCTGGGGCCT CAGTGAAGCTCTCCTGCAAGGCTTCTGGATATACCTTCACTAGCTACTACATG CACTGGGTGAGACAGGCCCTGGACAAGGCCTTGAGTGGATAGGACTGATCT CCCCTTACAATGGTAACAAAGGCTACGCACAGAACTCCAGGGCAGAGTCAC CATAACTACGGACACGTCCACGAGCACAGGCTACATGGAGCTGAGCAGCCTG AGATCTGAGGACACAGCCGTGTATTACTGTACGAGAGACACAGTGTGAAAAC CACATCCTGAGAGTGTGAGAAACCC
	Z12305	Z12305 Z12 305 H.sapie ns germline lg H-chain V- region (DP- 3)	gaggctcagctggtacagtctggggctgaggtgaagaagcctggggctacagtgaatctcctgca aggttcttgatacaccttcaccgactactacatgactgggtgcaacaggccctggaaaagggtt gagtggtgggacttgtgatcctgaagatgggtgaacaataacgcagagaagtccagggcagag tcaccataaccgcgacacgtctacagacacagcctacatggagctgagcagcctgagatctgagga cacggccgtgtattactgtgaaca
IGHV2	L21969	gnl Mmul_0 51212 Chr7: 169382524- 169383979	atggacacgcttctccacgctcctgctgctgacctccctcatgtgagtgtgtgatcagggactca ctcagggaataaacatctgttctctccttggggcttcactcttctgcttctccacagggtcttctcc CAGGTACCTTGAAGGAGTCTGATCCTGCGCTAGTGAAACCCACACAGACCCT CACGCTGACCTGCACCTTCTCTGGGTTCTCACTCAGCACTTCTGGAATGGGTG TGGACGGGATCCGTACGCCCCAGGGAAGGCCCTGGAGTGGCTTGCACGCA TTGATTGGGATGATGATAAATACTACAGCCCATCGCTGAAGAGCAGGCTCAC CATCTCCAAGGACACCTCCAAGAACCAGGTTGTCCTTACAATGACCAACATGG ACCTGTGGACACAGCCACATATTACTGTGCACGGAGACCACAGAGACACA
	M99648	gnl- Mmul_0512 12-Chr7- 169382453- 169383999	atggacacgcttctccacgctcctgctgctgacctccctcatgtgagtgtgtgatcagggactca ctcagggaataaacatctgttctctccttggggcttcactcttctgcttctccacagggtcttctcc CAGGTACCTTGAAGGAGTCTGATCCTGCGCTAGTGAAACCCACACAGACCCT CACGCTGACCTGCACCTTCTCTGGGTTCTCACTCAGCACTTCTGGAATGGGTG TGGACGGGATCCGTACGCCCCAGGGAAGGCCCTGGAGTGGCTTGCACGCA TTGATTGGGATGATGATAAATACTACAGCCCATCGCTGAAGAGCAGGCTCAC CATCTCCAAGGACACCTCCAAGAACCAGGTTGTCCTTACAATGACCAACATGG ACCTGTGGACACAGCCACATATTACTGTGCACGGAGACCACAGAGACACAG CTCAGGGCGCCTCCTGTACAAGAACCAGGCTGCTTCTTATTGGTGCTCCCTC CCCACCAATGCAGAACA
	M99647	gnl Mmul_0 51212 Chr7: 169594442- 169596070	atggacacgcttctccacgctcctgctgctgacctccctcctgtgagtgtgtggtcaagcactac ttcagggtgaaatatctgcacttttcttgggtttcatcttctatgcttctccacagggtcttgtcc CAGGTGACCTTGAAGGAGTCTGGTCCTGCGCTGGTGAAACCCACACAGACCC TCACGCTGACCTGCACCTTCTCTGGGTTCTCACTCAGCACTAGTGGTATGGGT GTGGGCTGGATCCGTACGCCCTCACGGAAGACCCTGGAGTGGCTTGCACACA TTTATTGGGATGATGATAAGCGCTACAGCACATCTCTGAAGAGCAGGCTCAC CATCTCCAAGGACACCTCCAAAACCAGGTGGTTCTAACAATGACCAACATGG ACCCTATGGACACAGCCACATATTACTGTGCACGGAGACCACAGAGACACAG TCCCGGGCGCCTCCTGTACAAGAACCAGGCTGCTTCTCAGGGGTGCTCCCTC CCCATCTCTGCAGAACAGGAAAGAGTGGCTGAGATGCCTTTTCTGCCAGGG
	X62111	gnl- Mmul_0512 12-Chr7- 169594471- 169596072	atggacacgcttctccacgctcctgctgctgacctccctcctgtgagtgtgtggtcaagcactac ttcagggtgaaatatctgcacttttcttgggtttcatcttctatgcttctccacagggtcttgtcc CAGGTGACCTTGAAGGAGTCTGGTCCTGCGCTGGTGAAACCCACACAGACCC TCACGCTGACCTGCACCTTCTCTGGGTTCTCACTCAGCACTAGTGGTATGGGT GTGGGCTGGATCCGTACGCCCTCACGGAAGACCCTGGAGTGGCTTGCACACA TTTATTGGGATGATGATAAGCGCTACAGCACATCTCTGAAGAGCAGGCTCAC CATCTCCAAGGACACCTCCAAAACCAGGTGGTTCTAACAATGACCAACATGG ACCCTATGGACACAGCCACATATTACTGTGCACGGAGACCACAGAGACACAG TCCCGGGCGCCTCCTGTACAAGAACCAGGCTGCTTCTCAGGGGTGCTCCCTC CCCATCTCTGCAGAACAGGAAAGAGTGGCTGAGATGCCTTTTCTGCCAGGG

			TTTATTGGGATGATGATAAGCGCTACAGCACATCTCTGAAGAGCAGGCTCAC CATCTCCAAGGACACCTCCAAAAACCAGGTGGTTCTAACAAATGACCAACATGG ACCCTATGGACACAGCCACATATTACTGTGCACGGAGACCACAGAGACACAG TCCCGGGCGCCTCCTGTACAAGAACCAGGCTGCTTCTCAGGGGTGCTCCCTC CCCATCTCTGCAGAACAGGAAAG
IGHV3	L06618	gnl Mmul_0 51212 Chr7: 169485479- 169486838	atgaggaaaaagtcagtgtagccagacacaaacctccttataggaacgctgccaggaaatcagctg caggggtcactcaggaccactgatcatgtacatgaatatgagtgagagaaacagtggtatgggtg gcataatccactgtggatatgggtggcagtttctgaccaggggtgtctctatgtttgcagggtgccagtgG AGGTGCAGCTGGTGGAGTCTGGGGGAGCCTCGGTACAGCCTGGGGCGTCCC TGAGACTCTGCTGTGCAGCCTCTGGATTACCTTCAGTAGCACCAGGATGAAC TGGATCCGACAGGCTCCAGGGAAGAAGCTGGAGTGGGTGACTGACATAAAG TAGGATGGAAGTGAGAAATACTATGTAGACTCTGTGAAGGGCCGATTACCA TCTCCAGAGACAATGCCAAGAAGCTCCCTCTATCTGCAAATGAACAGCCTGAGA GCTGAGGGGACACAGCCGTGTATTACTGTGTGAGAGACACAG
	M83134	gnl-Mmul- 051212- Chr7- 169513624- 169515410	atggagtttgccctgagctgggttttctgtgtctattttaaagtgattcatgaggaaatagagata ttgagtgtagtgagacatgagtgagagaaacagtggtatgttgccagtttctgaccttggtgtctttt tggttgagggtgccagggtGAGGTGCAGCTGGTGGAGTCTGGGGCGGCTTGGCA AAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCTCTGGATTACCTTCAG TAACTACTGGATGCACTGGGTTCCGACGACACAGGGAAGGGGCTGGAGTG GATCTCAGCTATTAATAGTGCTGGTAGTAGCACATACTATGCAGACTCCGTGA AGGGCAGATTACCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCA AATGGACGGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGA CACAGTGAGGGGAGGTCAGTGTGAGCCAGACACAAACCTCCCTGCAGGAA CGCCGGGGAAATCAGCGGCAGGGGGCGCTCGGGACCCACTGATCAGAGTCA ACCCAGAGGCCGGTGCACACGGAGGCTGATTTCTGTGAGGATGTGGGAC
	M99649	gnl Mmul_0 51212 Chr7: 169485525- 169486839	atgaggaaaaagtcagtgtagccagacacaaacctccttataggaacgctgccaggaaatcagctg caggggtcactcaggaccactgatcatgtacatgaatatgagtgagagaaacagtggtatgggtg gcataatccactgtggatatgggtggcagtttctgaccaggggtgtctctatgtttgcagggtgccagtgG AGGTGCAGCTGGTGGAGTCTGGGGGAGCCTCGGTACAGCCTGGGGCGTCCC TGAGACTCTGCTGTGCAGCCTCTGGATTACCTTCAGTAGCACCAGGATGAAC TGGATCCGACAGGCTCCAGGGAAGAAGCTGGAGTGGGTGACTGACATAAAG TAGGATGGAAGTGAGAAATACTATGTAGACTCTGTGAAGGGCCGATTACCA TCTCCAGAGACAATGCCAAGAAGCTCCCTCTATCTGCAAATGAACAGCCT
	M99651	gnlMmul- 051212- Chr7- 169622161- 169623566	atggagtttgctgtagctgggttttctgtgtctttttaaagtgattgagagtgagtagaata ctgtgtgtaggagacatgagtgagagaaacagtggtatgttgccagtttctgaccttggtgtctttt tggttgagggtgccagggtGAGGTGCAGCTGGTGGAGTCTGGGGCGGCTTGGCA AAGCCTGGGGGGTCCCTGAGACTCTCCTGCGCAGCCTCCGATTACCTTCAG TGACTACTACATGGACTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTG GGTCTCACGTATTAGTAATGGTGGTGGTAGCACATGGTACGCAGACTCCGTG AAGGGCAGATTACCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTC AAATGAACAGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGAAAGA CACAGTGAGGGGAAGTCAGTGTGAGCCAGACAGAAACCTCCCTGCAGGAA
	M99652	gnlMmul- 051212- Chr7- 169681086- 169682608	atggagttggggctgagctgggttttctgtgtctttttaaagtgattcagggaactagagag gttgaggatatgagtgagagaaacagtggtatgtgtggcagttttgaccaggggtgtctctttgtttgc agggtgccagtgGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTCCAGCC TGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACCTTTGGTGATT ATGGCATGCACTGGGTCCGCCAGGCTCCGGGAAAGGGACTGGAGTGGGTCT CATCCATTAGTAACACTGGTAAAACCGTATACTACGCTGACTCCGTGAAGGGC CGATTACCGTCTCCAGAGACAACGCCAAGAAGCTCGTGTCTCTGCAAATGAG CAGCCTGAGAGCCGAGGACACGGCCGTGTATTACTGTACTAGAGACACAGTG AGGCGAGGTCAGTGTGAGCCAGACACAAACCCCTGCGAGGG

	M99657	gnl-Mmul-051212-Chr7-169513727-169515232	atggagtttggcctgagctgggttttcctgttgctattttaaaagtgattcatgaggaaatagagata ttgagtgtgagtgacatgagtgagagaaacagtggatatttggcagtttctgaccttggtgtctttt tgtttgaggtgtccagggtGAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCA AAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACACCTTCAG TAACTACTGGATGCACTGGGTTCCGCCAGGCACCCAGGGAAGGGGCTGGAGTG GATCTCAGCTATTAATAGTGCTGGTAGTAGCACATACTATGCAGACTCCGTGA AGGGCAGATTCACCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCA AATGGACGGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGA CACAGTGAGGGGAGGTCACTGTGAGCCAGACACAAACCTCCCTGCAGG
	M99660	gnlMmul-051212-Chr7-169513727-169515238	atggagtttggcctgagctgggttttcctgttgctattttaaaagtgattcatgaggaaatagagata ttgagtgtgagtgacatgagtgagagaaacagtggatatttggcagtttctgaccttggtgtctttt tgtttgaggtgtccagggtGAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCA AAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACACCTTCAG TAACTACTGGATGCACTGGGTTCCGCCAGGCACCCAGGGAAGGGGCTGGAGTG GATCTCAGCTATTAATAGTGCTGGTAGTAGCACATACTATGCAGACTCCGTGA AGGGCAGATTCACCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCA AATGGACGGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGA CACAGTGAGGGGAGGTCACTGTGAGCCAGACACAAACCTCCCTGCAGG
	M99666	gnl-Mmul-051212-Chr7-169513670-169515232	atggagtttggcctgagctgggttttcctgttgctattttaaaagtgattcatgaggaaatagagata ttgagtgtgagtgacatgagtgagagaaacagtggatatttggcagtttctgaccttggtgtctttt tgtttgaggtgtccagggtGAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCA AAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACACCTTCAG TAACTACTGGATGCACTGGGTTCCGCCAGGCACCCAGGGAAGGGGCTGGAGTG GATCTCAGCTATTAATAGTGCTGGTAGTAGCACATACTATGCAGACTCCGTGA AGGGCAGATTCACCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCA AATGGACGGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGA CACAGTGAGGGGAGGTCACTGTGAGCCAGACACAAACCTCCCTGCAGGAA CGCCGGGGAAATCAGCGGCAGGGGGCGCTCGGGACCCACTGATCAGAGTCA ACCC
	M99669	gnl Mmul_051212 Chr7:169358346-169359837	atggaatttagctgagctgggttttcctgttgctgtttttaaagtgactcatggagaactagagata ttgagtgtgagtgacatgagtgagagaaacagtggatatttggcagtttctgaccttggtgtctctg tgtttgaggtgtccaggtGAGGTGCAGTTGGTGGAGTCTGGGGGAGGCTTGGTAC AGCCTGGGGGCTCCCTGAGACTCTCTGTACAGGCTCTGGATTACACCTTCAGT AGCTACTACATGTACTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGG GTCTCAGCTATTAATACTGGTGGGGGTAGCACATGGTACACAGACTCCGTGA AGGGCCGATTACCATCTCCAAAGAGAACGCCAAGAACACACTGTATCTTCAA ATGGACAGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGAAAGACA CACAGTGAGGAGAAGTCAGTGTGAGCCAGACACAAACCTCCCTG
	M99672	gnl Mmul_051212 Chr7:169513725-169515378	atggagtttggcctgagctgggttttcctgttgctattttaaaagtgattcatgaggaaatagagata ttgagtgtgagtgacatgagtgagagaaacagtggatatttggcagtttctgaccttggtgtctttt tgtttgaggtgtccagggtGAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCA AAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACACCTTCAG TAACTACTGGATGCACTGGGTTCCGCCAGGCACCCAGGGAAGGGGCTGGAGTG GATCTCAGCTATTAATAGTGCTGGTAGTAGCACATACTATGCAGACTCCGTGA AGGGCAGATTCACCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCA AATGGACGGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGA CACAGTGAGGGGAGGTCACTGTGAGCCAGACACAAACCTCCCTGCAGGAA
	M99676	gnl Mmul_051212 Chr7:169701616-169703359	atggagttggggctgagctgggttttcctgttgctctttttaaagtaattcatggagaactagagata ctgagtgtgaggggacatgagtgagagaaacagtggatatttggcagtttctgagcggtgtctc tgtgttcgaggtgtgaggtGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGT CCAGCCTGGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCCGATTACACCTTCA GTGACTACTACATGTACTGGGTCCGCCAGGCTCCAGGAAAGGGGCTGGAGT

			GGGTGGGTTTCATTAGAAGCAAAGCTTACGGTGGGACAGCAGAATACGCCG CGTCTGTGAAAGGCAGATTCAACATCTCCAGAGATGATTCCAAAAGCATCGCC TATCTGCAAATGAGCAGCCTGAAAACCGAGGACACGGCCGTGTATTACTGTA CTAGGGACACAGTGAGGCGAGGTCAAGTGTGAGCCCAGACACAAACCTCCCTG CAGGGGCGCACGGGGCCACAGGGGGCGCT
	M99679	gnl Mmul_0 51212 Chr7: 169358238- 169359884	atggaatttaggctgagctgggttttcctgtgtgtgttttaaagtgactcatggagaactagagata ttgagtgtgagtgacatgagtgaaagaaacagtggatatttggcagtttctgaccttggtgtctctg tgtttgaggtgtccagtgtGAGGTGCAGTTGGTGGAGTCTGGGGGAGGCTTGGTAC AGCCTGGGGGCTCCCTGAGACTCTCTGTACAGGCTCTGGATTACCTTCAGT AGCTACTACATGTACTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGG GTCTCAGCTATTAATACTGGTGGGGTAGCACATGGTACACAGACTCCGTGA AGGGCCGATTCAACATCTCCAAAGAGAACGCCAAGAACACACTGTATCTTCAA ATGGACAGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGAAAGACA CACAGTGAGGAGAAGTCAGTGTGAGCCCAGACACAAACCTCCCTGCAGGGA CGCGGGGGGAAATCAGCGGCAGGGGGCGCTCGGGACCCACTGATCAGAG T CAACCCAGAGGCAGGTGCAGACGGAGGCTGATTTTCTGTCAGGATGTGGG
	M99682	gnl Mmul_0 51212 Chr7: 169513671- 169515285	atggagtttgccctgagctgggttttcctgtgtgttttaaagtgattcatggagaactagagata ttgagtgtgagtgacatgagtgagagaaacagtggatatttggcagtttctgaccttggtgtctttt tgtttgaggtgtccagggtGAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCA AAGCCTGGGGGTCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACCTTCAG TAACTACTGGATGCACTGGGTTCCGCCAGGCACCCAGGGAAGGGGCTGGAGTG GATCTCAGCTATTAATAGTGCTGGTAGTACACATACTATGCAGACTCCGTGA AGGGCAGATTCAACATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCA AATGGACGGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGA CACAGTGAGGGGAGGTCAAGTGTGAGCCCAGACACAAACCTCCCTGCAGGAA CGCCGGGGAAATCAGCGGCAGGGGGCGCTCGGGACCCACTGATCAGAGTCA ACC
	M99655	gnl Mmul_0 51212 Chr7: 169358546- 169359851	atggaatttaggctgagctgggttttcctgtgtgttttaaagtgactcatggagaactagagata ttgagtgtgagtgacatgagtgaaagaaacagtggatatttggcagtttctgaccttggtgtctctg tgtttgaggtgtccagtgtGAGGTGCAGTTGGTGGAGTCTGGGGGAGGCTTGGTAC AGCCTGGGGGCTCCCTGAGACTCTCTGTACAGGCTCTGGATTACCTTCAGT AGCTACTACATGTACTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGG GTCTCAGCTATTAATACTGGTGGGGGTAGCACATGGTACACAGACTCCGTGA AGGGCCGATTCAACATCTCCAAAGAGAACG
	M99656	gnl- Mmul_0512 12-Chr7- 169513646- 169515068	atgagttagagaaacagtggatatttggcagtttctgaccttggtgtcttttgggttgaggtgtccag ggtGAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCAAAGCCTGGGGGG TCCCTGAGACTCTCCTGTGCAGCCTCTGGATTACCTTCAGTAATACTACTGGATG CACTGGGTTCGCCAGGCACCCAGGGAAGGGGCTGGAGTGGATCTCAGCTATT AATAGTGCTGGTAGTAGCACATACTATGCAGACTCCGTGAAGGGCAGATTCA CCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCAAATGGACGGCCT GAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGACACAGTGAGGGG AGGTCAAGTGTGAGCCCAGACACAAACCTCCCTGCAGGAACGCCGGGGAAATC AGCGGCAGGGGGCGCTCGGGACCCACTGATCAGAGTCAACCCAGAGGCCGG TGACACGGAGGCTG
	X92206	gnl Mmul_0 51212 Chr7: 169711580- 169713134	atggagtttgactgagctgggttttcctgtgtgccatttttaaagtgattcatggagaactagagatg tgagtgtaggggacatgagtgagagaaacagtggatatttggcagtttctgaccaggggtgtctc tgtgttgaggtgtccagtgtGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGT CCAGCCTGGCGGGTCCCTGAGACCCTCCTGTGCAGCCTCTGGATTACCTTCA GTAGTTCTGCTATGCACTGGGTCCGCCAGGCTTCTGGGAAAAGGACTGGAGTG GGTTGGACGTATTAGAAGCAAATCTAACAATTACGCGACAGAATACGCCGCA TCGGTGAAAGGCAGATTCAACATCTCCAGAGATGATTCAAAGAACACGGCGT

			ATCTGCAAATGAACAGCCTGAAAACCGAGGACACGGCCGTGTATTACTGTGC TAGAGACACAGTGAGGGGAGGTCAGTGTGAGCCCAGACACAAACCTCC
	X92216	gnl- Mmul_0512 12-Chr7- 169534974- 169536475	atggagttggggctgagctgggtttcctgttgctcttttaaaggtgattcatggagaactagagata ctgagtgtgagggacatgagttagagaaacagtgatattgttggtcttttaacacagtgctttt gtgtttgcaggtgtgcagtgGAGGTGCAGCTGGTGGAGTCTGGGGCAGGCTTGGTC CAGCCTGGCGGGTCCCTGAGACTCTCTGTGCAGCCTCTGGATTACCTTCAG TAACTCCTGGATGAGCTGGGTCCGCCAGGCTCCAGGCAAGGGGCTAGAGTG GGTTGCCCCGATTAAAAAGGAAAGCTGACGGTGAGACAGCAGATTACGCCGCA TCTGTGAAAGGCAGATTACCATCTCAAGAGACGATTCAAAGAACACACTGT ACCTGCAAATGAACAGCCTGAAAACCGAGGACACGGCCGTGTATTACTGCAC CACAGACACAGTGAGGCGAGGTCAATGTGAGCCCAGACACAAACC
	X92217	gnl- Mmul_0512 12-Chr7- 169358191- 169359838	atggaatttaggctgagctgggtttcctgttgctgttttaaaggtgactcatggagaactagagata ttgagtgtgagtgacatgagtgaagaaacagtgatattgttggtcagtttctgaccttggtgtctctg tggtttgcaggtgtccagtgGAGGTGCAGTTGGTGGAGTCTGGGGGAGGCTTGGTAC AGCCTGGGGGCTCCCTGAGACTCTCTGTACAGCCTCAAGATTACCTTCAGT AGCTACTACATGTACTGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGG GTCTCAGCTATTAATACTGGTGGGGGTAGCACATGGTACACAGACTCCGTGA AGGGCCGATTACCATCTCAAAGAGAACGCCAAGAACACACTGTATCTTCAA ATGGACAGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGAAAGACA CACAGTGAGGAGAAGTCAGTGTGAGCCCAGACACAAACCTCCCTGCAGGGA CGCGGGGGGAAATCAGCGGCAGGGGGCGCTCGGGACCCACTGATCAGAG T CAACCCAGAGGCAGGTGCAGACGGAGGCTGATTTTCTGTCAGGATGTGGG ACTTTCTTTTGCAGTTTCTCTGGTGAACGTCTCTAAGTTCGAAATT
	X92218	gnl Mmul_0 51212 Chr7: 169513734- 169515232	atggagtttgccctgagctgggtttcctgttgctattttaaaggtgattcatgaggaaatagagata ttgagtgtgagtgacatgagttagagaaacagtgatattgttggtcagtttctgaccttggtgtctttt tggtttgcaggtgtccaggtGAGGTGCAGCTGGTGGAGTCTGGGGGCGGCTTGGCA AAGCCTGGGGGGTCCCTGAGACTCTCTGTGCAGCCTCTGGATTACCTTCAG TAACTACTGGATGCACTGGGTTCCGCCAGGCACCCAGGGAAGGGGCTGGAGTG GATCTCAGCTATTAATAGTGCTGGTAGTAGCACATACTATGCAGACTCCGTGA AGGGCAGATTACCATCTCCAGAGAGAACGCCAAGAACACACTGTATCTTCA AATGGACGGCCTGAGAGCTGAGGACACGGCTGTCTATTACTGTGCGGGAGA CACAGTGAGGGGAGGTGAGTGTGAGCCCAGACACAAACCTCC
	Z14073	gnlMmul- 051212- Chr7- 169400831- 169402118	atggagtttgactgagctgggtttcctgttgctattttagaaggtgattcatggaaaactagagaga ttgagtgtgagtgcatatgacttagagaaacagtgatattgttggtcagtttctgaccaggtgtctct ttgtttgcaggtgtccagtgGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTA CAGCCTGGCCGGTCCCTGAGACCCTCTGTGCAGCCTCTGGATTACCTTCAG TAGCTATGGCATGCACTGGGTCCGCCAGGCTCCGGAAGAGGGGCTGGTGTG GGTTTCATACATTGGTAGTAGTACCATGTACTACGCAGACTCCGTGAAGGGCC GATTCACCATCTCCAGAGACAATGCCAAGAACTCGCTGTATCTGCAAATGAAC AGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGTGTGAGA
	Z27508	gnl Mmul_0 51212 Chr7: 169711624- 169712923	atggagtttgactgagctgggtttcctgttgctattttaaaggtgattcatggagaatagagatg tgagtgtaggggacatgagttagagaaacagtgatattgttggtcagtttctgaccaggtgtctc tgtgtttgcaggtgtccagtgGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGT CCAGCCTGGCGGGTCCCTGAGACCCTCTGTGCAGCCTCTGGATTACCTTCA GTAGTTCTGCTATGCACTGGGTCCGCCAGGCTTCTGGGAAAGGACTGGAGTG GGTTGGACGTATTAGAAGCAAATCTAACAATTACGCGACAGAATACGCCGCA TCGGTGAAAGGCAGATTACCATCTCCAGAGATGATTCAAAGAACACGGCGT ATCTGCAAATGAACAGCCTGAAAACCGAGGACACGGCCGTGTATTACTGTGC TAGA
IGHV4	L10088	gnl-Mmul-	atgaagcacctgtggttcttctcctcctggtggcagctccagatgtgagtgtctcggggatgcagata

		051212- Chr7- 169579884- 169581183	tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGACACAGTGA
	L10089	gnl-Mmul- 051212- Chr7- 169579892- 169581187	atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagata tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGA
	L10094	gnl-Mmul- 051212- Chr7- 169579880- 169581209	atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagata tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGACACAGTGAGGGG
	M29811	gnl-Mmul- 051212- Chr7- 169579617- 169581472	atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagata tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGACACAGTGAGGGGAGGT GAGTGTGAGCCCAGACACAACTCCCTGCAGGGAGGCGGAGGGGGCGGG CGCAGGTGCCGCTCAGGACCAGCAGGGGGCGCGGGGGCCACAGAGCAG GAGGCCGGGTCAGGAGCAGGTGCAGGGAGGGCGGGGCTTCTCATCTGCTC AGTGCTCTCCCTCCTACCAGCACCTCAGCTGTCCCAGGGCTCCTTTCTTT GATATCTGTGGTTCTGCTTCTCACATTCTTGCCAGAAAAGAAAGGAGGAA G
	M99685	gnl Mmul_0 51212 Chr7: 169579848- 169581496	atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagata tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGACACAGTGAGGGGAGGT GAGTGTGAGCCCAGACACAACTCCCT
	X05713	gnl-Mmul- 051212- Chr7- 169579853- 169581478	atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagata tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA

			TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGACACAGTGAGGGGAGGT GAGTGTGAGCCCAGACACAAACC
	X05714	gnl-Mmul- 051212- Chr7- 169579853- 169581478	Atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagat atgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctgttcacaggggtcctgtc cCAGCTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGACACAGTGAGGGGAGGT GAGTGTGAGCCCAGACACAAACC
	X92278	gnlMmul- 051212- Chr7- 169579790- 169581024	ATGAAGCACCTGTGTTTCTTCTCCTCCTGTTGGCAGCTCCCAGATGTGAGTG TCTCGGGGATGCAGATATGACGATATGGGAGGCTGTCTCTGATCCCAGGGCT CACTGTGGGTCTCTCTGTTCACAGGGGTCTGTCCAGCTGCAGCTGCAGGA GTCGGGCCAGGACTGGTGAAGCCTTCGGAGACCCTGTCCCTCACCTGCGCT GTCTCTGGTGGCTCCATCAGCAGTAACTACTGGAGCTGGATCCGCCAGCCCC AGGGAAGGGACTGGAGTGGATTGGACGTATCTCTGGTAGTGGTGGGAGCAC CGACTACAACCCCTCCCTCAAGAGTCGAGTCACCATTTCAACAGACACGTCCA AGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACCGCGCGGACACGGCCGT GTATTACTGTGCGAGAGACACAGTGAGGGGAGGTGAGTGTGAGCCCAGACA CAAACCTCCCTGCAGGGAGGCGGAGGGGGCGGGCGCAGGTGCCGCTCAGG ACCAGCAGGGGGCGCGCGG
	Z12367	gnl-Mmul- 051212- Chr7- 169579894- 169581183	atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagata tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGA
	Z14238	gnl-Mmul- 051212- Chr7- 169579864- 169581327	atgaagcacctgtggttcttctcctcctggtggcagctcccagatgtgagtgctcggggatgcagata tgacgatatgggaggctgtctctgatcccagggtcactgtgggtctctgttcacaggggtcctgtcc CAGCTGCAGCTGCAGGAGTCGGGCCCAGGACTGGTGAAGCCTTCGGAGACC CTGTCCCTCACCTGCGCTGTCTCTGGTGGCTCCATCAGCAGTAACTACTGGAG CTGGATCCGCCAGCCCCAGGGAAGGGACTGGAGTGGATTGGACGTATCTCT GGTAGTGGTGGGAGCACCGACTACAACCCCTCCCTCAAGAGTCGAGTCACCA TTTCAACAGACACGTCCAAGAACCAGTTCTCCCTGAAGCTGAGCTCTGTGACC GCCGCGGACACGGCCGTGTATTACTGTGCGAGAGACACAGTGAGGGGAGGT GAGTGTGAGCCC
IGHV5	M99686	gnl-Mmul- 051212- Chr7- 169477778- 169479473	atgggggtcaactgccatctcaccctcctactggctgtttctccaggtcagttctgccagggttgagg tcacagaaaagaatggtagaaggagccccaattcaaattttgtatctccccacaggagctgtgtc cGAGGTGCAGCTGGTGCAGTCTGGAGCAGAGGTGAAAAGGCCGGGGAGTC TCTGAAGATCTCCTGTAAGACTTCTGGATACAGCTTTACCAGCTACTGGATCA GCTGGGTGCGCCAGATGCCCGGAAAAGGCCTGGAGTGGATGGGGGCGATTG ATCCTAGTGATTCTGATACCAGATACAGCCCGTCTTCCAAGGCCAGGTCACC ATCTCAGCCGACAAGTCCATCAGCACCGCTACCTGCAGTGGAGCAGCCTGA AGGCCTCGGACACCGCCACGTATTACTGCGCGAAAGACACAGTGAGAGAAAC CAGCACTGAGAGCACCAAAAACCTCCACAACGCAGGTGCTGAATGAGCTGC CAGAGACTCCCTCTCAGGGACCTCTCT
	X992213	gnl Mmul_0	atgggggtccacgtcaccctcgccctcctcctgactgttctccaaggtcagtcctgccagggttgagg

		51212 Chr7: 169744786- 169746573	tcacagagaagaacgggcagaaagggggccctggtgcaaatttggtgtctccacacaggtgtctctgc cGAGGTGCAGCTGGTGCAGTCTGGAGCAGAGGTGAAAAGGCCCGGGGAGTC TCTGAGGATCTCCTGTAAGACTTCTGGATACAGCTTTACCAGCTACTGGATCA GCTGGGTGCGCCAGATGCCCCGGGAGAGAACTGGAGTGGATGGGGAGCATCC ATCTGGGACCTCTGATACCAGATACAACCCGTCCTTCCAAGGCCAGGTCACC ATCTCAGCCGACAGCTCCATCAGCACCGTCTACCTGCAGTGGAGCAGCTGAA GGCCTCGGACTCCGCCATGTATTATTGTGCAAGAGGGACCCTAGAACCCCC CACGGTGCAGGTGCAGAGTGAGGTGCCAGAGACACCCTCCCCAGGGACCTCT CTATCATCCAGGAGGAAACACCGGCTGTCTGTCTCTCAGGAGCCAGAACCA CAGAACCAGTGGGAG
	X92227	gnl Mmul- 051212- Chr7- 169477877- 169479179	atggggtcaactgccatctcacctcctactggctgtttctccaggtcagttctgccaggggtgagg tcacagaaagaatggtgagaaaggagccccaattcaaatttgtatctccccacaggagctgtgc cGAGGTGCAGCTGGTGCAGTCTGGAGCAGAGGTGAAAAGGCCCGGGGAGTC TCTGAAGATCTCTGTAAGACTTCTGGATACAGCTTTACCAGCTACTGGATCA GCTGGGTGCGCCAGATGCCCCGAAAAGGCCTGGAGTGGATGGGGGCGATTG ATCCTAGTGATTCTGATACCAGATACAGCCCGTCTTCCAAGGCCAGGTCACC ATCTCAGCCGACAAGTCCATCAGCACCGCTACCTGCAGTGGAGCAGCCTGA AGGCCTCGGACACCGCCACGTATTACTGCGCGA
	X92224	gnl Mmul_0 51212 Chr7: 169248630- 169250249	atgtctgtctcttctcatctctccccgtgctgggctcccatgggggtcagtgtagggcgatccgt attcacagcagcattcacagagttaggggtgttaactttgtctgttctctccaggtgtctgtgc aCAGGTGCAGCTGCAGGAGTCAAGTCCAGGACTGGTGAAGCCCTCACAGACC CTGTCACTCACCTGTCCATCTCCGGGACAGTGTCTCTAGCAACAGTGCTAC TTGGAAGTGGATCAGGCAGTCGCCATCGAGAGGCCTTGAGTGGCTGGGAAG GACATACTACAGGTCCAAGTGGTATAATGATTATGCACAATCTGTGCAAAATC GAATAAGCATCAACCCAGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAAC TCTGTGACCCCGAGGACATGGCCGTGTATTACTGTGCAAGAGACACAGTGA GGGGAAGTCAAGTGTGAGCCCAGACACAAACCTCCCTGCAGGGAGGCTCAGG ACCCAGAGGGCACCCAGCACTACCAGCCCGGGGCCAGACCAGGAGCAGG TGTGGAGTTAAGCAAAAATGGAACCTCTTGCCGTGT
	Z27509	gnl Mmul_0 51212 Chr7: 169767568- 169768861	atggacctcacctggagcatccttttcttgggtggcagcagcaaatgtaaggggtcccaggtcccagg gcggaggcagaaaccaggccagtcagtgagacttcacgcaatcctgtgtctctccacaggtgccca ctccCAGGTGCAGCTGGTGCAGTCTGGGCCTGAGGTGAAGCAGCCTGGGGCC TCAGTGAAGGTCTCTGCAAGGCTTCTGGTTACAGCTTACCACCTATGGTAT GAATTGGGTGCGACAGGCCCTGGACAGGGCCTTGAGTGGATGGGATGGAT GAACACCTACACTGGGAACCCAACGTATGCCAGGGCTTACAGAACGGTTT GTCTTCTCCATGGACACCTCTGTGACACAGTGTACCTGCAGATCAGCAGCCT AAAGGGTGAGGACACGGCCGTGTATTACTGTGCGAGA
	M99643	gnl Mmul_0 51212 Chr7: 169767589- 169769066	atggacctcacctggagcatccttttcttgggtggcagcagcaaatgtaaggggtcccaggtcccagg gcggaggcagaaaccaggccagtcagtgagacttcacgcaatcctgtgtctctccacaggtgccca ctccCAGGTGCAGCTGGTGCAGTCTGGGCCTGAGGTGAAGCAGCCTGGGGCC TCAGTGAAGGTCTCTGCAAGGCTTCTGGTTACAGCTTACCACCTATGGTAT GAATTGGGTGCGACAGGCCCTGGACAGGGCCTTGAGTGGATGGGATGGAT GAACACCTACACTGGGAACCCAACGTATGCCAGGGCTTACAGAACGGTTT GTCTTCTCCATGGACACCTCTGTGACACAGTGTACCTGCAGATCAGCAGCCT AAAGGGTGAGGACACG
	M99644	gnl Mmul_0 51212 Chr7: 169767079- 169768729	atggacctcacctggagcatccttttcttgggtggcagcagcaaatgtaaggggtcccaggtcccagg gcggaggcagaaaccaggccagtcagtgagacttcacgcaatcctgtgtctctccacaggtgccca ctccCAGGTGCAGCTGGTGCAGTCTGGGCCTGAGGTGAAGCAGCCTGGGGCC TCAGTGAAGGTCTCTGCAAGGCTTCTGGTTACAGCTTACCACCTATGGTAT GAATTGGGTGCGACAGGCCCTGGACAGGGCCTTGAGTGGATGGGATGGAT

			GAACACCTACACTGGGAACCCAACGTATGCCAGGGCTTCACAGAACGGTTT GTCTTCTCCATGGACACCTCTGTGACACAGTGTACCTGCAGATCAGCAGCCT AAAGGGTGAGGACACGGCCGTGTATTACTGTGCGAGATACACAGTGTGGAA ACCCACATCCTGAGAGTGTGAGAAGCCCTGAGGGAAGAGGTAGCTGTGCTGA GCTGAGGCAGTGGTGCAGCAGTTTTCTAAACTTCCGTATGATCTCATTTTTGCA TTGACTTCCACTTTTATATTAGCCAAGAATGTGGGATAGACAGATGCTCCTAA GAGATCCTTAACCTTGCCCATTTTGATAGGATTTCCAGAAGATGCGAAAAGCCA CTTTGTTAAACAAAGCATCCCAAAGCCATGCCCTGCTCCACACACGTAAGTGCC CATTTCTGGTCTTTAACTGACAAGCTCTCATCAGAGCACCTGGGCTAATTCCA CATCAGGTAGAGGAATGTGTTCTAAAGGAAAGCTAAAGTTGCAATAGCAAGT CCTGCTTAATAACCTTCAGCTTCATTGTTTTGTGTGATCCATCAACTAATTATG TTAGATCAAGCTTCTCCATGACAGTTTTCTAATAAATATAAGGGATGTATAG
	Z12323	gnl Mmul_0 51212 Chr7: 169767568- 169768861	atggacctcacctggagcatccttttcttggtggcagcagcaacatgtaaggggtcccaggtcccagg gcggaggcagaaaccaggccagtcagtgagacttcacgcaatcctgtgtcctctccacaggtgccca ctccCAGGTGCAGCTGGTGCAGTCTGGGCCTGAGGTGAAGCAGCCTGGGGCC TCAGTGAAGGTCTCCTGCAAGGCTTCTGGTTACAGCTTCACCACCTATGGTAT GAATTGGGTGCGACAGGCCCTGGACAGGGCCTTGAGTGGATGGGATGGAT GAACACCTACACTGGGAACCCAACGTATGCCAGGGCTTCACAGAACGGTTT GTCTTCTCCATGGACACCTCTGTGACACAGTGTACCTGCAGATCAGCAGCCT AAAGGGTGAGGACACGGCCGTGTATTACTGTGCGAGA

What is claimed is:

1. A nucleic acid molecule having a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.
2. A nucleic acid molecule having a nucleotide sequence comprising a sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10, wherein the oligonucleotide is capable of annealing to a *Macaca mulatta* immunoglobulin sequence, or a portion thereof.
3. A nucleic acid molecule comprising a nucleotide sequence which is at least 80% identical to a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.
4. A nucleic acid molecule comprising a fragment of at least 8 nucleotides of a nucleic acid molecule having a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.
5. A set of primers selected from the group consisting of two or more oligonucleotides of oligonucleotides shown in Tables 1-3 and 8-10.
6. A vector comprising a nucleic acid molecule of claims 1-4.
7. The nucleic acid molecule of claims 1-4, further comprising a label.
8. The oligonucleotides of claim 5, further comprising a label.
9. The nucleic acid molecule of claim 7, wherein the label is selected from the group consisting of a fluorescent group, digoxigenin, biotin, radioactive labels, chemiluminescent groups, enzymes, antibodies, luminescent agents, precipitating agents, and dyes.
10. The oligonucleotides of claim 8, wherein the label is selected from the group consisting of a fluorescent group, digoxigenin biotin, radioactive labels, chemiluminescent groups, enzymes) antibodies, luminescent agents, precipitating agents, and dyes.
11. An oligonucleotide for amplifying a nucleic acid sequence in a sample, wherein the sample contains a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10, or complement thereof.

12. The oligonucleotide of claim 11, wherein the sample is from a rhesus monkey.
13. The oligonucleotide of claim 12, wherein the rhesus monkey sample is selected from the group consisting of splenocytes, lymph nodes, and lymphocytes.
14. An isolated VL kappa sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby, selected from the group of sequences set forth in Tables 4-7.
15. An isolated VL lambda sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby, selected from the group of sequences set forth in Tables 4-7.
16. An isolated VH sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby, selected from the group of sequences set forth in Tables 4-7 and 12.
17. An isolated CL lambda sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby, selected from the group of sequences set forth in Tables 4-7.
18. An isolated CH lambda sequence encoded by a nucleic acid sequence and amino acid sequence encoded thereby, selected from the group of sequences set forth in Tables 4-7.
19. A library of nucleic acid molecules encoding a plurality of antigen binding polypeptides, the library made by a method comprising the steps of:
 - a.) isolating nucleic acid molecules encoding an immunoglobulin from a rhesus monkey;
 - b.) amplifying the variable regions of the immunoglobulin's heavy chain and the lambda and/or kappa light chains using PCR with a set of primers designed for the variable regions, wherein primers for the VL kappa chains are represented by any one of SEQ ID NOs: 1-18, primers for the VL lambda chains are represented by any one of SEQ ID NOs: 19-38, primers for the VH chains are represented by any one of SEQ ID NOs: 39-50, and/or their analogues or combination, wherein the amplified variable regions incorporate a flexible linker into the 3' end of the V_L amplicons and the 5' end of the V_H amplicons or vice versa;
 - c.) using the flexible linker, randomly linking the V_H and V_L amplicons.

20. The library of claim 19, further comprising scFv DNA constructs generated from randomly combined VH and VL chains with each scFv chain constituting a library member.
21. The library of claim 19, wherein the library is a library of display packages displaying the scFv chains, wherein a library member comprises a nucleic acid encoding scFv chain, and the scFv chain is displayed from the package.
22. The library of claim 21, wherein the population of nucleic acids are cloned into a DNA construct.
23. The library of claim 22, wherein the DNA construct is a plasmid, phagemid, or expression cassette.
24. The library of claim 22, wherein the DNA construct is a phage display vector or bacteriophage or yeast that express individual scFv to form phage or yeast display libraries expressing the generated scFv constructs.
25. The library of claim 19, wherein a nucleic acid sequence and amino acid sequence encoded thereby representing said VH chain is selected from the group of sequences set forth in Tables 4-7 and 12 and wherein a nucleic acid sequence and amino acid sequence encoded thereby representing said VL chain is selected from the group of sequences set forth in Tables 4-7.
26. The library of claim 19, wherein a library member further comprises a nucleic acid segment encoding a tag linked to the nucleic acid encoding the ScFv chain, wherein the tag is the same in different library members.
27. The library of claim 26, further comprising contacting library members with an agent having specific affinity for the tag and isolating a subpopulation of library members that bind to the agent.
28. A nucleic acid sequence encoding an scFv isolated from the library of claim 19.
29. The scFv of claim 28, wherein a nucleic acid sequence and amino acid sequence encoded thereby representing said VH chain is selected from the group of sequences set forth in Tables 4-7 and 12.

30. The scFv of claim 28, wherein a nucleic acid sequence and amino acid sequence encoded thereby representing said VL chain is selected from the group of sequences set forth in Tables 4-7.
31. An isolated scFv polypeptide encoded by the nucleic acid of claim 28.
32. A kit for the detection of *Macaca mulatta* immunoglobulin sequences, comprising at least one nucleic acid molecule, wherein the nucleic acid molecule comprises a nucleotide sequence which is at least 80% identical to a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.
33. A kit for the detection of *Macaca mulatta* immunoglobulin sequences, comprising a set of oligonucleotides, wherein the oligonucleotides comprise a nucleotide sequence which is at least 80% identical to a nucleotide sequence selected from the group consisting of oligonucleotides shown in Tables 1-3 and 8-10.
34. The kit of claim 32 or 33, further comprising at least one labeled oligonucleotide for detecting amplified *Macaca mulatta* immunoglobulin nucleic acid.
35. The kit of claim 32 or 33, wherein said kit further comprises a component selected from the group consisting of a plurality of nucleotides and a nucleic acid polymerase.
36. The kit of claim 35, wherein said kit further comprises at least one other component for conducting a polymerase amplification reaction.
37. The kit of claim 36, wherein said kit comprises a thermostable DNA polymerase.
38. The kit of claim 32 or 33, comprising instructions for use.
39. A method for amplifying a target nucleic acid, the method comprising: combining a target nucleic acid under conditions which allow for an amplification reaction to occur with: a) one or more nucleic acid primer sequences which are at least 80% identical to the sequences set forth as oligonucleotides shown in Tables 1-3 and 8-10; b) a nucleic acid polymerase; and c) a plurality of nucleotides, thereby resulting in an amplified target nucleic acid.

40. The method of claim 39, further comprising a step for generating an scFv, wherein the amplified target nucleic acid is combined under conditions which allow for an amplification reaction to occur with: a) one or more nucleic acid primer sequences which are at least 80% identical to the sequences set forth as SEQ ID NOs: 51-52; b) a nucleic acid polymerase; and c) a plurality of nucleotides, thereby generating an scFv.
41. The method of claim 39 or 40, wherein the target nucleic acid is from the *Macaca mulatta* genome or transcriptome.
42. The method of claim 41, wherein the target nucleic acid is a *Macaca mulatta* immunoglobulin sequence.
43. The method of claim 39 or 40, wherein said polymerase is selected from the group consisting of reverse transcriptase and thermostable DNA polymerase.
44. The method of claim 39 or 40, wherein the amplified *Macaca mulatta* immunoglobulin nucleic acid is sequenced.
45. The method of claim 39 or 40, wherein the amplified *Macaca mulatta* immunoglobulin nucleic acid is cloned into a vector.
46. The vector of claim 45.
47. A method for preparing single-chain variable fragment (scFv) nucleic acid molecules, comprising the steps of:
- a.) isolating nucleic acid molecules encoding an immunoglobulin from a rhesus monkey;
 - b.) amplifying the variable regions of the immunoglobulin's heavy chain and the lambda and/or kappa light chains using PCR with a set of primers designed for the variable regions, wherein primers for the VL kappa chains are represented by any one of SEQ ID NOs: 1-18, primers for the VL lambda chains are represented by any one of SEQ ID NOs: 19-38, primers for the VH chains are represented by any one of SEQ ID NOs: 39-50, and/or their analogues or combination, wherein the amplified variable regions incorporate a flexible linker into the 3' end of the V_L amplicons and the 5' end of the V_H amplicons or vice versa; and
 - c.) using the flexible linker, randomly linking the V_H and V_L amplicons.

48. The method of claim 47, wherein the isolated nucleic acid molecules are selected from the group consisting of genomic DNA, RNA, and cDNA.
49. The method of claim 47, wherein the nucleic acid molecules encoding an immunoglobulin from a rhesus monkey are obtained from rhesus monkey splenocytes or lymphocytes.
50. The method of claim 49, wherein the rhesus monkey is immunized with a specific antigen.
51. The method of claim 50, wherein the antigen is selected from the group consisting of a bacterial antigen, a fungal antigen, a parasitic antigen, a viral antigen, a toxin, a toxoid, and a self-antigen administered with an adjuvant such that tolerance is broken.
52. The method of claim 47, wherein the amplified variable regions incorporate enzyme restriction sites into the 5' end of the V_L amplicons and the 3' end of the V_H amplicons.
53. The method of claim 47, further comprising the step of cloning the single chain variable fragment into a DNA construct.
54. The method of claim 53, whereby the DNA construct is a plasmid, phagemid, or expression cassette.
55. A method for mutagenizing the plurality of scFv according to claim 19, the method comprising:
- a.) mutagenizing genes encoding the individual scFv members of the library; and
 - b.) expressing the genes to produce mutagenized chimeric scFv.
56. The method of claim 55 further comprising the step of screening said mutagenized chimeric scFv antibodies to select for a desired structure or function.
57. A method of identifying an antigen-specific scFv, comprising the step of panning the library of claim 19 against a target antigen, identifying an antigen-specific scFv; and isolating the identified antigen-specific scFv.
58. The method of claim 57, whereby panning is performed using phage display.

59. A method for specifically detecting *Macaca mulatta* immunoglobulin nucleic acids in a sample comprising: a) contacting said sample with one or more nucleic acid sequences which are at least 80% identical to the sequences set forth as SEQ ID NOs: 1-50, under conditions such that said *Macaca mulatta* immunoglobulin nucleic acids can hybridize with said primers; b) reverse transcribing and amplifying said nucleic acids to obtain amplified *Macaca mulatta* immunoglobulin nucleic acids; and c) detecting the presence of said amplified *Macaca mulatta* immunoglobulin nucleic acids.

60. The method of claim 59, wherein detecting the presence of said amplified *Macaca mulatta* immunoglobulin nucleic acids comprises: a) contacting said amplified *Macaca mulatta* immunoglobulin nucleic acids with a labeled oligonucleotide to obtain labeled *Macaca mulatta* immunoglobulin nucleic acids; and b) identifying said labeled nucleic acids.

61. The method of claim 59, wherein amplifying said nucleic acids is accomplished by nucleic acid sequence based amplification (NASBA), a polymerase chain reaction (PCR), Transcription Mediated Amplification (TMA) or Ligase chain reaction.

62. The method of claim 59, further comprising the step of sequencing the amplified *Macaca mulatta* immunoglobulin nucleic acids.

63. The method of claim 62, further comprising evaluating the sequence for mutations.

64. The use of an scFv according to claim 31, for the manufacture of a medicament for the treatment and/or prophylaxis of a disease involving aberrant expression or aberrant activity of an antigen recognised by said antibody.

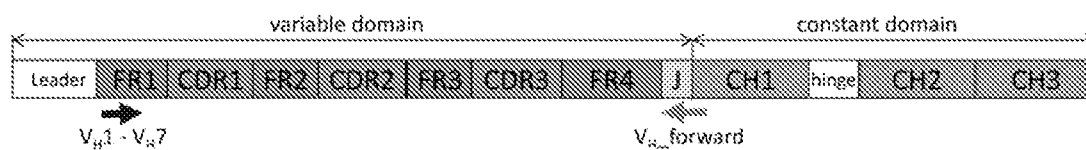
65. A method of screening for and/or diagnosis or prognosis of a disease in a subject, and/or monitoring the effectiveness of therapy for said disease, which comprises the step of detecting and/or quantifying in a biological sample obtained from said subject, the expression of an antigen recognised by an scFv according to claim 31.

1/4

Figure 1

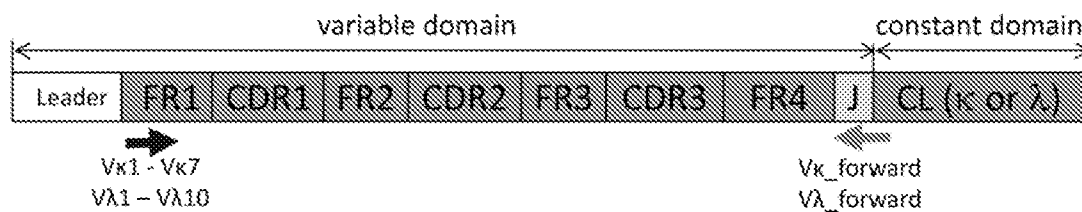
A

Heavy chain



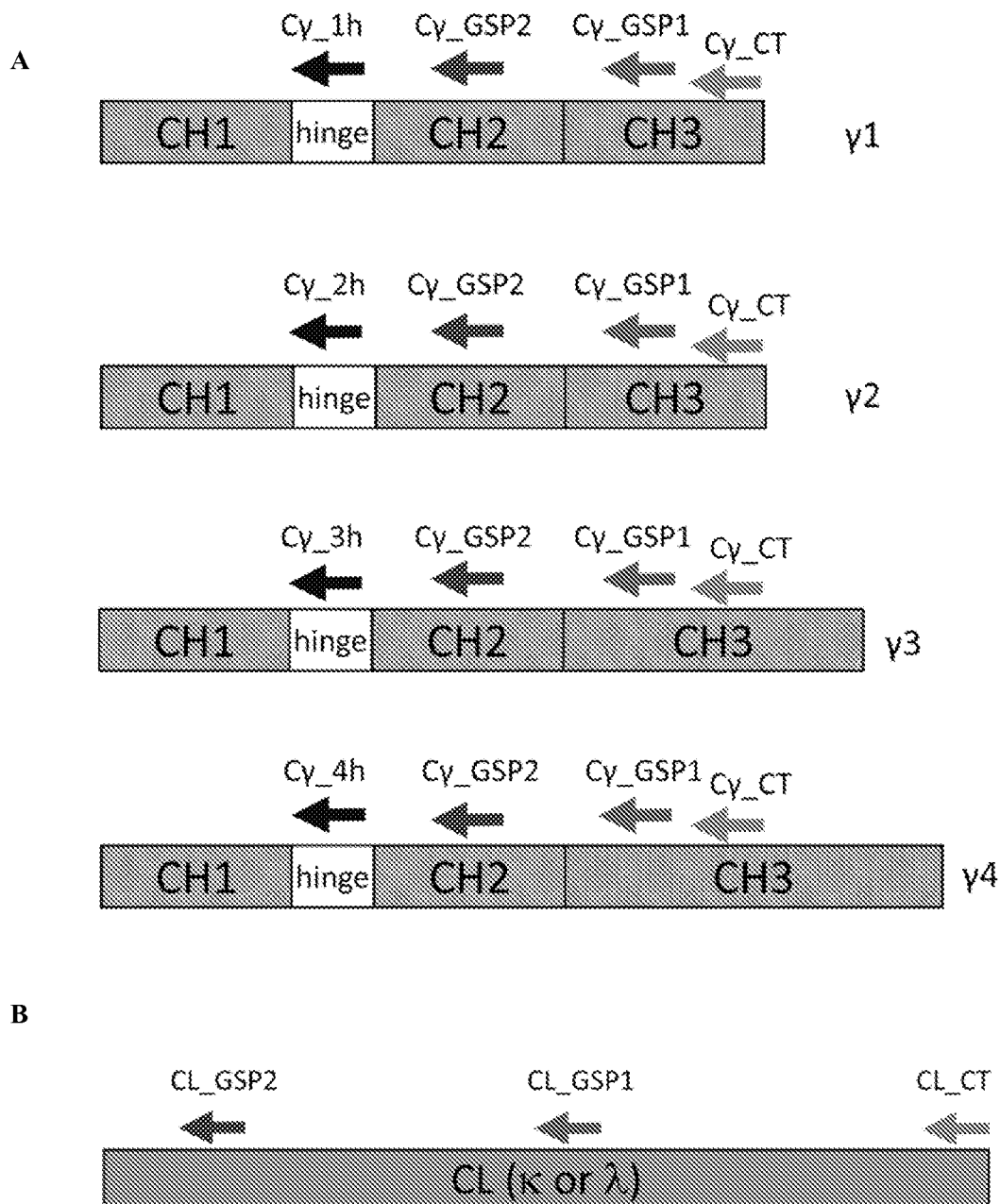
B

Light chain



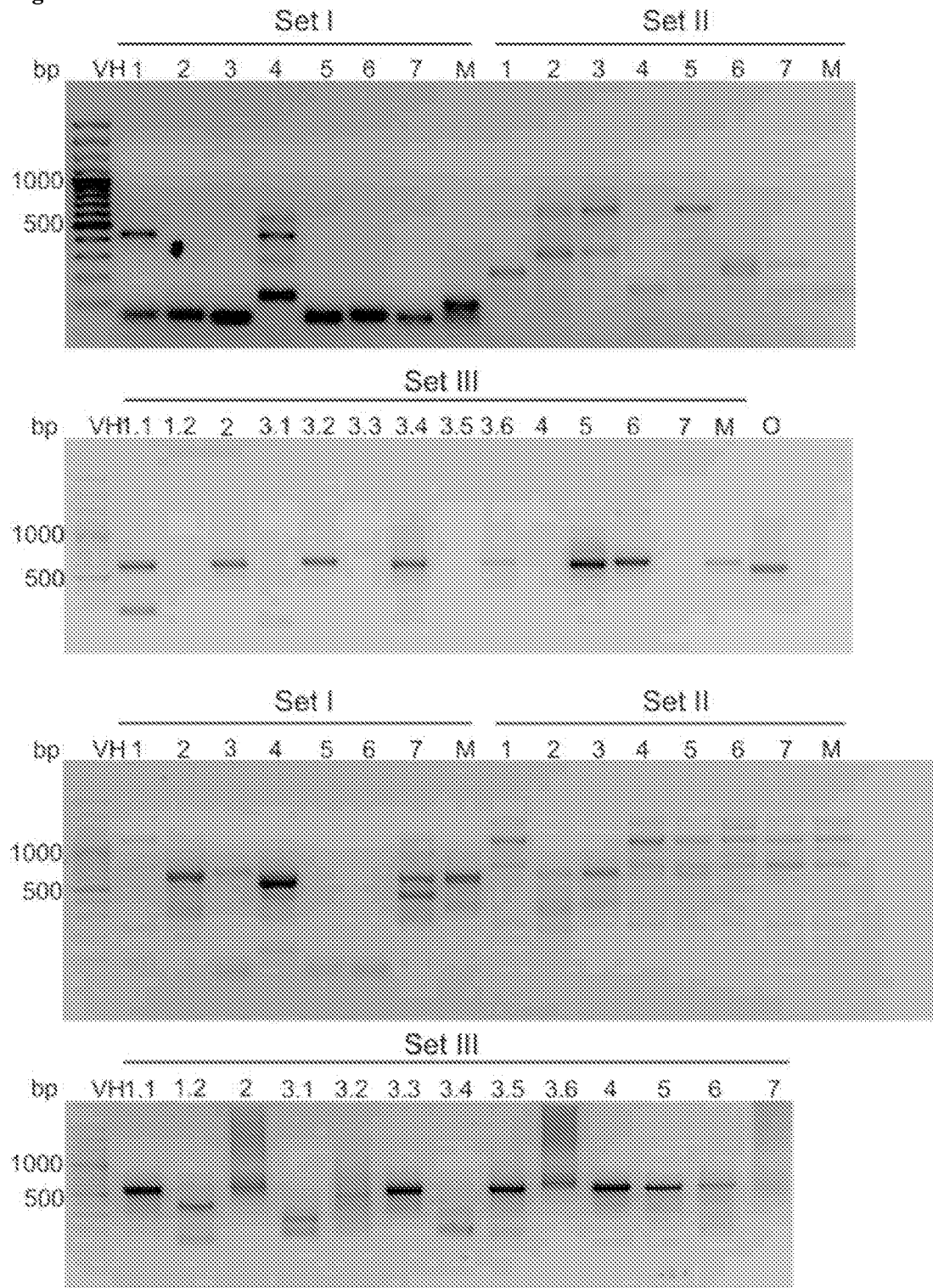
2/4

Figure 2



3/4

Figure 3



4/4

Figure 4

