



US 20060257937A1

(19) **United States**

(12) **Patent Application Publication**
Ladner

(10) **Pub. No.: US 2006/0257937 A1**

(43) **Pub. Date: Nov. 16, 2006**

(54) **FOCUSED LIBRARIES OF GENETIC
PACKAGES**

(60) Provisional application No. 60/256,380, filed on Dec.
18, 2000.

(75) Inventor: **Robert Charles Ladner**, Ijamsville,
MD (US)

Publication Classification

Correspondence Address:

FISH & RICHARDSON PC

P.O. BOX 1022

MINNEAPOLIS, MN 55440-1022 (US)

(51) **Int. Cl.**

C40B 30/06 (2006.01)

C40B 40/10 (2006.01)

G01N 33/53 (2006.01)

(52) **U.S. Cl. 435/7.1**

(73) Assignee: **Dyax Corp., a Delaware corporation**

(57) **ABSTRACT**

(21) Appl. No.: **11/416,460**

(22) Filed: **May 1, 2006**

Related U.S. Application Data

(63) Continuation of application No. 10/026,925, filed on
Dec. 18, 2001, now abandoned.

Focused libraries of vectors or genetic packages that display,
display and express, or comprise a member of a diverse
family of antibody peptides, polypeptides or proteins and
collectively display, display and express, or comprise at least
a portion of the focused diversity of the family. The libraries
have length and sequence diversities that mimic that found
in native human antibodies.

FOCUSED LIBRARIES OF GENETIC PACKAGES

[0001] This application claims the benefit under 35 USC § 120 of U.S. provisional application 60/256,380, filed Dec. 18, 2001. The provisional application and the Tables attached to it are specifically incorporated by reference herein.

[0002] The present invention relates to focused libraries of genetic packages that each display, display and express, or comprise a member of a diverse family of peptides, polypeptides or proteins and collectively display, display and express, or comprise at least a portion of the focused diversity of the family. The focused diversity of the libraries of this invention comprises both sequence diversity and length diversity. In a preferred embodiment, the focused diversity of the libraries of this invention is biased toward the natural diversity of the selected family. In a more preferred embodiment, the libraries are biased toward the natural diversity of human antibodies and are characterized by variegation in their heavy chain and light chain complementarity determining regions ("CDRs").

[0003] The present invention further relates to vectors and genetic packages (e.g., cells, spores or viruses) for displaying, or displaying and expressing a focused diverse family of peptides, polypeptides or proteins. In a preferred embodiment the genetic packages are filamentous phage or phagemids or yeast. Again, the focused diversity of the family comprises diversity in sequence and diversity in length.

[0004] The present invention further relates to methods of screening the focused libraries of the invention and to the peptides, polypeptides and proteins identified by such screening.

BACKGROUND OF THE INVENTION

[0005] It is now common practice in the art to prepare libraries of genetic packages that individually display, display and express, or comprise a member of a diverse family of peptides, polypeptides or proteins and collectively display, display and express, or comprise at least a portion of the amino acid diversity of the family. In many common libraries, the peptides, polypeptides or proteins are related to antibodies (e.g., single chain Fv (scFv), Fv, Fab, whole antibodies or minibodies (i.e., dimers that consist of V_H linked to V_L)). Often, they comprise one or more of the CDRs and framework regions of the heavy and light chains of human antibodies.

[0006] Peptide, polypeptide or protein libraries have been produced in several ways in the prior art. See e.g., Knappik et al., *J. Mol. Biol.*, 296, pp. 57-86 (2000), which is incorporated herein by references. One method is to capture the diversity of native donors, either naive or immunized. Another way is to generate libraries having synthetic diversity. A third method is a combination of the first two. Typically, the diversity produced by these methods is limited to sequence diversity, i.e., each member of the library differs from the other members of the family by having different amino acids or variegation at a given position in the peptide, polypeptide or protein chain. Naturally diverse peptides, polypeptides or proteins, however, are not limited to diversity only in their amino acid sequences. For example, human antibodies are not limited to sequence diversity in their amino acids, they are also diverse in the lengths of their amino acid chains.

[0007] For antibodies, diversity in length occurs, for example, during variable region rearrangements. See e.g., Corbett et al., *J. Mol. Biol.*, 270, pp. 587-97 (1997). The joining of V genes to J genes, for example, results in the inclusion of a recognizable D segment in CDR3 in about half of the heavy chain antibody sequences, thus creating regions encoding varying lengths of amino acids. The following also may occur during joining of antibody gene segments: (i) the end of the V gene may have zero to several bases deleted or changed; (ii) the end of the D segment may have zero to many bases removed or changed; (iii) a number of random bases may be inserted between V and D or between D and J; and (iv) the 5' end of J may be edited to remove or to change several bases. These rearrangements result in antibodies that are diverse both in amino acid sequence and in length.

[0008] Libraries that contain only amino acid sequence diversity are, thus, disadvantaged in that they do not reflect the natural diversity of the peptide, polypeptide or protein that the library is intended to mimic. Further, diversity in length may be important to the ultimate functioning of the protein, peptide or polypeptide. For example, with regard to a library comprising antibody regions, many of the peptides, polypeptides, proteins displayed, displayed and expressed, or comprised by the genetic packages of the library may not fold properly or their binding to an antigen may be disadvantaged, if diversity both in sequence and length are not represented in the library.

[0009] An additional disadvantage of prior art libraries of genetic packages that display, display and express, or comprise peptides, polypeptides and proteins is that they are not focused on those members that are based on natural occurring diversity and thus on members that are most likely to be functional. Rather, the prior art libraries, typically, attempt to include as much diversity or variegation at every amino acid residue as possible. This makes library construction time-consuming and less efficient than possible. The large number of members that are produced by trying to capture complete diversity also makes screening more cumbersome than it needs to be. This is particularly true given that many members of the library will not be functional.

SUMMARY OF THE INVENTION

[0010] One objective of this invention is focused libraries of vectors or genetic packages that encode members of a diverse family of peptides, polypeptides or proteins wherein the libraries encode populations that are diverse in both length and sequence. The diverse length comprising components that contain motifs that are likely to fold and function in the context of the parental peptide, polypeptide or protein.

[0011] Another object of this invention is focused libraries of genetic packages that display, display and express, or comprise a member of a diverse family of peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the focused diversity of the family. These libraries are diverse not only in their amino acid sequences, but also in their lengths. And, their diversity is focused so as to more closely mimic or take into account the naturally-occurring diversity of the specific family that the library represents.

[0012] Another object of this invention is diverse, but focused, populations of DNA sequences encoding peptides,

polypeptides or proteins suitable for display or display and expression using genetic packages (such as phage or phagemids) or other regimens that allow selection of specific binding components of a library.

[0013] A further object of this invention is focused libraries comprising the CDRs of human antibodies that are diverse in both their amino acid sequence and in their length (examples of such libraries include libraries of single chain Fv (scFv), Fv, Fab, whole antibodies or minibodies (i.e., dimers that consist of V_H linked to V_L)). Such regions may be from the heavy or light chains or both and may include one or more of the CDRs of those chains. More preferably, the diversity or variegation occurs in all of the heavy chain and light chain CDRs.

[0014] It is another object of this invention to provide methods of making and screening the above libraries and the peptides, polypeptides and proteins obtained in such screening.

[0015] Among the preferred embodiments of this invention are the following:

[0016] 1. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a heavy chain CDR1 selected from the group consisting of:

[0017] (1) $\langle 1 \rangle Y_2 \langle 1 \rangle_3 M_4 \langle 1 \rangle_5$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y;

[0018] (2) $(S/T)_1 (S/G/X)_2 (S/G/X)_3 Y_4 Y_5 W_6 (S/G/X)_7$, wherein (S/T) is a 1:1 mixture of S and T residues, (S/G/X) is a mixture of 0.2025 S, 0.2025 G and 0.035 of each of amino acid residues A, D, E, F, H, I, K, L, M, N, P, Q, R, T, V, W, and Y;

[0019] (3) $V_1 S_2 G_3 G_4 S_5 I_6 S_7 \langle 1 \rangle_8 \langle 1 \rangle_9 \langle 1 \rangle_{10} Y_{11} Y_{12} W_{13} \langle 1 \rangle_{14}$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; and

[0020] (4) mixtures of vectors or genetic packages characterized by any of the above DNA sequences, preferably in the ratio: HC CDR1s (1):(2):(3)::0.80:0.17:0.02.

[0021] 2. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a heavy chain CDR2 selected from the group consisting of:

[0022] (1) $\langle 2 \rangle I \langle 2 \rangle \langle 3 \rangle SGG \langle 1 \rangle T \langle 1 \rangle YADSVKGG$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; $\langle 2 \rangle$ is an equimolar mixture of each of amino acid residues Y, R, W, V, G, and S; and $\langle 3 \rangle$ is an equimolar mixture of each of amino acid residues P, S, and G or an equimolar mixture of P and S;

[0023] (2) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle G \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle YADSVKGG$, herein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; $\langle 4 \rangle$ is an equimolar mixture of residues D, I, N, S, W, Y; and $\langle 5 \rangle$ is an equimolar mixture of residues S, G, D and N;

[0024] (3) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle G \langle 5 \rangle \langle 1 \rangle YNPSLKGG$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and $\langle 4 \rangle$ and $\langle 5 \rangle$ are as defined above;

[0025] (4) $\langle 1 \rangle I \langle 8 \rangle S \langle 1 \rangle \langle 1 \rangle GGGY \langle 1 \rangle YASVKGG$, wherein $\langle 1 \rangle$ is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; $\langle 8 \rangle$ is 0.27 R and 0.027 of each of ADEFGHIKLMNPQSTVWY; and

[0026] (5) mixtures of vectors or genetic packages characterized by any of the above DNA sequences, preferably in the ratio: HC CDR2s: (1)/(2) (equimolar): (3):(4)::0.54:0.43:0.03.

[0027] 3. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a heavy chain CDR3 selected from the group consisting of:

[0028] (1) YYCA21111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

[0029] (2) YYCA211111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

[0030] (3) YYCA2111111YFDAYTG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

[0031] (4) YYCAR111S2S311YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of S and G; and 3 is an equimolar mixture of Y and W;

[0032] (5) YYCA2111CSG11CY1YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

[0033] (6) YYCA211S1TIFG1111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

[0034] (7) YYCAR111YY2S334411YFDYWG, wherein 1 is an equimolar mixture of each amino acid

residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; 2 is an equimolar mixture of D and S; and 3 is an equimolar mixture of S and G;

[0035] (8) YYCAR1111YC2231CY111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; 2 is an equimolar mixture of S and G; and 3 is an equimolar mixture of T, D and G; and

[0036] (9) mixtures of vectors or genetic packages characterized by any of the above DNA sequences, preferably the HC CDR3s (1) through (8) are in the following proportions in the mixture:

[0037] (1) 0.10

[0038] (2) 0.14

[0039] (3) 0.25

[0040] (4) 0.13

[0041] (5) 0.13

[0042] (6) 0.11

[0043] (7) 0.04 and

[0044] (8) 0.10; and more preferably the HC CDR3s (1) through (8) are in the following proportions in the mixture:

[0045] (1) 0.02

[0046] (2) 0.14

[0047] (3) 0.25

[0048] (4) 0.14

[0049] (5) 0.14

[0050] (6) 0.12

[0051] (7) 0.08 and

[0052] (8) 0.11.

[0053] Preferably, 1 in one or all of HC CDR3s (1) through (8) is 0.095 of each of G and Y and 0.048 of each of A, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, and W.

[0054] 4. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encodes a kappa light chain CDR1 selected from the group consisting of:

[0055] (1) RASQ<1>V<2><2><3>LA

[0056] (2) RASQ<1>V<2><2><2><3>LA;

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <2> is 0.2 S and 0.044 of each of ADEFGHIKLMNPQRTVWY; and <3> is 0.2Y and 0.044 each of ADEFGHIKLMNPQRTVW and Y; and

[0057] (3) mixtures of vectors or genetic packages characterized by any of the above DNA sequences, preferably in the ratio CDR1s (1):(2)::0.68:0.32.

[0058] 5. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a kappa light chain CDR2 having the sequence:

[0059] <1>AS<2>R<4><1>,

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <2> is 0.2 S and 0.044 of each of ADEFGHIKLMNPQRTVWY; and <4> is 0.2 A and 0.044 each of DEFGHIKLMNPQRSTVWY.

[0060] 6. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a kappa light chain CDR3 selected from the groups consisting of:

[0061] (1) QQ<3><1><1><1>P<1>T,

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <3> is 0.2 Y and 0.044 each of ADEFGHIKLMNPQRTVW;

[0062] (2) QQ33111P, wherein 1 and 3 are as defined in (1) above;

[0063] (3) QQ3211PP1T, wherein 1 and 3 are as defined in (1) above and 2 is 0.2 S and 0.044 each of ADEFGHIKLMNPQRTVWY; and

[0064] (4) mixtures of vectors or genetic packages characterized by any of the above DNA sequences, preferably in the ratio CDR3s (1):(2):(3)::0.65:0.1:0.25.

[0065] 7. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a lambda light chain CDR1 selected from the group consisting of:

[0066] (1) TG<1>SS<2>VG<1><3><2><3>VS,

wherein <1> is 0.27 T, 0.27 G and 0.027 each of ADEFGHIKLMNPQRSVWY; <2> is 0.27 D, 0.27 N and 0.027 each of ADEFGHIKLMNPQRSTVWY; and <3> is 0.36 Y and 0.036 each of ADEFGHIKLMNPQRSTVW;

[0067] (2) G<2><4>L<4><4><4><3><4><4>,

wherein <2> is as defined in (1) above and <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; and

[0068] (3) mixtures of vectors or genetic packages characterized by any of the above DNA sequences, preferably in the ratio CDR1s (1):(2)::0.67:0.33.

[0069] 8. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a lambda light chain CDR2 has the sequence:

[0070] $\langle 4 \rangle \langle 4 \rangle \langle 4 \rangle \langle 2 \rangle \text{RPS}$,

wherein $\langle 2 \rangle$ is 0.27 D, 0.27 N, and 0.027 each of A E F G H I K L M P Q R S T V W Y and $\langle 4 \rangle$ is an equimolar mixture of amino acid residues A D E F G H I K L M N P Q R S T V W.

[0071] 9. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a lambda light chain CDR3 selected from the group consisting of:

[0072] (1) $\langle 4 \rangle \langle 5 \rangle \langle 4 \rangle \langle 2 \rangle \langle 4 \rangle \text{S} \langle 4 \rangle \langle 4 \rangle \langle 4 \rangle \langle 4 \rangle \text{V}$,

wherein $\langle 2 \rangle$ is 0.27 D, 0.27 N, and 0.027 each of A E F G H I K L M P Q R S T V W Y; $\langle 4 \rangle$ is an equimolar mixture of amino acid residues A D E F G H I K L M N P Q R S T V W; and $\langle 5 \rangle$ is 0.36 S and 0.0355 each of A D E F G H I K L M N P Q R T V W Y;

[0073] (2) $\langle 5 \rangle \text{SY} \langle 1 \rangle \langle 5 \rangle \text{S} \langle 5 \rangle \langle 1 \rangle \langle 4 \rangle \text{V}$, wherein $\langle 1 \rangle$ is an equimolar mixture of A D E F G H I K L M N P Q R S T V W Y; and $\langle 4 \rangle$ and $\langle 5 \rangle$ are as defined in (1) above; and

[0074] (3) mixtures of vectors or genetic packages characterized by any of the above DNA sequences, preferably in the ratio CDR3s (1):(2)::1:1.

[0075] 10. A focused library comprising variegated DNA sequences that encode a heavy chain CDR selected from the group consisting of:

[0076] (1) one or more of the heavy chain CDR1s of paragraph 1 above;

[0077] (2) one or more of the heavy chain CDR2s of paragraph 2 above;

[0078] (3) one or more of the heavy chain CDR3s of paragraph 3 above; and

[0079] (4) mixtures of vectors or genetic packages characterized by (1), (2) and (3).

[0080] 11. The focused library comprising one or more of the variegated DNA sequences that encodes a heavy chain CDR of paragraphs 1, 2 and 3 and further comprising variegated DNA sequences that encodes a light chain CDR selected from the group consisting of

[0081] (1) one or more the kappa light chain CDR1s of paragraph 4;

[0082] (2) the kappa light chain CDR2 of paragraph 5;

[0083] (3) one or more of the kappa light chain CDR3s of paragraph 6;

[0084] (4) one or more of the kappa light chain CDR1s of paragraph 7;

[0085] (5) the lambda light chain CDR2 of paragraph 8;

[0086] (6) one or more of the lambda light chain CDR3s of paragraph 9; and

[0087] (7) mixtures of vectors and genetic packages characterized by one or more of (1) through (6).

[0088] 12. A population of variegated DNA sequences as described in paragraphs 1-11 above.

[0089] 13. A population of vectors comprising the variegated DNA sequences as described in paragraphs 1-11 above.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0090] Antibodies ("Ab") concentrate their diversity into those regions that are involved in determining affinity and specificity of the Ab for particular targets. These regions may be diverse in sequence or in length. Generally, they are diverse in both ways. However, within families of human antibodies the diversities, both in sequence and in length, are not truly random. Rather, some amino acid residues are preferred at certain positions of the CDRs and some CDR lengths are preferred. These preferred diversities account for the natural diversity of the antibody family.

[0091] According to this invention, and as more fully described below, libraries of vectors and genetic packages that more closely mirror the natural diversity, both in sequence and in length, of antibody families, or portions thereof are prepared and used.

Human Antibody Heavy Chain Sequence and Length Diversity

[0092] (a) Framework

[0093] The heavy chain ("HC") Germ-Line Gene (GLG) 3-23 (also known as VP-47) accounts for about 12% of all human Abs and is preferred as the framework in the preferred embodiment of the invention. It should, however, be understood that other well-known frameworks, such as 4-34, 3-30, 3-30.3 and 4-30.1, may also be used without departing from the principles of the focused diversities of this invention.

[0094] In addition, JH4 (YFDYWGQGTLVTVSS) occurs more often than JH3 in native antibodies. Hence, it is preferred for the focused libraries of this invention. However, JH3 (AFDIWGQGTMTVTVSS) could as well be used.

[0095] (b) Focused Length Diversity: CDR1, 2 and 3

[0096] (i) CDR1

[0097] For CDR1, GLGs provide CDR1s only of the lengths 5, 6, and 7. Mutations during the maturation of the V-domain gene, however, can lead to CDR1s having lengths as short as 2 and as long as 16. Nevertheless, length 5 predominates. Accordingly, in the preferred embodiment of this invention, the preferred HC CDR1 is 5 amino acids, with less preferred CDR is having lengths of 7 and 14. In the most preferred libraries of this invention, all three lengths are used in proportions similar to those found in natural antibodies.

[0098] (ii) CDR2

[0099] GLGs provide CDR2s only of the lengths 15-19, but mutations during maturation may result in CDR2s of lengths from 16 to 28 amino acids. The lengths 16 and 17 predominate in mature Ab genes. Accordingly, length 17 is the preferred length for HC CDR2 of the present invention. Less preferred HC CDR2s of this invention have lengths 16 and 19. In the most preferred focused libraries of this invention, all three lengths are included in proportions similar to those found in natural antibody families.

[0100] (iii) CDR3

[0101] HC CDR3s vary in length. About half of human HCs consist of the components: V::nz::D::ny::JHn where V is a V gene, nz is a series of bases (mean 12) that are essentially random, D is a D segment, often with heavy editing at both ends, ny is a series of bases (mean 6) that are essentially random, and JH is one of the six JH segments, often with heavy editing at the 5' end. The D segments appear to provide spacer segments that allow folding of the IgG. The greatest diversity is at the junctions of V with D and of D with JH.

[0102] In the preferred libraries of this invention both types of HC CDR3s are used. In HC CDR3s that have no identifiable D segment, the structure is V::nz::JHn where JH is usually edited at the 5' end. In HC CDR3s that have an identifiable D segment, the structure is V::nz::D::ny::JHn.

[0103] (c) Focused Sequence Diversity: CDR1, 2 and 3**[0104]** (i) CDR1

[0105] In 5 amino acid length CDR1, examination of a 3D model of a humanized Ab showed that the side groups of residues 1, 3, and 5 were directed toward the combining pocket. Consequently, in the focused libraries of this invention, each of these positions may be selected from any of the native amino acid residues, except cysteine ("C"). Cysteine can form disulfide bonds, which are an important component of the canonical Ig fold. Having free thiol groups could, thus, interfere with proper folding of the HC and could lead to problems in production or manipulation of selected Abs. Thus, in the focused libraries of this invention cysteine is excluded from positions 1, 3 and 5 of the preferred 5 amino acid CDR1s. The other 19 natural amino acids residues may be used at positions 1, 3 and 5. Preferably, each is present in equimolar ratios in the variegated libraries of this invention.

[0106] 3D modeling also suggests that the side groups of residue 2 in a 5 amino acid CDR1 are directed away from the combining-pocket. Although this position shows substantial diversity, both in GLG and mature genes, in the focused libraries of this invention this residue is preferably Tyr (Y) because it occurs in 681/820 mature antibody genes. However, any of the other native amino acid residues, except Cys (C), could also be used at this position.

[0107] For position 4, there is also some diversity in GLG and mature antibody genes. However, almost all mature genes have uncharged hydrophobic amino acid residues: A, G, L, P, F, M, W, I, V, at this position. Inspection of a 3D model also shows that the side group of residue 4 is packed into the innards of the HC. Thus, in the preferred embodiment of this invention which uses framework 3-23, residue 4 is preferably Met because it is likely to fit very well into

the framework of 3-23. With other frameworks, a similar fit consideration is used to assign residue 4.

[0108] Thus, the most preferred HC CDR1 of this invention consists of the amino acid sequence <1>Y<1>M<1> where <1> can be any one of amino acid residues: A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y (not C), preferably present at each position in an equimolar amount. This diversity is shown in the context of a framework 3-23:JH4 in Table 1. It has a diversity of 6859-fold.

[0109] The two less preferred HC CDR1s of this invention have length 7 and length 14. For length 7, a preferred variegation is (S/T)₁(S/G/<1>)₂(S/G/<1>)₃Y₄Y₅W₆(S/G/<1>); where (S/T) indicates an equimolar mixture of Ser and Thr codons; (S/G/<1>) indicates a mixture of 0.2025 S, 0.2025 G, and 0.035 for each of A, D, E, F, H, I, K, L, M, N, P, Q, R, T, V, W, Y. This design gives a predominance of Ser and Gly at positions 2, 3, and 7, as occurs in mature HC genes. For length 14, a preferred variegation is VSGGSIS<1><1><1>YYW<1>, where <1> is an equimolar mixture of the 19 native amino acid residues, except Cys (C).

[0110] The DNA that encodes these preferred HC CDR1s is preferably synthesized using trinucleotide building blocks so that each amino acid residue is present in essentially equimolar or other described amounts. The preferred codons for the <1> amino acid residues are gct, gat, gag, ttt, ggt, cat, att, aag, ctt, atg, aat, cct, cag, cgt, tct, act, gtt, tgg, and tat. Of course, other codons for the chosen amino acid residue could also be used.

[0111] The diversity oligonucleotide (ON) is preferably synthesized from BspEI to BstXI (as shown in Table 1) and can, therefore, be incorporated either by PCR synthesis using overlapping ONs or introduced by ligation of BspEI/BstXI-cut fragments. Table 2 shows the oligonucleotides that embody the specified variegations of the preferred length 5 HC CDR1s of this invention. PCR using ON-R1V1vg, ON-R1top, and ON-R1bot gives a dsDNA product of 73 base pairs, cleavage with BspEI and BstXI trims 11 and 13 bases from the ends and provides cohesive ends that can be ligated to similarly cut vector having the 3-23 domain shown in Table 1. Replacement of ON-R1V1vg with either ONR1V2vg or ONR1V3vg (see Table 2) allows synthesis of the two alternative diversity patterns—the 7 residue length and the 14 residue length HC CDR1.

[0112] The more preferred libraries of this invention comprise the 3 preferred HC CDR1 length diversities. Most preferably, the 3 lengths should be incorporated in approximately the ratios in which they are observed in antibodies selected without reference to the length of the CDRs. For example, one sample of 1095 HC genes have the three lengths present in the ratio: L=5:L=7:L=14::820:175:23::0.80:0.17:0.02. This is the preferred ratio in accordance with this invention.

[0113] (ii) CDR2

[0114] Diversity in HC CDR2 was designed with the same considerations as for HC CDR1: GLG sequences, mature sequences and 3D structure. A preferred length for CDR2 is 17, as shown in Table 1. For this preferred 17 length CDR2, the preferred variegation in accordance with the invention is: <2>I<2><3>SGG<1>T<1>YADSVKG, where <2> indicates

any amino acid residue selected from the group of Y, R, W, V, G and S (equimolar mixture), <3> is P, S and G or P and S only (equimolar mixture), and <1> is any native amino acid residue except C (equimolar mixture).

[0115] ON-R2V1vg shown in Table 3 embodies this diversity pattern. It is preferably synthesized so that fragments of dsDNA containing the BstXI and XbaI site can be generated by PCR. PCR with ON-R2V1vg, ON-R2top, and ONR2bot gives a dsDNA product of 122 base pairs. Cleavage with BstXI and XbaI removes about 10 bases from each end and produces cohesive ends that can be ligated to similarly cut vector that contains the 3-23gene shown in Table 1.

[0116] In an alternative embodiment for a 17 length HC CDR2, the following variegation may be used: <1>I<4><1><1>G<5><1><1>YADSVKGG, where <1> is as described above for the more preferred alternative of HC CDR2; <4> indicates an equimolar mixture of DINSWY, and <5> indicates an equimolar mixture of SGD N. This diversity pattern is embodied in ON-R2V2vg shown in Table 3. Preferably, the two embodiments are used in equimolar mixtures in the libraries of this invention.

[0117] Other preferred HC CDR2s have lengths 16 and 19. Length 16: <1>I<4><1><1>G<5><1><1>YNPSLKG; Length 19: <1>I<8>S<1><1><1>GGYY<1>YAASVKG, wherein <1> is an equimolar mixture of all native amino acid residues except C; <4> is a equimolar mixture of DINSWY; <5> is an equimolar mixture of SGD N; and <8> is 0.27 R and 0.027 of each of residues ADEFGHIKLMNPQSTVWY. Table 3 shows ON-R2V3vg which embodies a preferred CDR2 variegation of length 16 and ON-R2V4vg which embodies a preferred CDR2 variegation of length 19. To prepare these variegations ON-R2V3vg may be PCR amplified with ON-R2top and ON-R2bo3 and ON-R2V4vg may be PCR amplified with ON-R2top and ON-R2-bo4. See Table 3. In the most preferred embodiment of this invention, all three HC CDR2 lengths are used. Preferably, they are present in a ratio 17:16:19::579:464:31::0.54:0.43:0.03.

[0118] (iii) CDR3

[0119] The preferred libraries of this invention comprise several HC CDR3 components. Some of these will have only sequence diversity. Others will have sequence diversity with embedded D segments to extend the length, while also incorporating sequences known to allow Igs to fold. The HC CDR3 components of the preferred libraries of this invention and their diversities are depicted in Table 4: Components 1-8.

[0120] This set of components was chosen after studying the sequences of 1383 human HC sequences. The proposed components are meant to fulfill the following goals:

[0121] 1) approximately the same distribution of lengths as seen in native Ab genes;

[0122] 2) high level of sequence diversity at places having high diversity in native Ab genes; and

[0123] 3) incorporation of constant sequences often seen in native Ab genes.

[0124] Component 1 represents all the genes having lengths 0 to 8 (counting from the YYCAR motif at the end of FR3 to the WG dipeptide motif near the start of the J region, i.e., FR4). Component 2 corresponds the all the

genes having lengths 9 or 10. Component 3 corresponds to the genes having lengths 11 or 12 plus half the genes having length 13. Component 4 corresponds to those having length 14 plus half those having length 13. Component 5 corresponds to the genes having length 15 and half of those having length 16. Component 6 corresponds to genes of length 17 plus half of those with length 16. Component 7 corresponds to those with length 18. Component 8 corresponds to those having length 19 and greater. See Table 4.

[0125] For each HC CDR3 residue having the diversity <1>, equimolar ratios are preferably not used. Rather, the following ratios are used 0.095 [G and Y] and 0.048 [A, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, and W]. Thus, there is a double dose of G and Y with the other residues being in equimolar ratios. For the other diversities, e.g., KR or SG, the residues are present in equimolar mixtures.

[0126] In the preferred libraries of this invention the eight components are present in the following fractions: 1 (0.10), 2 (0.14), 3 (0.25), 4 (0.13), 5 (0.13), 6 (0.11), 7 (0.04) and 8 (0.10). See Table 4.

[0127] In the more preferred embodiment of this invention, the amounts of the eight components is adjusted because the first component is not complex enough to justify including it as 10% of the library. For example, if the final library were to have 1×10^9 members, then 1×10^8 sequences would come from component 1, but it has only 2.6×10^5 CDR3 sequences so that each one would occur in ~385 CDR1/2 contexts. Therefore, the more preferred amounts of the eight components are 1(0.02), 2(0.14), 3(0.25), 4(0.14), 5(0.14), 6(0.12), 7(0.08), 8(0.11). In accordance with the more preferred embodiment component 1 occurs in ~77 CDR1/2 contexts and the other, longer CDR3s occur more often.

[0128] Table 5 shows vgDNA that embodies each of the eight HC CDR3 components shown in Table 4. In Table 5, the oligonucleotides (ON) Ctop25, CtpmB, CBpmB, and CBot25 allow PCR amplification of each of the variegated ONs (vgDNA): C1t08, C2t10, C3t12, C4t14, C5t15, C6t17, C7t18, and C8t19. After amplification, the dsDNA can be cleaved with AflII and BstEII (or KpnI) and ligated to similarly cleaved vector that contains the remainder of the 3-23 domain. Preferably, this vector already contains diversity in one, or both, of CDR1 and CDR2 as disclosed herein. Most preferably, it contains diversity in both the CDR1 and CDR2 regions. It is, of course, to be understood that the various diversities can be incorporated into the vector in any order.

[0129] Preferably, the recipient vector originally contains a stuffer in place of CDR1, CDR2 and CDR3 so that there will be no parental sequence that would then occur in the resulting library. Table 6 shows a version of the V3-23 gene segment with each CDR replaced by a short segment that contains both stop codons and restriction sites that will allow specific cleavage of any vector that does not have the stuffer removed. The stuffer can either be short and contain a restriction enzyme site that will not occur in the finished library, allowing removal of vectors that are not cleaved by both AflII and BstEII (or KpnI) and religated. Alternatively, the stuffer could be 200-400 bases long so that uncleaved or once cleaved vector can be readily separated from doubly cleaved vector.

Human Antibody Light Chain: Sequence and Length Diversity

[0130] (i) Kappa Chain**[0131]** (a) Framework

[0132] In the preferred embodiment of this invention, the kappa light chain is built in an A27 framework with a JK1 region. These are the most common V and J regions in the native genes. Other frameworks, such as 012, L2, and A11, and other J regions, such as JK4, however, may be used without departing from the scope of this invention.

[0133] (b) CDR1

[0134] In native human kappa chains, CDR1s with lengths of 11, 12, 13, 16, and 17 were observed with length 11 being predominant and length 12 being well represented. Thus, in the preferred embodiments of this invention LC CDR1s of length 11 and 12 are used in an and mixture similar to that observed in native antibodies), length 11 being most preferred. Length 11 has the following sequence: RASQ<1>V<2><2><3>LA and Length 12 has the following sequence: RASQ<1>V<2><2><2><3>LA, wherein <1> is an equimolar mixture of all of the native amino acid residues, except C, <2> is 0.2 S and 0.044 of each of ADEFGHIKLMNPQRSTVWY, and <3> is 0.2 Y and 0.044 each of A, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W and Y. In the most preferred embodiment of this invention, both CDR1 lengths are used. Preferably, they are present in a ratio of 11:12::154:73::0.68:0.32.

[0135] (c) CDR2

[0136] In native kappa, CDR2 exhibits only length 7. This length is used in the preferred embodiments of this invention. It has the sequence <1>AS<2>R<4><1>, wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <2> is 0.2 S and 0.004 of each of ADEFGHIKLMNPQRSTVWY; and <4> is 0.2 A and 0.044 of each of DEFGHIKLMNPQRSTUWY.

[0137] (d) CDR3

[0138] In native kappa, CDR3 exhibits lengths of 1, 4, 6, 7, 8, 9, 10, 11, 12, 13, and 19. While any of these lengths and mixtures of them can be employed in this invention, we prefer lengths 8, 9 and 10, length 9 being more preferred. For the preferred Length 9, the sequence is QQ<3><1><1><1>P<1>T, wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY and <3> is 0.2 Y and 0.044 each of ADEFGHIKLMNPQRSVW. Length 8 is preferably QQ3311P and Length 10 is preferably QQ3211PPIT, wherein 1 and 3 are as defined for Length 9 and 2 is S (0.2) and 0.044 each of ADEFGHIKLMNPQRSTVWY. A mixture of all 3 lengths being most preferred (ratios as in native antibodies), i.e., 8:9:10::28:166:63::0.1:0.65:0.25.

[0139] Table 7 shows a kappa chain gene of this invention, including a PlacZ promoter, a ribosome-binding site, and signal sequence (M13 III signal). The DNA sequence encodes the GLG amino acid sequence, but does not comprise the GLG DNA sequence. Restriction sites are designed to fall within each framework region so that diversity can be cloned into the CDRs. XmaI and EspI are in FR1, SexAI is in FR2, RsrII is in FR3, and KpnI (or Acc65I) are in FR4.

Additional sites are provided in the constant kappa chain to facilitate construction of the gene.

[0140] Table 7 also shows a suitable scheme of variegation for kappa. In CDR1, the most preferred length 11 is depicted. However, most preferably both lengths 11 and 12 are used. Length 12 in CDR1 can be construed by introducing codon 51 as <2> (i.e. a Ser-biased mixture). CDR2 of kappa is always 7 codons. Table 7 shows a preferred variegation scheme for CDR2. Table 7 shows a variegation scheme for the most preferred CDR3 (length 9). Similar variegations can be used for CDRs of length 8 and 10. In the preferred embodiment of this invention, those three lengths (8, 9 and 10) are included in the libraries of this invention in the native ratios, as described above.

[0141] Table 9 shows series of diversity oligonucleotides and primers that may be used to construct the kappa chain diversities depicted in Table 7.

[0142] (ii) Lambda Chain**[0143]** (a) Framework

[0144] The lambda chain is preferably built in a 2a2 framework with an L2J region. These are the most common V and J regions in the native genes. Other frameworks, such as 31, 4b, 1a and 6a, and other J regions, such as L1J, L3J and L7J, however, may be used without departing from the scope of this invention.

[0145] (b) CDR1

[0146] In native human lambda chains, CDR1s with length 14 predominate, lengths 11, 12 and 13 also occur. While any of these can be used in this invention, lengths 11 and 14 are preferred. For length 11 the sequence is: TG<2><4>L<4><4><4><3><4><4> and for Length 14 the sequence is: TG<1>5S<2>VG<1><3><2><3>VS, wherein <1> is 0.27 T, 0.27 G and 0.027 each of ADEFGHIKLMNPQRSVWY; <2> is 0.27 D, 0.27 N and 0.027 each of AEF GHIKLMNPQRSTVWY; <3> is 0.36 Y and 0.0355 each of ADEFGHIKLMNPQRSTVW; and <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY. Most preferably, mixtures (similar to those occurring in native antibodies) preferably, the ratio is 11:14::23:46::0.33: 0.67 of the three lengths are used.

[0147] (c) CDR2

[0148] In native human lambda chains, CDR2s with length 7 are by far the most common. This length is preferred in this invention. The sequence of this Length 7 CDR2 is <4><4><4><2>RPS, wherein <2> is 0.27 D, 0.27 N, and 0.027 each of AEF GHIKLMNPQRSTVWY and <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVW.

[0149] (d) CDR3

[0150] In native human lambda chains, CDR3s of length 10 and 11 predominate, while length 9 is also common. Any of these three lengths can be used in the invention. Length 11 is preferred and mixtures of 10 and 11 more preferred. The sequence of Length 11 is <4><5><4><2><4>S<4><4><4><4>V, where <2> and <4> are as defined for the lambda CDR1 and <5> is 0.36 S and 0.0355 each of ADEFGHIKLMNPQRSTVWY. The sequence of Length 10 is <5>SY<1><5>S<5><1><4>V, wherein <1> is an equimolar mixture of ADEFGHIKLMNPQRSTVWY;

and <4> and <5> are as defined for Length 11. The preferred mixtures of this invention comprise an equimolar mixture of Length 10 and Length 11. Table 8 shows a preferred focused lambda light chain diversity in accordance with this invention.

[0151] Table 9 shows a series of diversity oligonucleotides and primers that may be used to construct the lambda chain diversities depicted in Table 7.

Method of Construction of the Genetic Package

[0152] The diversities of heavy chain and the kappa and lambda light chains are best constructed in separate vectors. First a synthetic gene is designed to embody each of the synthetic variable domains. The light chains are bounded by restriction sites for ApaLI (positioned at the very end of the signal sequence) and AscI (positioned after the stop codon). The heavy chain is bounded by SfiI (positioned within the Pe1B signal sequence) and NotI (positioned in the linker between CH1 and the anchor protein). Signal sequences other than Pe1B may also need, e.g., a M13 pIII signal sequence.

[0153] The initial genes are made with "stuffer" sequences in place of the desired CDRs. A "Stuffer" is a sequence that is to be cut away and replaced by diverse DNA but which does not allow expression of a functional antibody gene. For example, the stuffer may contain several stop codons and restriction sites that will not occur in the correct finished library vector. For example, in Table 10, the stuffer for CDR1 of kappa A27 contains a StuI site. The vgDNA for CDR1 is introduced as a cassette from EspI, XmaI, or AflII to either SexAI or KasI. After the ligation, the DNA is cleaved with StuI; there should be no StuI sites in the desired vectors.

[0154] The sequences of the heavy chain gene with stuffers is depicted in Table 6. The sequences of the kappa

light chain gene with stuffers is depicted in Table 10. The sequence of the lambda light chain gene with stuffers is depicted in Table 11.

[0155] In another embodiment of the present intention the diversities of heavy chain and the kappa or lambda light chains are constructed in a single vector or genetic packages (e.g., for display or display and expression) having appropriate restriction sites that allow cloning of these chains. The processes to construct such vectors are well known and widely used in the art. Preferably, a heavy chain and Kappa light chain library and a heavy chain and lambda light chain library would be prepared separately. The two libraries, most preferably, will then be mixed in equimolar amounts to attain maximum diversity.

[0156] Most preferably, the display is had on the surface of a derivative of M13 phage. The most preferred vector contains all the genes of M13, an antibiotic resistance gene, and the display cassette. The preferred vector is provided with restriction sites that allow introduction and excision of members of the diverse family of genes, as cassettes. The preferred vector is stable against rearrangement under the growth conditions used to amplify phage.

[0157] In another embodiment of this invention, the diversity captured by the methods of the present invention may be displayed and/or expressed in a phagemid vector (e.g., pCES1) that displays and/or expresses the peptide, polypeptide or protein. Such vectors may also be used to store the diversity for subsequent display and/or expression using other vectors or phage.

[0158] In another embodiment of this invention, the diversity captured by the methods of the present invention may be displayed and/or expressed in a yeast vector.

TABLE 1

3-23:JH4 CDR1/2 diversity = 1.78×10^8	
20 21 22 A M A ctgtctgaac cc atg gcc Scab..... NcoI....	FR1 (VP47/V3-23)..... 23 24 25 26 27 28 29 30 E V Q L L E S G gaa ggt caa ttg tta gag tct ggt MfeI
-----FR1-----	
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 G G L V Q P G G S L R L S C A ggc ggt cct ggt cag cct ggt ggt tct tta cgt ctt tct tgc gtc	
Sites of variegation <1> <1> <1> <1> 6859-fold diversity ----FR1-----> ...CDR1..... ----FR2----- 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 A S G F T F S - Y - M - W V R gct tcc gga ttc act ttc tct - tac - atg - tgg ggt cgc BspEI BsiWI BstXI.	
Sites of variegation-><2> <2> <3> -----FR2-----> ...CDR2..... 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 Q A P G K G L E W V S - I - - caa gct cct ggt aaa ggt ttg gag tgg ggt tct - atc - - ...BstXI	

TABLE 1-continued

3-23:JH4 CDR1/2 diversity = 1.78×10^8																			
<1> <1> 25992-fold diversity in CDR2 CDR2..... ---FR3--- 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 S G G - T - Y A D S V K G R F tct ggt ggc - act - tat gct gac tcc ggt aaa ggt cgc ttc -----FR3----- 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 T I S R D N S K N T L Y L Q M act atc tct aga gac aac tct aag aat act ctc tac ttg cag atg XbaI ---FR3-----> 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 N S L R A E D T A V Y Y C A K aac agc tta agg gct gag tac acc gct gtc tac tac tgc gcc aaa AflII CDR3..... Replaced by the various components! 121 122 123 124 125 126 127 D Y E G T G Y gac tat gaa ggt act ggt tat -----FRV----(JH4)----- Y F D Y W G Q G T L V T V S S tat ttc gat tat tgg ggt caa ggt acc ctg gtc acc gtc tct agt ... KpnI BstEII																			

<1> = Codons for ADEFGHIKLMNPQRSTVWY (equimolar mixture)

<2> = Codons for YRWVGS (equimolar mixture)

<3> = Codons for PS or PS and G (equimolar mixture)

[0159]

TABLE 2

OligoNucleotides used to variegate CDR1 of human HC	
CDR1 - 5 residues	
(ON-R1V1vg):	5'-ct tcc gga ttc act ttc tct <1> tac <1> atg <1> tgg gtt cgc caa gct cct gg -3'
(ON-R1top):	5'-cctactgtct tcc gga ttc act ttc tct-3'
(ON-R1bot) [RC]:	5'-tgg gtt cgc caa gct cct gggtgctcactc-3'
CDR1 - 7 residues	
(ON-R1V2vg):	5'-ct tcc gga ttc act ttc tct <6> <7> <7> tac tac tgg <7> tgg gtt cgc caa gct cct gg-3'
CDR1 - 14 residues	
(ON-R1V3vg):	5'-ct tcc gga ttc act ttc tct atc agc ggt ggt tct atc tcc <1> <1> <1> -tac tac tgg <1> tgg gtt cgc caa gct cct gg-3'

<1> = Codons of ADEFGHIKLMNPQRSTVWY 1:1

<6> = Codons for ST, 1:1

<7> = 0.2025 (Codons for SG) + 0.035 (Codons for ADEFGHIKLMNPQRSTVWY)

<1> = Codons for ADEFGHIKLMNPQRSTVWY 1:1

[0160]

TABLE 3

Oligonucleotides used to variegate CDR2 of human HC	
<u>CDR2 - 17 residues</u>	
(ON-R2V1vg):	5'-ggg ttg gag tgg gtt tct <2> atc <2> <3> tct ggg ggc <1> act <1> tat gct - gac tcc gtt aaa gg-3'
(ON-R2top):	5'-ct tgg gtt cgc caa gct cct ggg aaa ggg ttg gag tgg gtt tct-3'
(ON-R2bot) [RC]:	5'-tat gct gac tcc gtt aaa ggg cgc ttc act atc tct aga ttcctgtcac-3'
(ON-R2V2vg):	5'-ggg ttg gag tgg gtt tct <1> atc <4> <1> <1> ggg <5> <1> <1> <1> tat gct - gac tcc gtt aaa gg-3'
<u>CDR2 - 16 residues</u>	
(ON-R2V3vg):	5'-ggg ttg gag tgg gtt tct <1> atc <4> <1> <1> ggg <5> <1> <1> tat aac cct tcc ctt aag gg-3'
(ON-R2bo3) [RC]:	5'-tat aac cct tcc ctt aag ggg cgc ttc act tct aga ttcctgtcac-3'
<u>CDR2 - 19 residues</u>	
(ON-R2V4vg):	5'-fft ttg gag tgg gtt tct <1> atc <8> agt <1> <1> <1> ggg ggg act act <1> tat gcc gct tcc gtt aag gg-3'
(ON-R2bo4) [RC]:	5'-tat gcc gct tcc gtt aag ggg cgc ttc act atc tct aga ttcctgtcac-3'
<1> = Codons for A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y (equimolar mixture) <2> = Codons for Y, R, W, V, G and S (equimolar mixture) <3> = Codons for P and S (equimolar mixture) or P, S and G (equimolar mixture) <4> = Codons for DINSWY (equimolar mixture) <5> = Codons for SGDN, (equimolar mixture) <1>, <2>, <3>, <4> and <5> are as defined above <8> is 0.27 R and 0.027 each of ADEFGHIKLMNPQSTVWY	

[0161]

TABLE 4

Preferred Components of HC CDR3					
Component			Preferred Fraction of Adjusted Library		
			Length	Complexity	Fraction
1	YYCA21111YFDYWG.	8	2.6×10^5	.10	.02
(1 = any amino acid residue, except C; 2 = K and R)					
2	YYCA211111YFDYWG.	10	9.4×10^7	.14	.14
(1 = any amino acid residue, except C; 2 = K and R)					
3	YYCA2111111YFDYTG.	12	3.4×10^{10}	.25	.25
(1 = any amino acid residue, except C; 2 = K and R)					
4	YYCAR111S2S3111YFDYWG.	14	1.9×10^8	.13	.14
(1 = any amino acid residue, except C; 2 = S and G 3 = Y and W)					
5	YYCA2111CSG11CY1YFDYWG.	15	9.4×10^7	.13	.14
(1 = any amino acid residue, except C; 2 = K and R)					
6	YYCA211S1TIFG11111YFDYWG.	17	1.7×10^{10}	.11	.12
(1 = any amino acid residue, except C; 2 = K and R)					

TABLE 4-continued

<u>Preferred Components of HC CDR3</u>					
Component		Length	Complexity	Fraction of Library	Preferred Adjusted Fraction
7	YYCAR111YY2S33YY111YFDYWG.	18	3.8×10^8	.04	.08
(1 = any amino acid residue, except C; 2 = D or G; 3 = S and G)					
8	YYCAR1111YC2231CY111YFDYWG.	19	2.0×10^{11}	.10	.11
(1 = any amino acid residue, except C; 2 = S and G; 3 = T, D and G)					

[0162]

TABLE 5

<u>Oligonucleotides used to variegate the eight components of HC CDR3</u>	
(Ctop25):	5'-gctctggtcaac tta agg gct gag g-3'
(CtprmA):	5'-gctctggtcaac tta agg gct gag gac acc gct gtc tac tac tgc gcc-3' AflII...
(CBprmB) [RC]:	5'- tac ttc gat tac tgg ggc caa ggg acc ctg gtc acc tcgctccacc-3' BstEII...
(CBot25) [RC]:	5'- ggg acc ctg gtc acc tcgctccacc-3'
The 20 bases at 3' end of CtprmA are identical to the most 5' 20 bases of each of the vgDNA molecules.	
Ctop25 is identical to the most 5' 25 bases of CtprmA.	
The 23 most 3' bases of CBprmB are the reverse complement of the most 3' 23 bases of each of the vgDNA molecules.	
CBot25 is identical to the 25 bases at the 5' end of CBprmB.	
<u>Component 1</u>	
(C1t08):	5'-cc gct gtc tac tac tgc ggc <2> <1> <1> <1> <1> tac ttc gat tac tgg ggc caa gg-3'
<u>Component 2</u>	
(C2t10):	5'-cc gct gtc tac tac tgc gcc <2> <1> <1> <1> <1> <1> <1> tac ttc gat tac tgg ggc caa gg-3'
<u>Component 3</u>	
(C3t12):	5'-cc gct gtc tac tac tgc gcc <2> <1> <1> <1> <1> <1> <1> <1> tac ttc gat tac tgg ggc caa gg-3'
<u>Component 4</u>	
(C4t140):	5'-cc gct gtc tac tac tgc gcc cgt <1> <1> <1> tct <2> tct <3> <1> <1> <1> tac ttc gat tac tgg ggc caa gg-3'
<u>Component 5</u>	
(C5t15):	5'-cc gct gtc tac tac tgc gcc <2> <1> <1> <1> tgc tct ggg <1> <1> tgc tat <1> tac ttc gat tac tgg ggc caa gg-3'
<u>Component 6</u>	
(C6t17):	5'-cc gct gtc tac tac tgc gcc <2> <1> <1> tct <1> act atc ttc ggg <1> <1> <1> <1> <1> tac ttc gat tac tgg ggc caa gg-3'
<u>Component 7</u>	
(C7t18):	5'-cc gct gtc tac tac tgc ggc cgt <1> <1> <1> tat tac <2> tct <3> <3> tac tat <1> <1> <1> <1> tac ttc gat tac tgg ggc caa gg-3'

TABLE 5-continued

Oligonucleotides used to variegate the eight components of HC CDR3	
<u>Component 8</u>	
(c8t19):	5'-cc gct gtc tac tac tgc gcc cgt <1> <1> <1> <1> tat tgc <2> <2> <3> <1> tgc tat - <1> <1> <1> tac ttc gat tac tgg ggc caa gg-3'
<1> = 0.095 Y + 0.095 G + 0.048 each of the residues ADEFHIKLMNPQRSTVW, no C; <2> = K and R (equimolar mixture) <1> = 0.095 Y + 0.095 G + 0.048 each of ADEFHIKLMNPQRSTVW, no C; <2> = K and R (equimolar mixture) <1> = 0.095 Y + 0.095 G + 0.048 each of ADEFHIKLMNPQRSTVW, no C; <2> = K and R (equimolar mixture) <1> = 0.095 Y + 0.095 G + 0.048 each of ADEFHIKLMNPQRSTVW, no C; <2> = S and G (equimolar mixture); <3> = Y and W (equimolar mixture) <1> = 0.095 Y + 0.095 G + 0.048 each of ADEFHIKLMNPQRSTVW, no C; <2> = K and R (equimolar mixture) <1> = 0.095 Y + 0.095 G + 0.048 each of ADEFHIKLMNPQRSTVW, no C; <2> = K and R (equimolar mixture) <1> = 0.095 Y + 0.095 G + 0.048 each of ADEFHIKLMNPQRSTVW, no C; <2> = D and G (equimolar mixture); <3> = S and G (equimolar mixture) <1> = 0.095 Y + 0.095 G + 0.048 each of ADEFHIKLMNPQRSTVW, no C; <2> = S and G (equimolar mixture); <3> = TDG (equimolar mixture);	

[0163]

TABLE 6

3-23::JH4 Stuffers in place of CDRs	
20 21 22 A M A ctgtctgaac cc atg gcc Scab..... NcoI....	
FR1 (DP47/V3-23) 23 24 25 26 27 28 29 30 E V Q L L E S G gaa gtt caa ttg tta gag tct ggg MfeI	
-----FR1----- 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 G G L V Q P G G S L R L S C A ggc ggg ctt ggt cag cct ggg tct tta cgt ctt tct tgc gct	
----FR1-----> ...CDR1 stuffer.... ---FR2----- 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 A S G F T F S S Y A W V R gct tcc gaa ttc act ttc tct tcg tac gct tag taa tgg gtt cgc BspEI BsiWI BstXI.	
-----FR2-----> ...CDR2 stuffer. 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 Q A P G K G L E W V S p r caa gct cct ggg aaa ggg ttg gag tgg gtt tct taa cct agg tag ... BstXI AvrII..	
...CDR2 stuffer..... ---FR3--- 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 T I S R D N S K N T L Y L Q M act atc tct aga gac aac tct aag aat act ctc tac ttg cag atg XbaI	
----FR3-----> CDR3 Stuffer-----> 106 107 108 109 110 N S L R A aac agc tta agg gct tag taa agg cct taa AflII StuI...	

3-23::JH4 Stuffers in place of CDRs

	----	FR4	----	(JH4)	-----											
	Y	F	D	Y	W	G	Q	G	T	L	V	T	V	S	S	
	tat	ttc	gat	tat	tgg	ggg	caa	ggt	acc	ctg	gtc	acc	gtc	tct	agt	
								KpnI				BstEII				

[0164]

TABLE 7

```

A27:JH1 Human Kappa light chain gene
-----
gaggacc attgggcccc ctccgagact ctcgagcgca
Scab..... Eco0109I      XhoI..
      ApaI.

acgcgaattaa tgtgagtttag ctcactcatt aggcacccca ggctttacac tttatgcttc
..-35..          Plac                      ...-10.

cggctcgatat gttgtgtgga attgtgagcg gataacaatt tcacacagga

aacagctatg accatgatta

cgccaagcctt tggagccttt tttttggaga ttttcaac
      PflMI.....
      Hind III

M13 III signal sequence (AA se|)----->
 1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
 M  K  K  L  L  F  A  I  P  L  V  V  P  F  Y
gtg aag aag ctc cta ttt gct atc ccg ctt gtc gtt ccg ttt tac

-- Signal-->FR1----->
 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
  S  H  S  A  Q  S  V  L  T  Q  S  P  G  T  L
|agc|cat|agt|gca|caa|tcc|gtc|ctt|act|caa|tct|cct|ggc|act|ctt|
      ApaLI...

----- FR1 ----->| CDR1----->
 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
  S  L  S  P  G  E  R  A  T  L  S  C  R  A  S
|tcg|cta|agc|ccg|ggg|gaa|cgt|gct|acc|tta|agt|tgc|cgt|gct|tcc|
      EspI.....          AflII...
      XmaI....

(CDR1 installed as AflII-(SexAI or KasI) cassette.) For the most
preferred 11 length codon 51 (XXX) is omitted; for the preferred
12 length this codon is <2>

----- CDR1 ----->|--- FR2 ----->
      <1>          <2> <2> xxx <3>
 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
  Q  -  V  -  -  -  -  L  A  W  Y  Q  Q  K  P
|cag| - |gtt| - | - | - | - |ctt|gct|tgg|tat|caa|cag|aaa|cct|
      SexAI...

CDR2 installed as (SexAI or KasI) to (BamHI or RsrII)
cassette.)

----- FR2 ----->|----- CDR2 ----->
                        <1>          <2>          <4>
 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
  G  Q  A  P  R  L  L  I  Y  -  A  S  -  R  -
|ggg|cag|gcg|ccg|cgt|tta|ctt|att|tat| - |gct|tct| - |cgc| - |
SexAI.... KasI....

CDR2-->|--- FR3 ----->

```

A27:JH1 Human Kappa light chain gene

----- FR3 ----->
 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
 F T L T I S R L E P E D F A V
 |ttt|acc|ctt|act|att|tct|aga|ttg|gaa|cct|gaa|gac|ttc|gct|ggt|
 XbaI...

For CDR3 (Length 10): QQ3211PP1T
1 and 3 as defined for Length 9
2 S (0.2) and 0.044 each of ADEFGHIKLMNPORVTWY

CDR3 installed as XbaI to (StyI or BsiWI) cassette.

```

----->|-----CDR3----->|-----FR4----->
              <3> <1> <1> <1> <1>
106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
  Y   Y   C   Q   Q   -   -   -   -   P   -   T   F   G   Q
|tat|tat|tgc|caa|cag| - | - | - | - |cct| - |act|ttc|ggg|caa|
      BstXI.....

```

```

-----FR4----->|
121 122 123 124 125 126 127      128 129 130 131 132 133 134
  G   T   K   V   E   I   K      R   T   V   A   A   P   S
|gt|acc|aag|ggt|gaa|atc|aag|    |cgt|acg|ggt|gcc|gct|ccc|agt|
  StyI....                      BsiWI..

```

135 136 137 138 139 140 141 142 143 144 145 146 147 148 149
V F I F P P S D E Q L K S G T
|gtg|ttt|atc|ttt|cct|cct|tct|gac|gaa|caa|ttg|aag|tca|ggt|act|
MfeI...

150 151 152 153 154 155 156 157 158 159 160 161 162 163 164
A S V V C L L N N F Y P R E A
|gct|tct|gtc|gta|tgt|ttg|ctc|aac|aat|ttc|tac|cct|cgt|gaa|gct|
BssSI...

165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
K V Q W K V D N A L Q S G N S
|aaa|gtt|cag|tgg|aaa|gtc|gat|aac|gcg|ttg|cag|tcg|ggt|aac|agt|
MluI....

180 181 182 183 184 185 186 187 188 189 190 191 192 193 194
Q E S V T E Q D S K D S T Y S
|caa|gaa|tcc|gtc|act|gaa|cag|gat|agt|aaq|gac|tct|acc|tac|tct

195 196 197 198 199 200 201 202 203 204 205 206 207 208 209
L S S T L T L S K A D Y E K H
|ttg|tcc|tct|act|ctt|act|tta|tca|aag|gct|gat|tat|gag|aag|cat

210 211 212 213 214 215 216 217 218 219 220 221 222 223 224
K V Y A C E V T H Q G L S S P
|aag|gtc|tat|Gct|TGC|gaa|gtt|acc|cac|cag|ggg|ctg|agc|tcc|cct|
SacI....

TABLE 7-continued

A27:JH1 Human Kappa light chain gene														
225	226	227	228	229	230	231	232	233	234					
V	T	K	S	F	N	R	G	E	C	.	.			
gtt acc aaa agt ttc aac cgt ggt gaa tgc taa tag										ggcgcgcc				
DsaI....										AscI....				
										BssHII				
acgcattctctaa gcgccgc aacaggaggag														
NotI....														
For CDR1:														
<1> ADEFGHIKLMNPQRSTVWY 1:1														
<2> S (0.2) ADEFGHIKLMNPQRTVWY (0.044 each)														
<3> Y (0.2) ADEFGHIKLMNPQRSTVW (0.044 each)														
For CDR2:														
<1> ADEFGHIKLMNPQRSTVWY 1:1														
<2> S (0.2) ADEFGHIKLMNPQRTVWY (0.044 each)														
<4> A (0.2) DEFGHIKLMNPQRSTVWY (0.044 each)														
For CDR3 (Length 9):														
<1> ADEFGHIKLMNPQRSTVWY 1:1														
<3> Y (0.2) ADEFGHIKLMNPQRTVW (0.044 each)														

[0165]

TABLE 8

2a2 JH2 Human lambda-chain gene

gaggaccatt gggcccc ttactccgtgac
Scab..... Eco0109I
ApaI..

-----FR1----->

S	A	Q	S	A	L	T	Q	P	A	S	V	S	G	S	P	G
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

agt|gca|caa|tcc|gct|ctc|act|cag|cct|gct|agc|ggt|tcc|ggg|tca|cct|ggt|
ApaLI... NheI... BstEII... SexAI.....

T G <1> S S <2> V G

-----FR1-----> |-----CDR1----->

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Q	S	I	T	I	S	C	T	G	-	S	S	-	V	G

|caa|agt|atc|act|att|tct|tgt|aca|ggt| - |tct|tct| - gtt|ggc|
BsrGI..

<1> <3> <2> <3> V S = vg Scheme #1, length = 14

-----CDR1-----> |-----FR2----->

31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
-	-	-	-	V	S	W	Y	Q	Q	H	P	G	K	A

| - | - | - | - |gtt|tct|tgg|tat|caa|caa|cac|ccg|ggc|aag|gcg|
XmaI.... KasI
AvaI....

<4> <4> <4> <2> R P S

-----FR2-----> |-----CDR2-----> |-----FR3----->

46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
P	K	L	M	I	Y	-	-	-	-	R	P	S	G	V

|ccg|aag|ttg|atg|atc|tac| - | - | - | - |cgt|cct|tct|ggt|gtt|
KasI....

FR3

61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
S	N	R	F	S	G	S	K	S	G	N	T	A	S	L

|agc|aat|cgt|ttc|tcc|gga|tct|aaa|tcc|ggt|aat|acc|gca|agc|tta|
BspEI.. HindIII.

BsaBI.....(blunt)

-----FR3----->1

76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
T	I	S	G	L	Q	A	E	D	E	A	D	Y	Y	C

TABLE 8-continued

2a2 JH2 Human lambda-chain gene																												
act atc tct ggt ctg cag gct gaa gac gag gct gac tac tat tgt																												
PstI...																												
<4> <5> <4> <2> <4> S <4> <4> <4> <4> V																												
-----CDR3----->										---FR4-----																		
91	92	93	94	95	96	97	98	99	100	101	102	103	104	105														
-	-	-	-	-	S	-	-	-	-	V	F	G	G	G														
	-		-		-		tct	-		-		-		gtc	ttc	ggc	ggg	ggg	KpnI...									
-----FR4----->																												
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120														
T	K	L	T	V	L	G	Q	P	K	A	A	P	S	V														
	acc	aaa	ctt	act	gtc	ctc	ggg	caa	cct	aag	gct	gct	cct	tcc	ggt													
KpnI...						HincII...						Bsu36I...																
121	122	123	124	125	126	127	128	129	130	131	132	133	134	135														
T	L	F	P	P	S	S	E	E	L	Q	A	N	K	A														
	act	ctc	ttc	cct	cct	agt	tct	gaa	gag	ctt	caa	gct	aac	aag	gct													
															SapI.....													
136	137	138	139	140	141	142	143	144	145	146	147	148	149	150														
T	L	V	C	L	I	S	D	F	Y	P	G	A	V	T														
	act	ctt	gtt	tgc	ttg	atc	agt	gac	ttt	tat	cct	ggg	gct	gtt	act													
															BclI....													
151	152	153	154	155	156	157	158	159	160	161	162	163	164	165														
V	A	W	K	A	D	S	S	P	V	K	A	G	V	E														
	gtc	gct	tgg	aaa	gcc	gat	tct	tct	cct	gtt	aaa	gct	ggg	gtt	gag													
															BsmBI....													
166	167	168	169	170	171	172	173	174	175	176	177	178	179	180														
T	T	T	P	S	K	Q	S	N	N	K	Y	A	A	S														
	acg	acc	act	cct	tct	aaa	caa	tct	aac	aat	aag	tac	gct	gcg	agc													
BsmBI....															SacI....													
181	182	183	184	185	186	187	188	189	190	191	192	193	194	195														
S	Y	L	S	L	T	P	E	Q	W	K	S	H	K	S														
	tct	tat	ctt	tct	ctc	acc	cct	gaa	caa	tgg	aag	tct	cat	aaa	tcc													
SacI...																												
196	197	198	199	200	201	202	203	204	205	206	207	208	209	210														
Y	S	C	Q	V	T	H	E	G	S	T	V	E	K	T														
	tat	tcc	tgt	caa	gtt	act	cat	gaa	ggg	tct	acc	gtt	gaa	aag	act													
															BspHI...													
211	212	213	214	215	216	217	218	219																				
V	A	P	T	E	C	S	.	.																				
	gtt	gcc	cct	act	gag	tgt	tct	tag	tga	ggcgcgcc																		
															AscI....													
															BssHII													
aacgatgttc aag gcggccgc aacaggaggag																												
NotI.... Scab.....																												

For CDR1 (length 14):

<1> = 0.27 T, 0.27 G, 0.027 each of ADEFHIKLMNPQRSVWY, no C
 <2> = 0.27 D, 0.27 N, 0.027 each of AEFHGIKLMNPQSTVWY, no C
 <3> = 0.36 Y, 0.0355 each of ADEFHGIKLMNPQSTVW, no C
 A second Vg scheme for CDR1 gives segments of length 11:
 T₂₂G<2><4>L<4><4><4><3><4><4> where
 <4> 32 e|uimolar mixture of each of ADEFHGIKLMNPQSTVWY, no C
 <3> 32 as defined above for the alternative CDR1

For CDR2:

<2> and <4> are the same variegation as for CDR1

CDR3 (Length 11):

<2> and <4> are the same variegation as for CDR1

<5> = 0.36 S, 0.0355 each of ADEFHGIKLMNPQRTVWY no C

CDR3 (Length 10): <5> SY <1> <5> S <5> <1> <4> V

<1> is an equimolar mixture of ADEFHGIKLMNPQSTVWY, no C

<4> and <5> are as defined for Length 11

[0166]

TABLE 9

Oligonucleotides For Kappa and Lambda Light Chain Variegation	
(Ctop25):	5'-gctctggtcaac tta agg gct gag g-3'
(CtprmA):	5'-gctctggtcaac tta agg gct gag gac acc gct gtc tac tac tgc gcc-3' AflII...
(CBprmB) [RC]:	5'- tac ttc gat tac ttg ggc caa ggg acc ctg gtc acc tcgctccacc-3' BstEII...
(CBot25) [RC]:	5'- ggg acc ctg gtc acc tcgctccacc-3'
Kappa chains: CDR1 ("1"), CDR2 ("2"), CDR3 ("3")	
CDR1	
(Ka1Top610):	5'-gggtctcagttg cta agc ccg ggg gaa cgt gct acc tta agt tgc cgt gct tcc cag-3'
(Ka1STp615):	5'-gggtctcagttg cta agc ccg ggg g-3'
(Ka1Bot620) [RC]:	5'-ctt gct tgg tat caa cag aaa cct ggg cag gcg ccaagtcgtgtc-3'
(Ka1SB625) [RC]:	5'-cct ggg cag gcg ccaagtcgtgtc-3'
(Ka1vg600):	5'-gct acc tta agt tgc cgt gct tcc cag- <1> gtt <2> <2> <3> ctt gct tgg tat caa cag aaa cc-3'
(Ka1vg600-12):	5'-gct acc tta agt tgc cgt gct tcc cag- <1> gtt <2> <2> <3> ctt gct tgg tat caa cag aaa cc-3'
CDR2	
(Ka2Tshort657):	5'-cacgagtccta cct ggg cag gc-3'
(Ka2Tlong655):	5'-cacgagtccta cct ggg cag gdg ccg cgt tta ctt att tat-3'
(Ka2Bshort660):	[RC]: 5'- gac cgt ttc tct ggg tctcacc-3' cgc <4> <1> ggg atc ccg gac cgt ttc tct ggg tctcacc-3'
CDR3	
(Ka3Tlon672):	5'-gacgagtccttct aga ttg gaa cct gaa gac ttc gct gtt tat tat tgc caa c-3'
(Ka3BotL682) [RC]:	5'-act ttc ggg caa ggg acc aag gtt gaa atc aag cgt acg tcacaggtgag-3'
(Ka3Bsho694) [RC]:	5'-gaa atc aag cgt acg tcacaggtgag-3'
(Ka3vg670):	5'-gac ttc gct gtt - tat tat tgc caa cag <3> <1> <1>a<1> cct <1> act ttc ggg caa - ggg acc aag gtt g-3'
(Ka3vg670-8):	5'-gac ttc gct gtt - tat tat tgc caa cag <3> <3> <1> <1> cct ttc ggg caa - ggg acc aag gtt g-3'
(Ka3vg670-10):	5'-gac ttc gct gtt tat - tat tgc caa cag <3> <2> <1> cct cct <1> act ttc ggg caa - ggg acc aag gtt g-3'
Lambda Chains: CDR1 ("1"), CDR2 ("2"), CDR3 ("3")	
CDR1	
(Lm1TPri75):	5'-gacgagtcctgg tca cct ggg -3'
(Lm1tlo715):	5'-gacgagtcctgg tca cct ggg caa agt atc act att tct tgt aca ggg-3'
(Lm1blo724) [rc]:	5'-gtt tct tgg tat caa caa cac ccg ggc aag gcg agatcttcacaggtgag-3'
(Lm1bsh737) [rc]:	5'-gc aag gcg agatcttcacaggtgag-3'
(Lm1vh710b):	5'-gt atc act att tct tgt aca ggg <2> <4> ctc <4> <4> <4> - <3> <4> <4> tgg tat caa caa cac cc-3'
(Lm1vh710):	5'-gt atc act att tct tgt aca ggg <1> tct tct <2> gtt ggc - <1> <3> <2> <3> gtt tct tgg tat caa caa cac cc-3'

TABLE 9-continued

Oligonucleotides For Kappa and Lambda Light Chain Variegation	
CDR2	
(Lm2TSh757):	5'-gagcagaggac ccg ggc aag gc-3'
(Lm2TL0753):	5'-gagcagaggac ccg ggc aag gcg ccg aag ttg atg atc tac -3'
(Lm2BL0762) [RC]:	5'-cgt cct tct ggt gtc agc aat cgt ttc tcc gga tcacaggtgag-3'
(Lm2BSh765) [RC]:	5'-cgt ttc tcc gga tcacaggtgag-3'
(Lm2vg750):	5'-g ccg aag ttg atg atc tac - <4> <4> <4> <2> cgt cct tct ggt gtc agc aat c-3'
CDR3	
(Lm3TSh822):	5'-ctg cag gct gaa gac gag gct gac-3'
(Lm3TL0819):	5'-ctg cag gct gaa gac gag gct gac tac tat tgt -3'
(Lm3BL0825) [RC]:	5'-gtc ttc ggc ggt ggg acc aaa ctt act gtc ctc ggt caa cct aag g- acacaggtgag-3'
(Lm3BSh832) [RC]:	5'-c ggt caa cct aag gacacaggtgag-3'
(Lm3vg817):	5'-gac gag gct gac tac tat tgt - <4> <5> <4> <2> <4> tct <4> <4> <4> <4> - Gtc ttc ggc ggt ggg acc aaa ctt ac-3'
(Lm3vg817-10):	5'-gac gag gct gac tac tat tgt - <5> agc tat <1> <5> tct<5> <1> <4> gtc ttc ggc ggt ggg - acc aaa ctt ac-3'

[0167]

TABLE 10

A27:JH1 Kappa light chain gene with stuffers in place of CDRs	
Each stuffer contains at least one stop codon and a restriction site that will be unique within the diversity vector.	
gaggacc attgggcccc ctccgagact ctcgagcgca	
Scab.....Eco0109I	
ApaI.	
XhoI..	
acgcaattaa tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc	
...-35.. Plac ...-10.	
cggtcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatgac	
catgatta cgccaagcctt tggagccttt tttttggaga ttttcaac	
PflMI.....	
Hind3.	
M13 III signal sequence (AA se)----->	
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	
M K K L L F A I P L V V P F Y	
gtg aag aag ctc cta ttt gct atc ccg ctt gtc gtt ccg ttt tac	
--Signal--> FR1----->	
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
S H S A Q S V L T Q S P G T L	
agc cat agt gca caa tcc gtc ctt act caa tct cct ggc act ctt	
ApaLI...	
----- FR1 -----> -----Stuffer-->	
31 32 33 34 35 36 37 38 39 40 41 42 43	
S L S P G E R A T L S	
tcg cta agc ccg ggg gaa cgt gct acc tta agt tag taa gct ccc	
EspI..... AflII...	

TABLE 10-continued

A27:JH1 Kappa light chain gene with stuffers in place of CDRs	
XmaI....	
- Stuffer for CDR1-->FR2 ----- FR2 -----> -----Stuffer for CDR2	
59 60 61 62 63 64 65 66	
K P G Q A P R	
agg cct actt tga tct g aaa cct ggt cag gcg ccg cgt taa tga aagcgctaataggccaacagtg	
StuI... SexAI... KasI.... AfeI.. MscI..	
Stuffer--> --- FR3 ----->	
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90	
T G I P D R F S G S G S G T D	
act ggg atc ccg gac ccgt ttc tct ggc tct ggt tca ggt act gac	
BamHI... RsrII.....	
----- FR3 -----> -----STUFFER for CDR3----->	
91 92 93 94 95 96 97	
F T L T I S R	
ttt acc ctt act att tct aga taa tga gttaac tag acc tacgta acc tag	
XbaI... HpaI.. SnaBI.	
-----CDR3 stuffer-----> -----FR4--->	
118 119 120	
F G Q	
ttc ggt caa	
-----FR4-----> <----- Ckappa -----	
121 122 123 124 125 126 127	128 129 130 131 132 133 134
G T K V E I K	R T V A A P S
ggt acc aag gtt gaa atc aag	cgt acg gtt gcc gct cct agt
StyI.... BsiWI..	
135 136 137 138 139 140 141 142 143 144 145 146 147 148 149	
V F I F P P S D E Q L K S G T	
gtg ttt atc ttt cct cct tct gac gaa caa ttg aag tca ggt act	
MfeI...	
acgcacatctctaa gcggccgc aacaggaggag	
NotI....	
EagI..	

[0168]

TABLE 11

2a2:JH2 Human lambda-chain gene with stuffers in place of CDRs	
gaggaccatt gggcccc ttactccgtgac	
Scab..... Eco0109I	
ApaI..	
-----FR1----->	
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	
S A Q S A L T Q P A S V S G S P G	
agt gca caa tcc gct ctc act cag cct gct agc gtt tcc ggg tca cct ggt	
ApaLI... NheI... BstEII... SexAI....	
-----FR1-----> -----stuffer for CDR1-----	
16 17 18 19 20 21 22 23	
Q S I T I S C T	
caa agt atc act att tct tgt aca tct tag tga ctc	
BsrGI..	
-----Stuffer-----> -----FR2----->	
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45	
R S P	H P G K A
aga tct taa tga ccg tag	cac ccg ggc aag gcg
BglII XmaI.... KasI....	
AvaI....	

TABLE 11-continued

2a2:JH2 Human lambda-chain gene with stuffers in place of CDRs

-----Stuffer for CDR2 ----->

P

|ccg|taa|tga|atc tgc tac g ct|ggg|gtt|

KasI.... BsiWI...

-----FR3-----

61 62 63 64 65 66 67 68 69 70 71 72 73 74 75

S N R F S G S K S G N T A S L

|agc|aat|cgt|ttc|tcc|gga|tct|aaa|tcc|ggg|aat|acc|gca|agc|tta|

BspEI.. HindIII.

BsaBI.....(blunt)

-----FR3----->|-----Stuffer for CDR3----->|

76 77 78 79 80 81 82 83 84 85 86 87 88 89 90

T I S G L Q

|act|atc|tct|ggg|ctg|cag|gtt ctg tag ttc caattg ctt tag tga ccc

PstI... MfeI..

-----Stuffer----->|-----FR4-----

103 104 105

G G G

|ggc|ggg|ggg|

KpnI...

-----FR4----->

106 107 108 109 110 111 112 113 114 115 116 117 118 119 120

T K L T V L G Q P K A A P S V

|acc|aaa|ctt|act|gtc|ctc|ggg|caa|cct|aag|gct|gct|cct|tcc|gtt|

KpnI... HincII..

Bsu36I...

121 122 123 124 125 126 127 128 129 130 131 132 133 134 135

T L F P P S S E E L Q A N K A

|act|ctc|ttc|cct|cct|agt|tct|gaa|gag|ctt|caa|gct|aac|aag|gct|

SapI.....

136 137 138 139 140 141 142 143 144 145 146 147 148 149 150

T L V C L I S D F Y P G A V T

|act|ctt|gtt|tgc|ttg|atc|agt|gac|ttt|tat|cct|ggg|gct|gtt|act|

BclI....

[0169]

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 99

<210> SEQ ID NO 1

<211> LENGTH: 14

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain CDR1 vector

<220> FEATURE:

<221> NAME/KEY: MOD_RES

<222> LOCATION: (8)..(10)

<223> OTHER INFORMATION: Any amino acid except Cys

<220> FEATURE:

<221> NAME/KEY: MOD_RES

<222> LOCATION: (14)

<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 1

Val Ser Gly Gly Ser Ile Ser Xaa Xaa Xaa Tyr Tyr Trp Xaa
 1 5 10

-continued

<210> SEQ ID NO 2
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
CDR2 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)
<223> OTHER INFORMATION: Tyr, Arg, Trp, Val Gly or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)
<223> OTHER INFORMATION: Tyr, Arg, Trp, Val, Gly or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)
<223> OTHER INFORMATION: Pro, Ser or Gly
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (10)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 2

Xaa Ile Xaa Xaa Ser Gly Gly Xaa Thr Xaa Tyr Ala Asp Ser Val Lys
1 5 10 15

Gly

<210> SEQ ID NO 3
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
CDR2 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)
<223> OTHER INFORMATION: Asp, Ile, Asn, Ser, Trp or Tyr
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)..(5)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (7)
<223> OTHER INFORMATION: Ser, Gly, Asp or Asn
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)..(10)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 3

Xaa Ile Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Tyr Ala Asp Ser Val Lys
1 5 10 15

Gly

<210> SEQ ID NO 4
<211> LENGTH: 16

-continued

<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
CDR2 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)
<223> OTHER INFORMATION: Asp, Ile, Asn, Ser, Trp or Tyr
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)..(5)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (7)
<223> OTHER INFORMATION: Ser, Gly, Asp or Asn
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)..(9)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 4

Xaa Ile Xaa Xaa Xaa Gly Xaa Xaa Xaa Tyr Asn Pro Ser Leu Lys Gly
1 5 10 15

<210> SEQ ID NO 5
<211> LENGTH: 19
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
CDR2 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (5)..(7)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (12)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 5

Xaa Ile Xaa Ser Xaa Xaa Xaa Gly Gly Tyr Tyr Xaa Tyr Ala Ala Ser
1 5 10 15

Val Lys Gly

<210> SEQ ID NO 6
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
CDR3 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (5)
<223> OTHER INFORMATION: Lys or Arg
<220> FEATURE:

<400> SEQUENCE: 6

<400> SEQUENCE: 7

Gly

<400> SEQUENCE: 8

Tyr Trp Gly

```
<210> SEQ ID NO 9
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
      CDR3 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)..(8)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (10)
<223> OTHER INFORMATION: Ser or Gly
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (12)
```

-continued

<223> OTHER INFORMATION: Tyr or Trp
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (13)..(15)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 9

Tyr Tyr Cys Ala Arg Xaa Xaa Xaa Ser Xaa Ser Xaa Xaa Xaa Xaa Tyr
1 5 10 15

Phe Asp Tyr Trp Gly
20

<210> SEQ ID NO 10
<211> LENGTH: 22
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
CDR3 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (5)
<223> OTHER INFORMATION: Lys or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)..(8)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (12)..(13)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (16)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 10

Tyr Tyr Cys Ala Xaa Xaa Xaa Xaa Cys Ser Gly Xaa Xaa Cys Tyr Xaa
1 5 10 15

Tyr Phe Asp Tyr Trp Gly
20

<210> SEQ ID NO 11
<211> LENGTH: 24
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
CDR3 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (5)
<223> OTHER INFORMATION: Lys or Arg
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)..(7)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (14)..(18)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 11

Tyr Tyr Cys Ala Xaa Xaa Xaa Ser Xaa Thr Ile Phe Gly Xaa Xaa Xaa
1 5 10 15

-continued

Xaa Xaa Tyr Phe Asp Tyr Trp Gly
20

<210> SEQ ID NO 12
 <211> LENGTH: 25
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
 CDR3 vector
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (6)..(8)
 <223> OTHER INFORMATION: Any amino acid except Cys
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (11)
 <223> OTHER INFORMATION: Asp or Gly
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (13)..(14)
 <223> OTHER INFORMATION: Ser or Gly
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (15)..(16)
 <223> OTHER INFORMATION: Any amino acid
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (17)..(19)
 <223> OTHER INFORMATION: Any amino acid except Cys
 <400> SEQUENCE: 12

Tyr Tyr Cys Ala Arg Xaa Xaa Xaa Tyr Tyr Xaa Ser Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Tyr Phe Asp Tyr Trp Gly
20 25

<210> SEQ ID NO 13
 <211> LENGTH: 26
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Heavy chain
 CDR3 vector
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (6)..(9)
 <223> OTHER INFORMATION: Any amino acid except Cys
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (12)..(13)
 <223> OTHER INFORMATION: Ser or Gly
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (14)
 <223> OTHER INFORMATION: Thr, Asp or Gly
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (15)
 <223> OTHER INFORMATION: Any amino acid except Cys
 <220> FEATURE:
 <221> NAME/KEY: MOD_RES
 <222> LOCATION: (18)..(20)
 <223> OTHER INFORMATION: Any amino acid except Cys
 <400> SEQUENCE: 13

Tyr Tyr Cys Ala Arg Xaa Xaa Xaa Xaa Tyr Cys Xaa Xaa Xaa Xaa Cys
1 5 10 15

Tyr Xaa Xaa Xaa Tyr Phe Asp Tyr Trp Gly

-continued

20 25

<210> SEQ ID NO 14
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Kappa light
chain CDR1 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (5)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (7)..(8)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)
<223> OTHER INFORMATION: Any amino acid except Cys and Ser

<400> SEQUENCE: 14

Arg Ala Ser Gln Xaa Val Xaa Xaa Xaa Leu Ala
1 5 10

<210> SEQ ID NO 15
<211> LENGTH: 12
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Kappa light
chain CDR1 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (5)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (7)..(9)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (10)
<223> OTHER INFORMATION: Any amino acid except Cys and Ser

<400> SEQUENCE: 15

Arg Ala Ser Gln Xaa Val Xaa Xaa Xaa Xaa Leu Ala
1 5 10

<210> SEQ ID NO 16
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Kappa light
chain CDR3 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)
<223> OTHER INFORMATION: Any amino acid except Cys or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)..(6)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (8)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 16

-continued

Gln Gln Xaa Xaa Xaa Xaa Pro Xaa Thr
1 5

<210> SEQ ID NO 17
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Kappa light
chain CDR3 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)
<223> OTHER INFORMATION: Any amino acid except Cys or Ser
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (5)..(6)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 17

Gln Gln Xaa Xaa Xaa Xaa Pro Pro Xaa Thr
1 5 10

<210> SEQ ID NO 18
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Lambda
light chain CDR1 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (3)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (6)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)..(12)
<223> OTHER INFORMATION: Any amino acid except Cys

<400> SEQUENCE: 18

Thr Gly Xaa Ser Ser Xaa Val Gly Xaa Xaa Xaa Val Ser
1 5 10

<210> SEQ ID NO 19
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Lambda
light chain CDR3 vector
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (1)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (4)..(5)
<223> OTHER INFORMATION: Any amino acid except Cys

-continued

<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (7)..(8)
<223> OTHER INFORMATION: Any amino acid except Cys
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (9)
<223> OTHER INFORMATION: Any amino acid except Cys or Tyr

<400> SEQUENCE: 19

Xaa Ser Tyr Xaa Xaa Ser Xaa Xaa Xaa Val
1 5 10

<210> SEQ ID NO 20
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Ser Ser
1 5 10

<210> SEQ ID NO 21
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

Ala Phe Asp Ile Trp Gly Gln Gly Thr Met Val Thr Val Ser Ser
1 5 10 15

<210> SEQ ID NO 22
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 22

Tyr Tyr Cys Ala Arg
1 5

<210> SEQ ID NO 23
<211> LENGTH: 323
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: 3-23: JH4 vector with CDR1/2 diversity
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (3)..(323)
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (99)..(101)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (105)..(107)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (111)..(113)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (156)..(158)

-continued

```

<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (162)..(167)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (177)..(179)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (183)..(185)
<223> OTHER INFORMATION: a, c, t, g, other or unknown

<400> SEQUENCE: 23

cc atg gcc gaa gtt caa ttg tta gag tct ggt ggc ggt ctt gtt cag      47
Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
  1             5             10             15

cct ggt ggt tct tta cgt ctt tct tgc gct gct tcc gga ttc act ttc      95
Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
      20             25             30

tct nnn tac nnn atg nnn tgg gtt cgc caa gct cct ggt aaa ggt ttg      143
Ser Xaa Tyr Xaa Met Xaa Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
      35             40             45

gag tgg gtt tct nnn atc nnn nnn tct ggt ggc nnn act nnn tat gct      191
Glu Trp Val Ser Xaa Ile Xaa Xaa Ser Gly Gly Xaa Thr Xaa Tyr Ala
      50             55             60

gac tcc gtt aaa ggt cgc ttc act atc tct aga gac aac tct aag aat      239
Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn
      65             70             75             80

act ctc tac ttg cag atg aac agc tta agg gct gag gac acc gct gtc      287
Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
      85             90             95

tac tac tgc gcc aaa gac tat gaa ggt act ggt tat                      323
Tyr Tyr Cys Ala Lys Asp Tyr Glu Gly Thr Gly Tyr
      100             105

<210> SEQ ID NO 24
<211> LENGTH: 108
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: 3-23: JH4
vector with CDR1/2 diversity
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (34)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (36)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (38)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (53)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (55)..(56)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (60)
<223> OTHER INFORMATION: Any amino acid

```

-continued

<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (62)
<223> OTHER INFORMATION: Any amino acid

<400> SEQUENCE: 24

Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
1 5 10 15
Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
20 25 30
Ser Xaa Tyr Xaa Met Xaa Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
35 40 45
Glu Trp Val Ser Xaa Ile Xaa Xaa Ser Gly Gly Xaa Thr Xaa Tyr Ala
50 55 60
Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn
65 70 75 80
Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
85 90 95
Tyr Tyr Cys Ala Lys Asp Tyr Glu Gly Thr Gly Tyr
100 105

<210> SEQ ID NO 25
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(45)

<400> SEQUENCE: 25

tat ttc gat tat tgg ggt caa ggt acc ctg gtc acc gtc tct agt 45
Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
1 5 10 15

<210> SEQ ID NO 26
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 26

Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
1 5 10 15

<210> SEQ ID NO 27
<211> LENGTH: 55
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (21)..(23)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (27)..(29)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (33)..(35)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 27

-continued

cttcggtt cactttctct nntacnna tgnntgggt tcgccaagct cctgg 55

<210> SEQ ID NO 28
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 28

cctactgtct tccgattca ctttctct 28

<210> SEQ ID NO 29
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 29

tgggttcgcc aagctcctgg ttgctcactc 30

<210> SEQ ID NO 30
<211> LENGTH: 61
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (23)..(29)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (39)..(41)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 30

cttcggtt cactttctct wnnnnnnnt actactggn ntgggttcgc caagctcctg 60

g 61

<210> SEQ ID NO 31
<211> LENGTH: 82
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (42)..(50)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (60)..(62)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 31

cttcggtt cactttctct atcagcgtg gttctatctc cnnnnnnnn tactactggn 60

ntgggttcg ccaagctcct gg 82

-continued

```

<210> SEQ ID NO 32
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (19)..(21)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (25)..(28)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (30)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (40)..(42)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (46)..(48)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 32

ggtttggagt gggtttctnn nactnnnnsn tctggtggcn nactnnnta tgetgactcc      60
gttaaagg                                         68

<210> SEQ ID NO 33
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide

<400> SEQUENCE: 33

cttgggttcg ccaagctcct ggtaaagggt tggagtgggt ttct                        44

<210> SEQ ID NO 34
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide

<400> SEQUENCE: 34

tatgtgact ccgttaaagg tcgttcact atctctagat tcctgtcac                    49

<210> SEQ ID NO 35
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (19)..(21)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (26)..(33)
<223> OTHER INFORMATION: a, c, t or g

```

-continued

```

<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (39)..(48)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 35

ggtttgaggt gggtttctnn natcdnnnnn nnnggtdvnn nnnnnnnnta tgctgactcc      60
gttaaagg                                           68

<210> SEQ ID NO 36
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
        oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (19)..(21)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (26)..(33)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (39)..(45)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 36

ggtttgaggt gggtttctnn natcdnnnnn nnnggtdvnn nnnntataa cccttcctt      60
aaggg                                           65

<210> SEQ ID NO 37
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
        oligonucleotide

<400> SEQUENCE: 37

tataaccctt ccottaagggt tcgcttcact atctctagat tcctgtcac      49

<210> SEQ ID NO 38
<211> LENGTH: 74
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
        oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (19)..(21)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (25)..(27)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (31)..(39)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (52)..(54)
<223> OTHER INFORMATION: a, c, t or g

```

-continued

<400> SEQUENCE: 38

ggtttgaggt gggtttctnn nactcnnnagt nnnnnnnnng gtggtactac tnnntatgcc 60
gcttccgtta aggg 74

<210> SEQ ID NO 39

<211> LENGTH: 49

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 39

tatgccgctt ccgtaaggg tcgcttcact atctctagat tcctgtcac 49

<210> SEQ ID NO 40

<211> LENGTH: 25

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 40

gctctggtca acttaagggc tgagg 25

<210> SEQ ID NO 41

<211> LENGTH: 48

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 41

gctctggtca acttaagggc tgaggacacc gctgtctact actgcgcc 48

<210> SEQ ID NO 42

<211> LENGTH: 46

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 42

tacttcgatt actggggcca aggtaccctg gtcacctcgc tccacc 46

<210> SEQ ID NO 43

<211> LENGTH: 25

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 43

ggtaccctgg tcacctcgct ccacc 25

<210> SEQ ID NO 44

<211> LENGTH: 58

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (23)..(35)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 44

ccgctgtcta ctactgcgcc mrnnnnnnnn nnnntactt cgattactgg ggccaagg 58

<210> SEQ ID NO 45
<211> LENGTH: 64
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (23)..(41)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 45

ccgctgtcta ctactgcgcc mrnnnnnnnn nnnnnnnnnn ntacttcgat tactggggcc 60
aagg 64

<210> SEQ ID NO 46
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (23)..(47)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 46

ccgctgtcta ctactgcgcc mrnnnnnnnn nnnnnnnnnn nnnnnntac ttcgattact 60
ggggccaagg 70

<210> SEQ ID NO 47
<211> LENGTH: 76
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (24)..(32)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (38)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (45)..(53)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 47

ccgctgtcta ctactgcgcc cgtnnnnnnn nntctdsntc ttrbnnnnnn nnntacttcg 60
attactgggg ccaagg 76

-continued

<210> SEQ ID NO 48
<211> LENGTH: 79
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (23)..(32)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (42)..(47)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (54)..(56)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 48

ccgctgtcta ctactgcgcc mrnnnnnnnn nntgctctgg tnnnnntgc tatnnntact 60

tcgattactg gggccaagg 79

<210> SEQ ID NO 49
<211> LENGTH: 85
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (23)..(29)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (33)..(35)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (48)..(62)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 49

ccgctgtcta ctactgcgcc mrnnnnnnnt ctnnnactat cttcggtnnn nnnnnnnnnn 60

nntacttcga ttactggggc caagg 85

<210> SEQ ID NO 50
<211> LENGTH: 88
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (24)..(32)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (41)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (47)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:

-continued

```

<221> NAME/KEY: modified_base
<222> LOCATION: (50)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (57)..(65)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 50

ccgctgtcta ctactgcgcc cgtnnnnnnn nntattacgr ntctdsndsn tactatnnnn      60
nnnnntactt cgattactgg ggccaagg                                           88

<210> SEQ ID NO 51
<211> LENGTH: 91
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
        oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (24)..(35)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (44)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (47)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (50)..(53)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (60)..(68)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 51

ccgctgtcta ctactgcgcc cgtnnnnnnn nnnntatttg cdsndsnrvn nntgtctatn      60
nnnnnnnnta cttcgattac tggggccaag g                                       91

<210> SEQ ID NO 52
<211> LENGTH: 242
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: 3-23: JH4
        vector with stuffers
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (3)..(107)
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (114)..(155)
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (159)..(164)
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (168)..(227)

<400> SEQUENCE: 52

cc atg gcc gaa gtt caa ttg tta gag tct ggt ggc ggt ctt gtt cag      47
Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
  1             5             10             15

```

-continued

```

cct ggt ggt tct tta cgt ctt tct tgc gct gct tcc gga ttc act ttc      95
Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
      20              25              30

tct tcg tac gct tagtaa tgg gtt cgc caa gct cct ggt aaa ggt ttg      143
Ser Ser Tyr Ala      Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
      35              40              45

gag tgg gtt tct taa cct agg tag act atc tct aga gac aac tct aag      191
Glu Trp Val Ser      Pro Arg      Thr Ile Ser Arg Asp Asn Ser Lys
      50              55              60

aat act ctc tac ttg cag atg aac agc tta agg gct tagtaaaggc cttaa      242
Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala
      65              70

```

```

<210> SEQ ID NO 53
<211> LENGTH: 72
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: 3-23: JH4
      vector with stuffers

```

```

<400> SEQUENCE: 53

```

```

Ala Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln
  1              5              10              15

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe
      20              25              30

Ser Ser Tyr Ala Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp
      35              40              45

Val Ser Pro Arg Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
      50              55              60

Leu Gln Met Asn Ser Leu Arg Ala
  65              70

```

```

<210> SEQ ID NO 54
<211> LENGTH: 952
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (206)..(907)
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (344)..(346)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (350)..(361)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (413)..(415)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (422)..(424)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (428)..(433)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (536)..(547)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base

```

-continued

<222> LOCATION: (551)..(553)

<223> OTHER INFORMATION: a, c, t, g, other or unknown

<400> SEQUENCE: 54

```

gaggaccatt gggccccctc cgagactctc gagcgcaacg caattaatgt gagttagctc      60
actcattagg caccaccaggc tttacacttt atgcttcggt ctcgtatggt gtgtggaatt      120
gtgagcggat aacaatttca cacaggaaac agctatgacc atgattacgc caagcttttg      180
agcctttttt ttggagattt tcaac  gtg aag aag ctc cta ttt gct atc ccg      232
                        Met Lys Lys Leu Leu Phe Ala Ile Pro
                        1             5

ctt gtc gtt ccg ttt tac agc cat agt gca caa tcc gtc ctt act caa      280
Leu Val Val Pro Phe Tyr Ser His Ser Ala Gln Ser Val Leu Thr Gln
 10             15             20             25

tct cct ggc act ctt tcg cta agc ccg ggt gaa cgt gct acc tta agt      328
Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser
             30             35             40

tgc cgt gct tcc cag nnn gtt nnn nnn nnn nnn ctt gct tgg tat caa      376
Cys Arg Ala Ser Gln Xaa Val Xaa Xaa Xaa Xaa Leu Ala Trp Tyr Gln
             45             50             55

cag aaa cct ggt cag gcg ccg cgt tta ctt att tat nnn gct tct nnn      424
Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile Tyr Xaa Ala Ser Xaa
 60             65             70

cgc nnn nnn ggg atc ccg gac cgt ttc tct ggc tct ggt tca ggt act      472
Arg Xaa Xaa Gly Ile Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr
 75             80             85

gac ttt acc ctt act att tct aga ttg gaa cct gaa gac ttc gct gtt      520
Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Phe Ala Val
 90             95             100             105

tat tat tgc caa cag nnn nnn nnn nnn cct nnn act ttc ggt caa ggt      568
Tyr Tyr Cys Gln Gln Xaa Xaa Xaa Xaa Pro Xaa Thr Phe Gly Gln Gly
110             115             120

acc aag gtt gaa atc aag cgt acg gtt gcc gct cct agt gtg ttt atc      616
Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile
125             130             135

ttt cct cct tct gac gaa caa ttg aag tca ggt act gct tct gtc gta      664
Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val
140             145             150

tgt ttg ctc aac aat ttc tac cct cgt gaa gct aaa gtt cag tgg aaa      712
Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys
155             160             165

gtc gat aac gcg ttg cag tcg ggt aac agt caa gaa tcc gtc act gaa      760
Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu
170             175             180             185

cag gat agt aag gac tct acc tac tct ttg tcc tct act ctt act tta      808
Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu
190             195             200

tca aag gct gat tat gag aag cat aag gtc tat gct tgc gaa gtt acc      856
Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr
205             210             215

cac cag ggt ctg agc tcc cct gtt acc aaa agt ttc aac cgt ggt gaa      904
His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu
220             225             230

tgc taatagggcg cgccacgcat ctctaagcgg ccgcaacagg aggag      952
Cys

```

-continued

```

<210> SEQ ID NO 55
<211> LENGTH: 234
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (47)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (49)..(52)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (70)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (73)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (75)..(76)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (111)..(114)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (116)
<223> OTHER INFORMATION: Any amino acid

<400> SEQUENCE: 55

Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
 1             5             10             15

His Ser Ala Gln Ser Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu
      20             25             30

Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Xaa Val
      35             40             45

Xaa Xaa Xaa Xaa Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro
      50             55             60

Arg Leu Leu Ile Tyr Xaa Ala Ser Xaa Arg Xaa Xaa Gly Ile Pro Asp
      65             70             75             80

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser
      85             90             95

Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Xaa Xaa
      100            105            110

Xaa Xaa Pro Xaa Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
      115            120            125

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
      130            135            140

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
      145            150            155            160

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
      165            170            175

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
      180            185            190

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
      195            200            205

His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
      210            215            220

```

-continued

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
225 230

<210> SEQ ID NO 56
<211> LENGTH: 732
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (30)..(686)
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (108)..(110)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (117)..(119)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (126)..(137)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (189)..(200)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (306)..(320)
<223> OTHER INFORMATION: a, c, t, g, other or unknown
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (324)..(335)
<223> OTHER INFORMATION: a, c, t, g, other or unknown

<400> SEQUENCE: 56

gaggaccatt gggccctta ctccgtgac agt gca caa tcc gct ctc act cag	53
Ser Ala Gln Ser Ala Leu Thr Gln	
1 5	
cct gct agc gtt tcc ggg tca cct ggt caa agt atc act att tct tgt	101
Pro Ala Ser Val Ser Gly Ser Pro Gly Gln Ser Ile Thr Ile Ser Cys	
10 15 20	
aca ggt nnn tct tct nnn gtt ggc nnn nnn nnn gtt tct tgg tat	149
Thr Gly Xaa Ser Ser Xaa Val Gly Xaa Xaa Xaa Val Ser Trp Tyr	
25 30 35 40	
caa caa cac ccg ggc aag gcg ccg aag ttg atg atc tac nnn nnn nnn	197
Gln Gln His Pro Gly Lys Ala Pro Lys Leu Met Ile Tyr Xaa Xaa Xaa	
45 50 55	
nnn cgt cct tct ggt gtt agc aat cgt ttc tcc gga tct aaa tcc ggt	245
Xaa Arg Pro Ser Gly Val Ser Asn Arg Phe Ser Gly Ser Lys Ser Gly	
60 65 70	
aat acc gca agc tta act atc tct ggt ctg cag gct gaa gac gag gct	293
Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu Gln Ala Glu Asp Glu Ala	
75 80 85	
gac tac tat tgt nnn nnn nnn nnn tct nnn nnn nnn gtc ttc	341
Asp Tyr Tyr Cys Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa Xaa Val Phe	
90 95 100	
ggc ggt ggt acc aaa ctt act gtc ctc ggt caa cct aag gct gct cct	389
Gly Gly Gly Thr Lys Leu Thr Val Leu Gly Gln Pro Lys Ala Ala Pro	
105 110 115 120	
tcc gtt act ctc ttc cct cct agt tct gaa gag ctt caa gct aac aag	437
Ser Val Thr Leu Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys	
125 130 135	
gct act ctt gtt tgc ttg atc agt gac ttt tat cct ggt gct gtt act	485

-continued

Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr Pro Gly Ala Val Thr	
140 145 150	
gtc gct tgg aaa gcc gat tct tct cct gtt aaa gct ggt gtt gag acg	533
Val Ala Trp Lys Ala Asp Ser Ser Pro Val Lys Ala Gly Val Glu Thr	
155 160 165	
acc act cct tct aaa caa tct aac aat aag tac gct gcg agc tct tat	581
Thr Thr Pro Ser Lys Gln Ser Asn Asn Lys Tyr Ala Ala Ser Ser Tyr	
170 175 180	
ctt tct ctc acc cct gaa caa tgg aag tct cat aaa tcc tat tcc tgt	629
Leu Ser Leu Thr Pro Glu Gln Trp Lys Ser His Lys Ser Tyr Ser Cys	
185 190 195 200	
caa gtt act cat gaa ggt tct acc gtt gaa aag act gtt gcc cct act	677
Gln Val Thr His Glu Gly Ser Thr Val Glu Lys Thr Val Ala Pro Thr	
205 210 215	
gag tgt tct tagtgaggcg cgccaacgat gttcaaggcg gccgcaacag gaggag	732
Glu Cys Ser	

```

<210> SEQ ID NO 57
<211> LENGTH: 219
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (27)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (30)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (33)..(36)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (54)..(57)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (93)..(97)
<223> OTHER INFORMATION: Any amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (99)..(102)
<223> OTHER INFORMATION: Any amino acid

```

```

<400> SEQUENCE: 57

```

Ser Ala Gln Ser Ala Leu Thr Gln Pro Ala Ser Val Ser Gly Ser Pro	
1 5 10 15	
Gly Gln Ser Ile Thr Ile Ser Cys Thr Gly Xaa Ser Ser Xaa Val Gly	
20 25 30	
Xaa Xaa Xaa Xaa Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro	
35 40 45	
Lys Leu Met Ile Tyr Xaa Xaa Xaa Xaa Arg Pro Ser Gly Val Ser Asn	
50 55 60	
Arg Phe Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser	
65 70 75 80	
Gly Leu Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Xaa Xaa Xaa Xaa	
85 90 95	
Xaa Ser Xaa Xaa Xaa Xaa Val Phe Gly Gly Gly Thr Lys Leu Thr Val	
100 105 110	
Leu Gly Gln Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro Pro Ser	

-continued

115	120	125
Ser Glu Glu Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser		
130	135	140
Asp Phe Tyr Pro Gly Ala Val Thr Val Ala Trp Lys Ala Asp Ser Ser		
145	150	155 160
Pro Val Lys Ala Gly Val Glu Thr Thr Thr Pro Ser Lys Gln Ser Asn		
	165	170 175
Asn Lys Tyr Ala Ala Ser Ser Tyr Leu Ser Leu Thr Pro Glu Gln Trp		
	180	185 190
Lys Ser His Lys Ser Tyr Ser Cys Gln Val Thr His Glu Gly Ser Thr		
	195	200 205
Val Glu Lys Thr Val Ala Pro Thr Glu Cys Ser		
210	215	

<210> SEQ ID NO 58
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> SEQUENCE: 58

gctctggtca acttaagggc tgagg

25

<210> SEQ ID NO 59
 <211> LENGTH: 48
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> SEQUENCE: 59

gctctggtca acttaagggc tgaggacacc gctgtctact actgcgcc

48

<210> SEQ ID NO 60
 <211> LENGTH: 46
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> SEQUENCE: 60

tacttcgatt acttgggcca aggtaccctg gtcacctcgc tccacc

46

<210> SEQ ID NO 61
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> SEQUENCE: 61

ggtaccctgg tcacctcgct ccacc

25

<210> SEQ ID NO 62
 <211> LENGTH: 56
 <212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 62

ggtctcagtt gctaagcccg ggtgaacgtg ctaccttaag ttgccgtgct tcccag 56

<210> SEQ ID NO 63
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 63

ggtctcagtt gctaagcccg ggtg 24

<210> SEQ ID NO 64
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 64

cttgcttggt atcaacagaa acctgggtcag gcgccaagtc gtgtc 45

<210> SEQ ID NO 65
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 65

cctgggtcagg cgccaagtcg tgtc 24

<210> SEQ ID NO 66
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (28)..(30)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (34)..(42)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 66

gtaccttaa gttgccgtgc ttcccagnnn gttnnnnnnn nncttgcttg gtatcaacag 60

aaacc 65

<210> SEQ ID NO 67
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (28)..(30)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (34)..(45)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 67

gctaccttaa gttgccgtgc ttcccagnnn gttnnnnnnn nnnnncttgc ttggtatcaa 60
cagaaacc 68

<210> SEQ ID NO 68
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 68

cacgagtcct acctgggtcag gc 22

<210> SEQ ID NO 69
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 69

cacgagtcct acctgggtcag gcgccggtt tacttattta t 41

<210> SEQ ID NO 70
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 70

gaccgtttct ctggttctca cc 22

<210> SEQ ID NO 71
<211> LENGTH: 76
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (25)..(27)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (34)..(36)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (40)..(45)

-continued

<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 71

caggcgccgc gtttacttat ttatnnngct tctnnncgcn nnnnngggat cccggaccgt 60

ttctctggtt ctcacc 76

<210> SEQ ID NO 72

<211> LENGTH: 53

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 72

gacgagtcct tctagattgg aacctgaaga cttcgctggt tattattgcc aac 53

<210> SEQ ID NO 73

<211> LENGTH: 50

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 73

actttcggtc aaggtagcaa ggttgaaatc aagcgtagct cacaggtgag 50

<210> SEQ ID NO 74

<211> LENGTH: 26

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 74

gaaatcaagc gtacgtcaca ggtgag 26

<210> SEQ ID NO 75

<211> LENGTH: 70

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<220> FEATURE:

<221> NAME/KEY: modified_base

<222> LOCATION: (28)..(39)

<223> OTHER INFORMATION: a, c, t or g

<220> FEATURE:

<221> NAME/KEY: modified_base

<222> LOCATION: (43)..(45)

<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 75

gacttcgctg ttattatttg ccaacagnnn nnnnnnnnnc ctnnnacttt cggtagaagg 60

accaaggttg 70

<210> SEQ ID NO 76

<211> LENGTH: 67

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

-continued

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (28)..(42)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 76

gacttcgctg tttattattg ccaacagnnn nnnnnnnnnn nncctttcgg tcaaggtacc 60

aagggttg 67

<210> SEQ ID NO 77
<211> LENGTH: 73
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (28)..(39)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (46)..(48)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 77

gacttcgctg tttattattg ccaacagnnn nnnnnnnnnc ctctnnnac tttcggtcaa 60

ggtaccaagg ttg 73

<210> SEQ ID NO 78
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 78

gacgagtcct ggtcacctgg t 21

<210> SEQ ID NO 79
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 79

gacgagtcct ggtcacctgg tcaaagtatc actatttctt gtacaggt 48

<210> SEQ ID NO 80
<211> LENGTH: 50
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 80

gtttcttggt atcaacaaca cccgggcaag gcgagatctt cacaggtgag 50

-continued

<210> SEQ ID NO 81
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 81

gcaaggcgag atcttcacag gtgag 25

<210> SEQ ID NO 82
<211> LENGTH: 67
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (24)..(29)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (33)..(50)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 82

gtatcactat ttcttgtaca ggtnnnnnnc tcnnnnnnnn nnnnnnnnnn tggatcaac 60
aacacccc 67

<210> SEQ ID NO 83
<211> LENGTH: 76
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (24)..(26)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (33)..(35)
<223> OTHER INFORMATION: a, c, t or g
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (42)..(53)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 83

gtatcactat ttcttgtaca ggtnnntctt ctnnngttgg cnnnnnnnnn nnnngtttctt 60
ggtatcaaca acacccc 76

<210> SEQ ID NO 84
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 84

gagcagagga cccgggcaag gc 22

-continued

<210> SEQ ID NO 85
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 85

gagcagagga cccgggcaag gcgccgaagt tgatgatcta c 41

<210> SEQ ID NO 86
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 86

cgctccttctg gtgtcagcaa tcgtttctcc ggatcacagg tgag 44

<210> SEQ ID NO 87
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 87

cgttttctccg gatcacaggt gag 23

<210> SEQ ID NO 88
<211> LENGTH: 53
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (20)..(31)
<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 88

gccgaagttg atgatctacn nnnnnnnnnn ncgtccttct ggtgtcagca atc 53

<210> SEQ ID NO 89
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 89

ctgcaggctg aagacgaggc tgac 24

<210> SEQ ID NO 90
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

-continued

<400> SEQUENCE: 90

ctgcaggctg aagacgaggc tgactactat tgt

33

<210> SEQ ID NO 91

<211> LENGTH: 57

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 91

gtcttcgctg gtggtaccaa acttactgtc ctcggtcaac ctaaggacac aggtgag

57

<210> SEQ ID NO 92

<211> LENGTH: 25

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 92

cgggtcaacct aaggacacag gtgag

25

<210> SEQ ID NO 93

<211> LENGTH: 77

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<220> FEATURE:

<221> NAME/KEY: modified_base

<222> LOCATION: (22)..(36)

<223> OTHER INFORMATION: a, c, t or g

<220> FEATURE:

<221> NAME/KEY: modified_base

<222> LOCATION: (40)..(51)

<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 93

gacgaggctg actactattg tnnnnnnnnn nnnnnntctn nnnnnnnnnn ngcttcggc

60

ggtgtacca aacttac

77

<210> SEQ ID NO 94

<211> LENGTH: 74

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<220> FEATURE:

<221> NAME/KEY: modified_base

<222> LOCATION: (22)..(24)

<223> OTHER INFORMATION: a, c, t or g

<220> FEATURE:

<221> NAME/KEY: modified_base

<222> LOCATION: (31)..(36)

<223> OTHER INFORMATION: a, c, t or g

<220> FEATURE:

<221> NAME/KEY: modified_base

<222> LOCATION: (40)..(48)

<223> OTHER INFORMATION: a, c, t or g

<400> SEQUENCE: 94

-continued

gacgaggctg actactattg tnnnagctat nnnnnntctn nnnnnnnngt ctteggcggt 60

ggtaccaaac ttac 74

<210> SEQ ID NO 95
 <211> LENGTH: 627
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: A27: JH1
 Kappa light chain gene with stuffers
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (206)..(328)
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (357)..(377)
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (405)..(470)
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (501)..(596)

<400> SEQUENCE: 95

gaggaccatt gggccccctc cgagactctc gagcgcaacg caattaatgt gagttagctc 60

actcattagg caccacagcg tttacacttt atgcttccgg ctcgatatgt gtgtggaatt 120

gtgagcggt aacaatttca cacaggaaac agctatgacc atgattacgc caagctttgg 180

agcctttttt ttggagattt tcaac gtg aag aag ctc cta ttt gct atc ccg 232
 Met Lys Lys Leu Leu Phe Ala Ile Pro
 1 5

ctt gtc gtt ccg ttt tac agc cat agt gca caa tcc gtc ctt act caa 280
 Leu Val Val Pro Phe Tyr Ser His Ser Ala Gln Ser Val Leu Thr Gln
 10 15 20 25

tct cct ggc act ctt tcg cta agc ccg ggt gaa cgt gct acc tta agt 328
 Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser
 30 35 40

tagtaagctc ccaggcctct ttgatctg aaa cct ggt cag gcg ccg cgt 377
 Lys Pro Gly Gln Ala Pro Arg
 45

taatgaaagc gctaattggc aacagtg act ggg atc ccg gac cgt ttc tct ggc 431
 Thr Gly Ile Pro Asp Arg Phe Ser Gly
 50 55

tct ggt tca ggt act gac ttt acc ctt act att tct aga taatgagta 480
 Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg
 60 65 70

actagaccta cgtaacctag ttc ggt caa ggt acc aag gtt gaa atc aag cgt 533
 Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
 75 80

acg gtt gcc gct cct agt gtg ttt atc ttt cct cct tct gac gaa caa 581
 Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
 85 90 95

ttg aag tca ggt act acgcatctct aagcggccgc aacaggagga g 627
 Leu Lys Ser Gly Thr
 100

<210> SEQ ID NO 96
 <211> LENGTH: 102
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: A27: JH1

-continued

Kappa light chain gene with stuffers

<400> SEQUENCE: 96

Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
 1 5 10 15

His Ser Ala Gln Ser Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu
 20 25 30

Ser Pro Gly Glu Arg Ala Thr Leu Ser Lys Pro Gly Gln Ala Pro Arg
 35 40 45

Thr Gly Ile Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
 50 55 60

Thr Leu Thr Ile Ser Arg Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
 65 70 75 80

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
 85 90 95

Gln Leu Lys Ser Gly Thr
 100

<210> SEQ ID NO 97

<211> LENGTH: 413

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: 2a2: JH2
 Human lambda-chain gene with stuffers in place of CDRs

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (30)..(104)

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (117)..(122)

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (177)..(239)

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (270)..(413)

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (129)..(131)

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (135)..(152)

<400> SEQUENCE: 97

gaggaccatt gggccccccta ctccgtgac agt gca caa tcc gct ctc act cag 53
 Ser Ala Gln Ser Ala Leu Thr Gln
 1 5

cct gct agc gtt tcc ggg tca cct ggt caa agt atc act att tct tgt 101
 Pro Ala Ser Val Ser Gly Ser Pro Gly Gln Ser Ile Thr Ile Ser Cys
 10 15 20

aca tcttagtgac tc aga tct taatga ccg tag cac ccg ggc aag gcg 149
 Thr Arg Ser Pro His Pro Gly Lys Ala
 25 30

ccg taatgaatct cgtacgctgg tggt agc aat cgt ttc tcc gga tct aaa 200
 Pro Ser Asn Arg Phe Ser Gly Ser Lys
 35 40

tcc ggt aat acc gca agc tta act atc tct ggt ctg cag gttctgtagt 249
 Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu Gln
 45 50 55

tcgaattgct ttagtgaccc ggc ggt ggt acc aaa ctt act gtc ctc ggt caa 302
 Gly Gly Gly Thr Lys Leu Thr Val Leu Gly Gln

-continued

	60	65	
cct aag gct gct cct tcc gtt act ctc ttc cct cct agt tct gaa gag			350
Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro Pro Ser Ser Glu Glu			
	70	75	80
ctt caa gct aac aag gct act ctt gtt tgc ttg atc agt gac ttt tat			398
Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr			
	85	90	95
cct ggt gct gtt act			413
Pro Gly Ala Val Thr			
	100		

<210> SEQ ID NO 98
 <211> LENGTH: 103
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: 2a2: JH2
 Human lambda-chain gene with stuffers in place of CDRs

<400> SEQUENCE: 98

Ser	Ala	Gln	Ser	Ala	Leu	Thr	Gln	Pro	Ala	Ser	Val	Ser	Gly	Ser	Pro
1				5					10					15	
Gly	Gln	Ser	Ile	Thr	Ile	Ser	Cys	Thr	Arg	Ser	Pro	His	Pro	Gly	Lys
	20					25							30		
Ala	Pro	Ser	Asn	Arg	Phe	Ser	Gly	Ser	Lys	Ser	Gly	Asn	Thr	Ala	Ser
	35					40						45			
Leu	Thr	Ile	Ser	Gly	Leu	Gln	Gly	Gly	Gly	Thr	Lys	Leu	Thr	Val	Leu
	50				55					60					
Gly	Gln	Pro	Lys	Ala	Ala	Pro	Ser	Val	Thr	Leu	Phe	Pro	Pro	Ser	Ser
	65			70					75					80	
Glu	Glu	Leu	Gln	Ala	Asn	Lys	Ala	Thr	Leu	Val	Cys	Leu	Ile	Ser	Asp
		85						90					95		
Phe	Tyr	Pro	Gly	Ala	Val	Thr									
		100													

<210> SEQ ID NO 99
 <211> LENGTH: 10
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> SEQUENCE: 99

ctgtctgaac

10

I claim:

1. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a heavy chain CDR1 selected from the group consisting of:

(1) $\langle 1 \rangle_1 Y_2 \langle 1 \rangle_3 M_4 \langle 1 \rangle_5$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y;

(2) $(S/T)_1(S/G/X)_2(S/G/X)_3 Y_4 Y_5 W_6(S/G/X)_7$, wherein (S/T) is a 1:1 mixture of S and T residues, (S/G/X) is a mixture of 0.025 S, 0.025, G and 0.035 of each of amino acid residues A, D, E, F, H, I, K, L, M, N, P, Q, R, T, V, W, and Y;

(3) $V_1 S_2 G_3 G_4 S_5 I_6 S_7 \langle 1 \rangle_8 \langle 1 \rangle_9 \langle 1 \rangle_{10} Y_{11} Y_{12} W_{13} \langle 1 \rangle_{14}$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; and

(4) mixtures of vectors or genetic packages characterized by any of the above DNA sequences.

2. The focused library according to claim 1, wherein HC CDR1s (1), (2) and (3) are present in the library in the ratio 0.80:0.17:0.02.

3. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a heavy chain CDR2 selected from the group consisting of:

- (1) <2>I<2><3>SGG<1>T<1>YADSVKG, wherein <1> is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; <2> is an equimolar mixture of each of amino acid residues Y, R, W, V, G, and S; and <3> is an equimolar mixture of each of amino acid residues P, S, and G or an equimolar mixture of P and S;
- (2) <1>I<4><1><1><G><5><1><1><1>YADSVKG, wherein <1> is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; <4> is an equimolar mixture of residues D, I, N, S, W, Y; and <5> is an equimolar mixture of residues S, G, D and N;
- (3) <1>I<4><1><1>G<5><1><1>YNPSLKG, wherein <1> is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and <4> and <5> are as defined above;
- (4) <1>I<8>S<1><1><1>GGYY<1>YAASVKG, wherein <1> is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; <8> is 0.27 R and 0.027 of each of ADEFGHIKLMNPQSTVWY; and

(5) mixtures of vectors or genetic packages characterized by any of the above DNA sequences.

4. The focused library according to claim 3 wherein a mixture of HC CDR2s (1)/(2) (equimolar), (3) and (4) are present in the library in a ratio of 0.54:0.43:0.03.

5. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a heavy chain CDR3 selected from the group consisting of:

- (1) YYCA21111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;
- (2) YYCA211111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;
- (3) YYCA2111111YFDAYTG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;
- (4) YYCAR111S2S3111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E,

F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of S and G; and 3 is an equimolar mixture of Y and W;

(5) YYCA2111CSG11CY1YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

(6) YYCA211S1TIFG11111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

(7) YYCAR111YY2S3344111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; 2 is an equimolar mixture of D and S; and 3 is an equimolar mixture of S and G;

(8) YYCAR1111YC2231CY111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; 2 is an equimolar mixture of S and G; and 3 is an equimolar mixture of T, D and G; and

(9) mixtures of vectors or genetic packages characterized by any of the above DNA sequences.

6. The focused library according to claim 5, wherein 1 in one or all of HC CDR3s (1) through (8) is 0.095 of each of G and Y and 0.048 of each of A, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, and W.

7. The focused library according to claim 5 or 6, wherein HC CDR3s (1) through (8) are present in the library in the following proportions:

- (1) 0.10
- (2) 0.14
- (3) 0.25
- (4) 0.13
- (5) 0.13
- (6) 0.11
- (7) 0.04 and
- (8) 0.10

8. The focused library according to claim 5 or 6, wherein the HC CDR3s (1) through (8) are present in the library in the following proportions:

- (1) 0.02
- (2) 0.14
- (3) 0.25
- (4) 0.14
- (5) 0.14
- (6) 0.12
- (7) 0.08 and
- (8) 0.11

9. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the

antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encodes a kappa light chain CDR1 selected from the group consisting of:

- (1) RASQ<1>V<2><2><3>LA
- (2) RASQ<1>V<2><2><2><3>LA;

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <2> is 0.2 S and 0.044 of each of ADEFGHIKLMNPQRTVWY; and <3> is 0.2Y and 0.044 each of ADEFGHIKLMNPQRTVW and Y; and

- (3) mixtures of vectors or genetic packages characterized by any of the above DNA sequences.

10. The focused library of claim 9, wherein CDR1s (1) and (2) are present in the library in a ratio of 0.68:0.32.

11. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a kappa light chain CDR2 having the sequence:

<1>AS<2>R<4><1>,

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <2> is 0.2 S and 0.044 of each of ADEFGHIKLMNPQRTVWY; and <4> is 0.2 A and 0.044 each of DEFGHIKLMNPQRSTVWY.

12. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a kappa light chain CDR3 selected from the groups consisting of:

- (1) QQ<3><1><1><1>P<1>T,

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <3> is 0.2 Y and 0.044 each of ADEFGHIKLMNPQRTVW;

- (2) QQ33111P, wherein 1 and 3 are as defined in (1) above;

- (3) QQ3211PP1T, wherein 1 and 3 are as defined in (1) above and 2 is 0.2 S and 0.044 each of ADEFGHIKLMNPQRTVWY; and

- (4) mixtures of vectors or genetic packages characterized by any of the above DNA sequences.

13. The focused library according to claim 12, wherein CDR3s (1), (2) and (3) are present in the library in a ratio of 0.65:0.1:0.25.

14. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being

characterized by variegated DNA sequences that encode a lambda light chain CDR1 selected from the group consisting of:

- (1) TG<1>SS<2>VG<1><3><2><3>VS,

wherein <1> is 0.27 T, 0.27 G and 0.027 each of ADEFGHIKLMNPQRSVWY, <2> is 0.27 D, 0.27 N and 0.027 each of AEF GHIKLMNPQRSTVWY, and <3> is 0.36 Y and 0.036 each of ADEFGHIKLMNPQRSTVW;

- (2) G<2><4>L<4><4><4><3><4><4>, wherein <2> is as defined in (1) above and <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; and

- (3) mixtures of vectors or genetic packages characterized by any of the above DNA sequences.

15. The focused library according to claim 14, where CDR1s (1) and (2) are present in the library in a ratio of 0.67:0.33.

16. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a lambda light chain CDR2 has the sequence:

<4><4><4><2>RPS,

wherein <2> is 0.27 D, 0.27 N, and 0.027 each of AEF GHIKLMNPQRSTVWY and <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVW.

17. A focused library of vectors or genetic packages that display, display and express, or comprise a member of a diverse family of human antibody related peptides, polypeptides and proteins and collectively display, display and express, or comprise at least a portion of the diversity of the antibody family, the vectors or genetic packages being characterized by variegated DNA sequences that encode a lambda light chain CDR3 selected from the group consisting of:

- (1) <4><5><4><2><4>S<4><4><4>V,

wherein <2> is 0.27 D, 0.27 N, and 0.027 each of AEF GHIKLMNPQRSTVWY; <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVW; and <5> is 0.36 S and 0.0355 each of ADEFGHIKLMNPQRTVWY;

- (2) <5>SY<1><5>S<5><1><4>V, wherein <1> is an equimolar mixture of ADEFGHIKLMNPQRSTVWY; and <4> and <5> are as defined in (1) above; and

- (3) mixtures of vectors or genetic packages characterized by any of the above DNA sequences.

18. The focused library according to claim 17, wherein CDR3s (1) and (2) are present in the library in an equimolar mixture.

19. The focused library according to claim 1 or 2 further comprising variegated DNA sequences that encode a heavy chain CDR selected from the group consisting of:

- (1) one or more of the heavy chain CDR2s defined in claim 3 or 4;

(2) one or more of the heavy chain CDR3s defined in claims 5, 6, 7, or 8; and

(3) mixtures of vectors or genetic packages characterized by (1) and (2).

20. The focused library according to claim 3 further comprising variegated DNA sequences that encodes one or more heavy chain CDR3s selected from the group defined in claims 5, 6, 7 or 8.

21. The focused library according to claim 19 or 20, further comprising variegated DNA sequences that encodes a light chain CDR selected from the group consisting of

(1) one or more of the kappa light chain CDR1s defined in claim 9 or 10;

(2) the kappa light chain CDR2 defined in claim 11;

(3) one or more of the kappa light chain CDR3s defined in claim 12 or 13;

(4) one or more of the kappa light chain CDR1s defined in claim 14 or 15;

(5) the lambda light chain CDR2 defined in claim 16;

(6) one or more of the lambda light chain CDR3s defined in claim 17 or 18; and

(7) mixtures of vectors and genetic packages characterized by one or more of (1) through (6).

22. A population of variegated DNA sequences that encode a heavy chain CDR1 selected from the group consisting of:

(1) $\langle 1 \rangle_1 Y_2 \langle 1 \rangle_3 M_4 \langle 1 \rangle_5$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y;

(2) $(S/T)_1 (S/G/X)_2 (S/G/X)_3 Y_4 Y_5 W_6 (S/G/X)_7$ wherein (S/T) is a 1:1 mixture of S and T residues, $(S/G/X)$ is a mixture of 0.2025 S, 0.2025 G and 0.035 of each of amino acid residues A, D, E, F, H, I, K, L, M, N, P, Q, R, T, V, W, and Y;

(3) $V_1 S_2 G_3 G_4 S_5 I_6 S_7 \langle 1 \rangle_8 \langle 1 \rangle_9 \langle 1 \rangle_{10} Y_{11} Y_{12} W_{13} \langle 1 \rangle_{14}$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; and

(4) mixtures of variegated DNA sequences characterized by any of the above DNA sequences.

23. The population of variegated DNA sequences according to claim 22, wherein HC CDR1s (1), (2) and (3) are present in the population in the ratio 0.80:0.17:0.02.

24. A population of variegated DNA sequences that encode a heavy chain CDR2 selected from the group consisting of:

(1) $\langle 2 \rangle \langle 1 \rangle \langle 2 \rangle \langle 3 \rangle SGG \langle 1 \rangle T \langle 1 \rangle YADSVKG$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; $\langle 2 \rangle$ is an equimolar mixture of each of amino acid residues Y, R, W, V, G, and S; and $\langle 3 \rangle$ is an equimolar mixture of each of amino acid residues P, S, and G or an equimolar mixture of P and S;

(2) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle G \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle YADSVKG$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, and Y; $\langle 4 \rangle$ is an equimolar mixture of

residues D, I, N, S, W, Y; and $\langle 5 \rangle$ is an equimolar mixture of residues S, G, D and N;

(3) $\langle 1 \rangle I \langle 4 \rangle \langle 1 \rangle \langle 1 \rangle G \langle 5 \rangle \langle 1 \rangle \langle 1 \rangle YNPSLKG$, wherein $\langle 1 \rangle$ is an equimolar mixture of each of amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and $\langle 4 \rangle$ and $\langle 5 \rangle$ are as defined above;

(4) $\langle 1 \rangle I \langle 8 \rangle S \langle 1 \rangle \langle 1 \rangle \langle 1 \rangle GGY \langle 1 \rangle YAASVKG$, wherein $\langle 1 \rangle$ is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; $\langle 8 \rangle$ is 0.27 R and 0.027 of each of ADEFGHIKLMNPQSTVWY; and

(5) mixtures of variegated DNA sequences characterized by any of the above DNA sequences.

25. The population of variegated DNA sequences according to claim 24, wherein a mixture of HC CDR2s (1)/(2) (equimolar), (3) and (4) are present in the population in a ratio of 0.54:0.43:0.03.

26. A population of variegated DNA sequences that encode a heavy chain CDR3 selected from the group consisting of:

(1) YYCA21111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

(2) YYCA211111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

(3) YYCA2111111YFDAYTG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

(4) YYCAR111S2S3111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of S and G; and 3 is an equimolar mixture of Y and W;

(5) YYCA2111CSG11CY1YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

(6) YYCA211S1TIFG1111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; and 2 is an equimolar mixture of K and R;

(7) YYCAR111YY2S334411YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; 2 is an equimolar mixture of D and S; and 3 is an equimolar mixture of S and G;

(8) YYCAR1111YC2231CY111YFDYWG, wherein 1 is an equimolar mixture of each amino acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y; 2 is an equimolar mixture of S and G; and 3 is an equimolar mixture of T, D and G; and

(9) mixtures of variegated DNA sequences characterized by any of the above DNA sequences.

27. The population of variegated DNA according to claim 26, wherein 1 in one or all of HC CDR3s (1) through (8) is 0.095 of each of G and Y and 0.048 of each of A, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, and W.

28. The population of variegated DNA sequences according to claim 26 or 27, wherein HC CDR3s (1) through (8) are present in the population in the following proportions:

- (1) 0.10
- (2) 0.14
- (3) 0.25
- (4) 0.13
- (5) 0.13
- (6) 0.11
- (7) 0.04 and
- (8) 0.10

29. The population of variegated DNA sequences according to claim 26 or 27, wherein the HC CDR3s (1) through (8) are present in the population in the following proportions:

- (1) 0.02
- (2) 0.14
- (3) 0.25
- (4) 0.14
- (5) 0.14
- (6) 0.12
- (7) 0.08 and
- (8) 0.11

30. A population of variegated DNA sequences that encode a kappa light chain CDR1 selected from the group consisting of:

- (1) RASQ<1>V<2><2><3>LA
- (2) RASQ<1>V<2><2><2><3>LA;

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <2> is 0.2 S and 0.044 of each of ADEFGHIKLMNPQRTVWY; and <3> is 0.2Y and 0.044 each of ADEFGHIKLMNPQRTVW and Y; and

- (3) mixtures of variegated DNA sequences characterized by any of the above DNA sequences.

31. The population of variegated DNA sequences of claim 30, wherein CDR1s (1) and (2) are present in the population in a ratio of 0.68:0.32.

32. A population of variegated DNA sequences that encode a kappa light chain CDR2 having the sequence:

<1>AS<2>R<4><1>,

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <2> is 0.2 S and 0.044 of each of ADEFGHIKLMNPQRTVWY; and <4> is 0.2 A and 0.044 each of DEFGHIKLMNPQRTVWY.

33. A population of variegated DNA sequences that encode a kappa light chain CDR3 selected from the groups consisting of:

- (1) QQ<3><1><1><1>P<1>T,

wherein <1> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; <3> is 0.2 Y and 0.044 each of ADEFGHIKLMNPQRTVW;

- (2) QQ33111P, wherein 1 and 3 are as defined in (1) above;

- (3) QQ3211PPIT, wherein 1 and 3 are as defined in (1) above and 2 is 0.2 S and 0.044 each of ADEFGHIKLMNPQRTVWY; and

- (4) mixtures of variegated DNA sequences characterized by any of the above DNA sequences.

34. The population of variegated DNA sequences according to claim 33, wherein CDR3s (1), (2) and (3) are present in the population in a ratio of 0.65:0.1:0.25.

35. A population of variegated DNA sequences that encode a lambda light chain CDR1 selected from the group consisting of:

- (1) TG<1>SS<2>VG<1><3><2><3>VS,

wherein <1> is 0.27 T, 0.27 G and 0.027 each of ADEFGHIKLMNPQRSVWY; <2> is 0.27 D, 0.27 N and 0.027 each of AEFGLHIKLMNPQRSTVWY; and <3> is 0.36 Y and 0.036 each of ADEFGHIKLMNPQRSTVW;

- (2) G<2><4>L<4><4><4><3><4><4>,

wherein <2> is as defined in (1) above and <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVWY; and

- (3) mixtures of variegated DNA sequences characterized by any of the above DNA sequences.

36. The population of variegated DNA sequences according to claim 35, where CDR1s (1) and (2) are present in the population in a ratio of 0.67:0.33.

37. A population of variegated DNA sequences that encode a lambda light chain CDR2 has the sequence:

<4><4><4><2>RPS,

wherein <2> is 0.27 D, 0.27 N, and 0.027 each of AEFGLHIKLMNPQRSTVWY and <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVW.

38. A population of variegated DNA sequences that encode a lambda light chain CDR3 selected from the group consisting of:

- (1) <4><5><4><2><4>S<4><4><4>V,

wherein <2> is 0.27 D, 0.27 N, and 0.027 each of AEFGLHIKLMNPQRSTVWY; <4> is an equimolar mixture of amino acid residues ADEFGHIKLMNPQRSTVW; and <5> is 0.36 S and 0.0355 each of ADEFGHIKLMNPQRTVWY;

- (2) <5>SY<1><5>S<5><1><4>V, wherein <1> is an equimolar mixture of ADEFGHIKLMNPQRSTVWY; and <4> and <5> are as defined in (1) above; and

- (3) mixtures of variegated DNA sequence characterized by any of the above DNA sequences.

39. The population of variegated DNA sequences according to claim 38, wherein CDR3s (1) and (2) are present in the population in an equimolar mixture.

40. The population of variegated DNA sequences according to claim 22 or 23 further comprising variegated DNA sequences that encode a heavy chain CDR selected from the group consisting of:

- (1) one or more of the heavy chain CDR2s defined in claim 24 or 25;
- (2) one or more of the heavy chain CDR3s defined in claims **26, 27, 28** or **29**; and
- (3) mixtures of variegated DNA sequence characterized by (1) and (2).

41. The population of variegated DNA sequences according to claim 24 further comprising variegated DNA sequences that encodes one or more heavy chain CDR3s selected from the group defined in claims **26, 27, 28** or **29**.

42. The population of variegated DNA sequences according to claim 40 or 41 further comprising variegated DNA sequences that encodes a light chain CDR selected from the

group consisting of (1) one or more the kappa light chain CDR1s defined in claim 30 or 31;

- (2) the kappa light chain CDR2 defined in claim 32;
- (3) one or more of the kappa light chain CDR3s defined in claim 33 or 34;
- (4) one or more of the kappa light chain CDR1s defined in claim 35 or 36;
- (5) the lambda light chain CDR2 defined in claim 37;
- (6) one or more of the lambda light chain CDR3s defined in claim 38 or 39; and
- (7) mixtures of variegated DNA sequences characterized by one or more of (1) through (6).

43. A population of vectors comprising the variegated DNA sequences of any one of claims **22-42**.

* * * * *