



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION

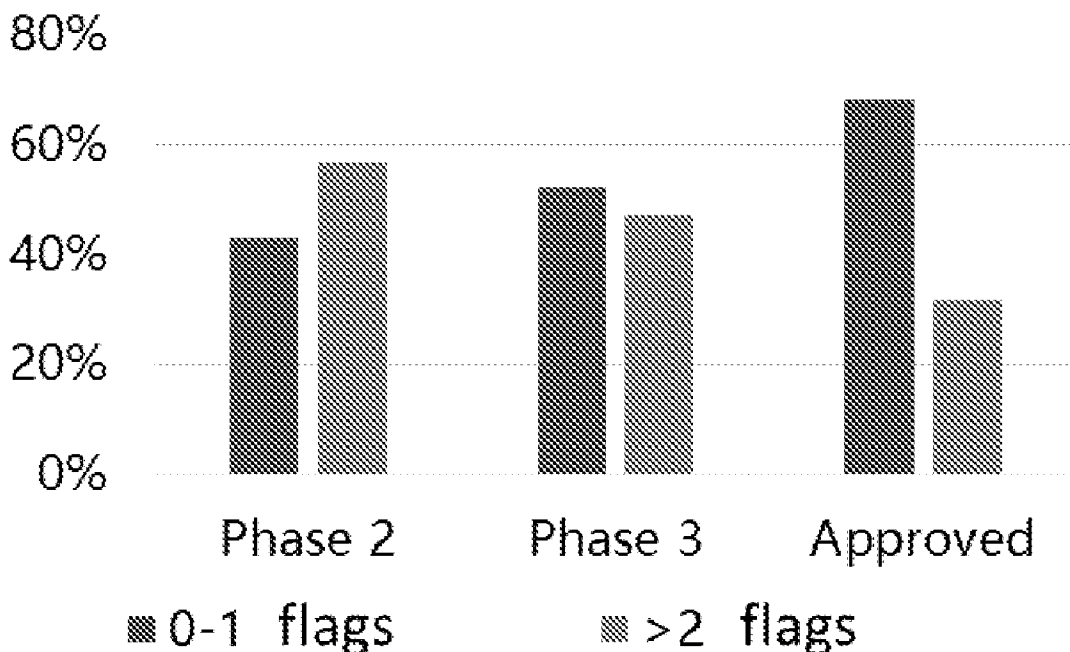
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2019/07/08
(87) Date publication PCT/PCT Publication Date: 2020/01/16
(85) Entrée phase nationale/National Entry: 2021/01/08
(86) N° demande PCT/PCT Application No.: US 2019/040843
(87) N° publication PCT/PCT Publication No.: 2020/014143
(30) Priorités/Priorities: 2018/07/08 (US62/695,065);
2019/03/22 (US62/822,671)

(51) Cl.Int./Int.Cl. *C07K 16/18* (2006.01),
C07K 16/46 (2006.01), *C12N 15/09* (2006.01),
C12P 21/08 (2006.01), *C40B 40/08* (2006.01),
C40B 40/10 (2006.01)
(71) Demandeur/Applicant:
SPECIFICA INC., US
(72) Inventeurs/Inventors:
BRADBURY, ANDREW RAYMON MORTON, US;
ERASMUS, MICHAEL FRANK, US;
TEIXEIRA, ANDRE, US
(74) Agent: BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.

(54) Titre : BANQUES D'ANTICORPS AYANT DES CARACTERISTIQUES MAXIMISEES D'APTITUDE AU
DEVELOPPEMENT D'ANTICORPS
(54) Title: ANTIBODY LIBRARIES WITH MAXIMIZED ANTIBODY DEVELOPABILITY CHARACTERISTICS

Figure 1



(57) **Abbrégé/Abstract:**

Antibody libraries comprising a plurality of heavy chain variable domains and/or a plurality of light chain variable domains, which comprise complementary determining regions (CDRs) found in naturally-occurring human antibodies, and methods of making

(57) **Abrégé(suite)/Abstract(continued):**

such antibody libraries. The antibody libraries are free of members that comprise one or more liabilities affecting one or more features of such members. Further, the antibody libraries comprise members having heavy chain and/or light chain CDRs not found in the same naturally-occurring human antibody.

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

16 January 2020 (16.01.2020)



(10) International Publication Number

WO 2020/014143 A1

(51) International Patent Classification:

C07K 16/18 (2006.01) C12P 21/08 (2006.01)

C07K 16/46 (2006.01) C40B 40/08 (2006.01)

C12N 15/09 (2006.01) C40B 40/10 (2006.01)

(21) International Application Number:

PCT/US2019/040843

(22) International Filing Date:

08 July 2019 (08.07.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/695,065 08 July 2018 (08.07.2018) US

62/822,671 22 March 2019 (22.03.2019) US

(71) Applicant: **SPECIFICA INC.** [US/US]; 512 Pacheco Street, Suite A203, Santa Fe, NM 87505 (US).(72) Inventors: **BRADBURY, Andrew Raymon, Morton**; c/o SPECIFICA INC., 512 Pacheco Street, Suite A203, Santa Fe, NM 87505 (US). **ERASMUS, Michael, Frank**; c/o SPECIFICA INC., 512 Pacheco Street, Suite A203, Santa Fe, NM 87505 (US). **TEIXEIRA, Andre**; c/o SPECIFICA INC., 512 Pacheco Street, Suite A203, Santa Fe, NM 87505 (US).(74) Agent: **CHEN, Yahua** et al.; POLSINELLI PC, One International Place, Suite 3900, Boston, MA 02210 (US).

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

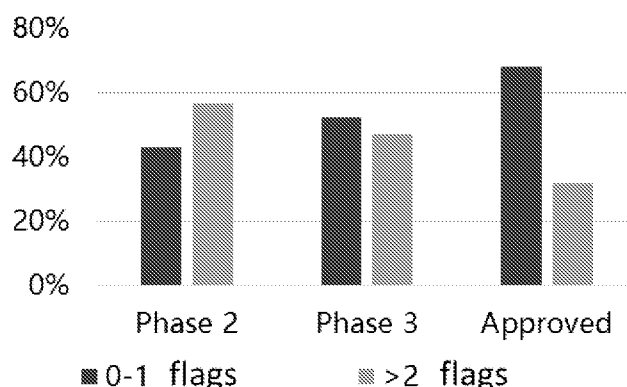
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,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

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: ANTIBODY LIBRARIES WITH MAXIMIZED ANTIBODY DEVELOPABILITY CHARACTERISTICS

Figure 1



(57) Abstract: Antibody libraries comprising a plurality of heavy chain variable domains and/or a plurality of light chain variable domains, which comprise complementary determining regions (CDRs) found in naturally-occurring human antibodies, and methods of making such antibody libraries. The antibody libraries are free of members that comprise one or more liabilities affecting one or more features of such members. Further, the antibody libraries comprise members having heavy chain and/or light chain CDRs not found in the same naturally-occurring human antibody.



WO 2020/014143 A1

ANTIBODY LIBRARIES WITH MAXIMIZED ANTIBODY DEVELOPABILITY CHARACTERISTICS

CROSS REFERENCE TO RELATED APPLICATIONS

5 This application claims the benefit of the filing dates of U.S. Provisional Application No. 62/695,065, filed on July 8, 2018, and U.S. Provisional Application No. 62/822,971, filed on March 22, 2019, the entire contents of each of which are incorporated by reference herein.

BACKGROUND OF THE INVENTION

10 There are currently three recombinant antibody platforms used to generate human antibodies for human therapeutic use: (1) the “humanization” of murine monoclonal antibodies; (2) the immunization of transgenic mice containing human antibody genes; and (3) *in vitro* selection from vast human antibody libraries. Immunization approaches depend upon the occurrence of appropriate *in vivo* immune responses, and may not yield antibodies
15 with desired characteristics. In contrast, *in vitro* selection has the advantage that antibodies with specific properties can be directly selected, and once selected, can be easily improved in terms of affinity or specificity.

In general, there are two types of antibody libraries: synthetic and natural antibody libraries. Synthetic antibody libraries can be constructed by introducing randomized
20 complementarity determining region (CDR) sequences into antibody frameworks. Such antibody libraries can have vast potential genetic diversity and improved expression *via* selection of well-behaved frameworks. However, synthetic antibody libraries also include many non-functional antibody members and exclude much natural diversity due to the formulaic manner used to generate diversity within a restricted set of framework scaffolds.
25 Antibody libraries created from natural sources, known as natural antibody libraries, have the advantage that the rearranged V genes undergo quality control in the B cell, and consequently a far higher proportion of the V genes are biologically functional, even if the potential diversity is lower. Disadvantages include the challenges of obtaining large numbers of B-cells to increase diversity, and the poor expression and biophysical properties of some
30 antibodies expressed recombinantly in *E. coli*, yeast or mammalian cells.

SUMMARY OF THE INVENTION

Provided herein are antibody libraries that comprise diversified heavy chain variable domains (VH) and/or light chain variable domains (VL), which comprise complementary determining regions (CDRs) obtained from naturally-occurring antibodies (*e.g.*, naturally-
5 occurring human antibodies or naturally-occurring camelid antibodies). Optionally, any of the VH CDRs and/or VL CDRs excludes at least members carrying one or more liabilities that affect one or more features of an antibody carrying such. Such antibody libraries, comprising CDRs from natural antibodies such as human antibodies, would have a high number of functional members and reflect natural diversity of human antibodies. Excluding
10 members carrying one or more liabilities as described herein would enhance the percentage of members having desired properties, for example, high yield when produced by recombinant technology, high stability, reduced aggregation capacity, reduced liabilities as described below *etc.* Thus, the antibody libraries described herein would maximize antibody developability characteristics.

15 Accordingly, one aspect of the present disclosure features an antibody heavy chain library, comprising a plurality of nucleic acids or a plurality of genetic packages comprising the nucleic acids. The plurality of nucleic acids encode a population of antibody heavy chain variable domains, which collectively (in combination) comprise a population of heavy chain CDR1s, a population of heavy chain CDR2s, and/or a population of heavy chain CDR3s
20 located at the CDR1 region, the CDR2 region, and the CDR3 region of a common antibody heavy chain variable domain gene. The heavy chain CDR1s, CDR2s, and/or CDR3s may be derived from naturally-occurring antibodies of a mammalian species, for example, human or camelid. In some embodiments, the plurality of nucleic acids encode a population of antibody heavy chain variable domains (*e.g.*, human antibody heavy chain variable domains),
25 which collectively (in combination) comprise a population of heavy chain CDR1s, a population of heavy chain CDR2s, and a population of heavy chain CDR3s.

In some embodiments, the common antibody heavy chain variable domain gene may be a human antibody heavy chain variable domain gene. Examples include VH1-24, VH2-70, VH3-7, VH4-30-4, VH5-51, VH1-18, VH1-69, VH3-23, VH5-10-1, VH3-9, or VH3-11.
30 In some instances, the human antibody heavy chain variable region gene may be derived from a therapeutic antibody, for example, abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, evoculumab, abituzumab, adalimumab, alemtuzumab,

alirocumab, bapineuzumab, benralizumab, brodalumab, canakinumab, certolizumab, clazakizumab, dacetuzumab, daclizumab, daratumumab, eculizumab, efalizumab, elotuzumab, epratuzumab, farletuzumab, fasinumab, ficlatuzumab, fletikumab, fresolimumab, fulranumab, gevokizumab, ibalizumab, lintuzumab, matuzumab, 5 mavrilimumab, mogamulizumab, motavizumab, natalizumab, nivolumab, obinutuzumab, ofatumumab, olokizumab, omalizumab, onartuzumab, otelixizumab, otlertuzumab, palivizumab, panitumumab, panobacumab, pertuzumab, pinatuzumab, polatuzumab, radretumab, ramucirumab, reslizumab, romosozumab, sarilumab, secukinumab, sifalimumab, tabalumab, tigatuzumab, til-drakizumab, tocilizumab, tovetumab, trastuzumab, vedolizumab, 10 veltuzumab, zalutumumab, or zanolimumab.

In some embodiments, the population of heavy chain CDR1s, the population of heavy chain CDR2s, and/or the population of heavy chain CDR3s can be free (*e.g.*, substantially free) of members comprising one or more of the following liabilities:

- (i) a glycosylation site (*e.g.*, comprising the motif NXS, NXT, or NXC, in which X 15 represents any naturally-occurring amino acid residue except for proline),
- (ii) a deamidation site (*e.g.*, comprising the motif of NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG),
- (iii) an isomerization site (*e.g.*, comprising the motif of DT, DH, DS, DG, or DD),
- (iv) an unpaired cysteine,
- 20 (v) net charge greater than 1 (*e.g.*, in LCDR1-2 and/or HCDR1-2),
- (vi) a tripeptide motif containing at least two aromatic residues (*e.g.*, HYF or HWH), which may affect viscosity;
- (vii) a motif that promotes aggregation (*e.g.*, comprising the motif of FHW);
- (viii) a polyspecificity site (*e.g.*, GG, GGG, RR, VG, W, WV, WW, WWW, YY, or 25 WXW, X referring to any amino acid residue),
- (ix) a protease sensitive site (*e.g.*, comprising the motif DX, in which X can be P, G, S, V, Y, F, Q, K, L, or D),
- (x) an integrin binding site (*e.g.*, comprising the motif RGD, LDV, or KGD),
- (xi) a lysine glycation site such as a lysine glycation site (*e.g.*, KE, EK, or ED),

(xii) a metal catalyzed fragmentation site (*e.g.*, comprises the motif of HS, SH, KT, HXS, or SXH, in which X represents any amino acid residue),

(xiii) a polyspecificity aggregation site (*e.g.*, the motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is F, I, L, V, W, or Y), and/or

5 (xiv) a streptavidin binding motif of (*e.g.*, comprises the motif HPQ, EPDW (SEQ ID NO: 117), PWXWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), or PWPWL (SEQ ID NO: 120)).

Alternatively, or in addition, the population of heavy chain CDR1s, the population of heavy chain CDR2s, and/or the population of heavy chain CDR3s in the antibody library
10 described herein is free of non-functional members.

In some embodiments, the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising at least two of (i)-(xiv). In some examples, at least the population of antibody heavy chain CDR1 fragments, the
15 population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising (i)-(ix), and optionally free of members comprising one or more of (x)-(xiv).

In some embodiments, the antibody library described herein comprise a population of heavy chain CDR1s and/or a population of heavy chain CDR2s that is free of members
20 comprising one or more of (i)-(xiv), and a population of heavy chain CDR3s is derived from naturally-occurring human antibodies (without removal of one or more the liability (i)-(xiv) disclosed herein), for example, derived from human B lymphocytes or precursor cells thereof.

In some instances, members of the antibody library described herein comprise heavy chain CDR1, heavy chain CDR2, and/or heavy chain CDR3 that are not from the same
25 naturally-occurring antibody. For examples, at least 50% of the members in the antibody library do not comprise heavy chain CDR1, heavy chain CDR2, and/or heavy chain CDR3 that are from the same naturally-occurring antibody.

In another aspect, the instant disclosure provides an antibody light chain library, comprising a plurality of nucleic acids or a plurality of genetic packages comprising the
30 nucleic acids. The plurality of nucleic acids encode a population of antibody light chain variable domains (*e.g.*, human antibody light chain variable domains), which collectively (in combination) comprise a population of light chain CDR1s, a population of light chain

CDR2s, and/or a population of light chain CDR3s located at the CDR1 region, the CDR2 region, and the CDR3 region of a common antibody light chain variable domain gene (*e.g.*, a human antibody light chain variable domain gene). The light chain CDR1s, CDR2s, and CDR3s, and optionally the common light chain variable domain gene may be derived from naturally-occurring antibodies of the same mammal species, for example, human. In some embodiments, the population of human antibody light chain variable domains collectively comprises a population of light chain CDR1s, a population of light chain CDR2s, and a population of light chain CDR3s. In some embodiments, the population of light chain CDR1s, the population of light chain CDR2s, and/or the population of light chain CDR3s is free of members comprising one or more of the liabilities described herein (*e.g.*, (i)-(xiv) described herein).

In some embodiments, the antibody light chain library described herein may comprise members having light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not from the same naturally-occurring antibody. For example, at least 50% of the members in the antibody light chain library do not comprise light chain CDR1, light chain CDR2, and/or light chain CDR3 that are from the same naturally-occurring antibody.

In some embodiments, the common antibody light chain variable domain gene used in the antibody light chain library may be a human antibody light chain variable domain gene. Examples include K1-12, K4-1, K2D-29, K3-11, K3-20, or L2-14. In some instances, the human antibody heavy chain variable region gene is derived from a therapeutic antibody such as those described herein.

Also provided herein is an antibody library that comprises (i) a first plurality of nucleic acids encoding the population of antibody heavy chain variable domains of the antibody heavy chain library described herein, and (ii) a second plurality of nucleic acids encoding the population of antibody light chain variable domains of the antibody light chain library described herein. Alternatively, the antibody library provided herein may comprise (i) a first plurality of nucleic acids encoding the population of antibody heavy chain variable domains of the antibody heavy chain library described herein, and (ii) a common light chain variable domain, which may be VK3-20.

Any of the antibody libraries disclosed herein may be of a suitable format, for example, a library of full-length antibodies, a library of antigen-binding fragments such as Fab fragments, a library of single-chain antibodies, or a library of single-domain antibodies

(*e.g.*, VHH antibodies). In some examples, the antibody library disclosed herein may be a human antibody library. In other examples, the antibody library disclosed herein may be a camelid VHH antibody library.

In another aspect, the present disclosure features a method for producing an antibody library, comprising:

providing (a) a first plurality of nucleic acids encoding a population of naturally-occurring antibody heavy chain complementary determining region 1 (CDR1) fragments, and/or (b) a second plurality of nucleic acids encoding a population of naturally-occurring antibody heavy chain complementary determining region 2 (CDR2) fragments; and

inserting the first plurality of nucleic acids and/or the second plurality of nucleic acids into the CDR1 region and/or the CDR2 region, respectively, of an antibody heavy chain variable domain gene (*e.g.*, those described herein), thereby producing an antibody library.

The method may further comprise:

providing a third plurality of nucleic acids encoding a population of naturally-occurring heavy chain complementary determining region 3 (CDR3) fragments, and

inserting the third plurality of nucleic acids into the CDR3 region of the heavy chain variable region gene.

The heavy chain CDR1 fragments, the heavy chain CDR2 fragments, and the heavy chain CDR3 fragments may be derived from naturally-occurring antibodies of a mammalian species as disclosed herein. In some embodiments, the heavy chain CDR1, CDR2, and CDR3 fragments can be derived from the same mammalian species. Optionally, the common heavy chain variable region gene may also be derived from the same mammalian species.

In some embodiments, the antibody library comprises members in each of which the heavy chain CDR1, the heavy chain CDR2, and/or the heavy chain CDR3 are not from the same naturally-occurring antibody. For example, at least 50% of the members in the antibody library do not contain heavy chain CDR1, the heavy chain CDR2, and/or the heavy chain CDR3 from the same naturally-occurring antibody.

In some embodiments, the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising one or more of the liabilities described herein, *e.g.*, (i)-(xiv) disclosed herein. For example, the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments,

and/or the population of antibody heavy chain CDR3 fragments may be free of members comprising at least two of (i)-(xiv). In some instances, the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising any of (i)-(ix), optionally further free of one or more of (x)-(xiv).

In other embodiments, the population of antibody heavy chain CDR1 fragments and/or the population of antibody heavy chain CDR2 fragments is free of members comprising one or more of the liabilities described herein (*e.g.*, at least two of (i) to (xiv) or all of (i) to (xiv)) and the population of human antibody heavy chain CDR3 fragments are from naturally-occurring human antibodies (*e.g.*, derived from B cells), which may not have the one or more liabilities excluded.

In some embodiments, the first plurality of nucleic acids, the second plurality of nucleic acids, and/or the third plurality of nucleic acids is produced by a process comprising:

(a) obtaining amino acid sequences of the heavy chain CDR1 regions, the heavy chain CDR2 regions, and/or the heavy chain CDR3 regions of a population of naturally-occurring antibodies (*e.g.*, naturally-occurring human antibodies);

(b) excluding from (a) the heavy chain CDR1 amino acid sequences, the heavy chain CDR2 amino acid sequences, and/or the heavy chain CDR3 amino acid sequences that comprise one or more of liabilities (i) to (xiv) to obtain liability-free heavy chain CDR1 sequences, heavy chain CDR2 sequences, and/or heavy chain CDR3 sequences; and

(c) synthesizing the first plurality of nucleic acids that encode the liability-free heavy chain CDR1 regions, the second plurality of nucleic acids that encode the liability-free heavy chain CDR2 regions, and/or the third plurality of nucleic acids that encode the liability-free heavy chain CDR3 regions.

In some instances, the process described above may further comprise (d) isolating functional members from the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions. For example, the functional members of the liability-free heavy chain CDR1, CDR2, and/or CDR3 can be isolated by expressing antibodies comprising the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions in host cells in a manner that the antibodies are displayed on surface of the host cells, isolating the antibodies that display on the host cells, show improved folding, and/or show reduced binding to polyspecificity reagents and

identifying the CDR1, CDR2, and/or CDR3 regions in the displayed antibodies, which are functional members of the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions.

The method for producing an antibody library as described herein may further comprise:

5 (i) providing a fourth plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 1 (CDR1) fragments, a fifth plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 2 (CDR2) fragments, and/or a sixth plurality of nucleic acids encoding a population of naturally-occurring human antibody
10 light chain complementary determining region 3 (CDR3) fragments, and

(ii) inserting the fourth plurality of nucleic acids, the fifth plurality of nucleic acids, and/or the sixth plurality of nucleic acids into the CDR1 region, the CDR2 region, and the CDR3 region, respectively, of an antibody light chain variable domain gene (*e.g.*, those described herein).

15 The light chain CDR1 fragments, the light chain CDR2 fragments, and the light chain CDR3 fragments may be derived from naturally-occurring antibodies of a mammalian species as disclosed herein. In some embodiments, the light chain CDR1, CDR2, and CDR3 fragments can be derived from the same mammalian species. Optionally, the common light chain variable region gene may also be derived from the same mammalian species.

20 The population of naturally-occurring antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments may be free of members comprising one or more of the liabilities described herein (*e.g.*, at least two or all of (i) to (xiv) described herein). The antibody library may comprise members, each of which comprises a light chain CDR1, a light chain
25 CDR2, and/or a light chain CDR3 that are not found in the same naturally-occurring antibody. For example, at least 50% of the members in the antibody library comprise light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not found in the same naturally-occurring antibody.

In some embodiments, the fourth plurality of nucleic acids, the fifth plurality of
30 nucleic acids, and/or the sixth plurality of nucleic acids is produced by a process comprising:

(a) obtaining amino acid sequences of the light chain CDR1, CDR2, and/or CDR3 regions of a population of naturally-occurring antibodies (*e.g.*, naturally-occurring human antibodies),

(b) excluding from (a) the light chain CDR1, CDR2, and/or CDR3 amino acid sequences that comprise one or more of (i) to (x) to obtain liability-free light chain CDR1, CDR2, and/or CDR3 sequences, and

(c) synthesizing the fourth plurality of nucleic acids, the fifth plurality of the nucleic acids, and/or the sixth plurality of nucleic acids that encode the liability-free light chain CDR1, CDR2, and/or CDR3 regions.

10 The above process may further comprise (d) isolating functional members from the liability-free light chain CDR1, CDR2, and/or CDR3 regions. For example, the functional members of the liability-free light chain CDR1, CDR2, and/or CDR3 are isolated by expressing antibodies comprising the liability-free light chain CDR1, CDR2, and/or CDR3 regions in host cells in a manner that the antibodies are displayed on surface of the host cells, isolating the antibodies that display on the host cells, show improved folding, and/or show reduced binding to polyspecificity reagents and identifying the CDR1, CDR2, and/or CDR3 regions in the displayed antibodies, which are functional members of the liability-free light chain CDR1, CDR2, and/or CDR3 regions.

Further, the present disclosure features a method for making an antibody light chain library, the method comprising:

(i) providing a first plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 1 (CDR1) fragments, a second plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 2 (CDR2) fragments, and/or a third plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 3 (CDR3) fragments, and

(ii) inserting the first plurality of nucleic acids, the second plurality of nucleic acids, and/or the third plurality of nucleic acids into the CDR1 region, the CDR2 region, and the CDR3 region, respectively, of an antibody light chain variable domain gene (*e.g.*, those described herein).

The light chain CDR1 fragments, the light chain CDR2 fragments, and the light chain CDR3 fragments may be derived from naturally-occurring antibodies of a mammalian

species as disclosed herein. In some embodiments, the light chain CDR1, CDR2, and CDR3 fragments can be derived from the same mammalian species. Optionally, the common light chain variable region gene may also be derived from the same mammalian species.

Also within the scope of the present disclosure are antibody libraries (*e.g.*, antibody heavy chain libraries, antibody light chain libraries, or a combination thereof) produced by any of the methods described herein.

The details of one or more embodiments of the invention are set forth in the description below. Other features or advantages of the present invention will be apparent from the following drawings and detailed description of several embodiments, and also from the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

The instant application contains at least one drawing executed in color.

Figure 1 is a chart showing the percentages of therapeutic antibodies in Phase 2 or Phase 3 clinical trials or approved therapeutic antibodies that are suitable or non-suitable for use as scaffold for antibody library construction. Suitable scaffolds for each of Phase 2, Phase 3, and Approved: 0-1 flags (left bar). Non-suitable scaffolds for each of Phase 2, Phase 3, and Approved: >2 flags (right bar).

Figure 2 is a schematic showing the seven scaffolds developed from the VH and/or VL gene of a selected therapeutic antibody. Except for LCDR2, all CDRs are as described by IMGT. LCDR2 uses the Kabat description.

Figure 3 is a schematic illustrating an exemplary design of cloning LCDR1 (as an example) into a selected light chain scaffold. Exemplary vector nucleic acids and corresponding amino acids are shown. The nucleic acid sequences correspond to SEQ ID NO: 1 (top) and SEQ ID NO: 2 (bottom). The amino acid sequence corresponds to SEQ ID NO: 3.

Figure 4 is a flow chart describing an exemplary process of identifying unique CDRs and informatically removing CDRs based on liabilities, occurrence rate, and/or anomalous length.

Figure 5 illustrates exemplary HCDR2 sequences comprising liabilities, such as deamidation sites, isomerization sites, glycosylation sites, or unpaired cysteine (highlighted). The first sequence of each column shows a graphical representation of a consensus sequence,

which is followed by exemplary variant sequences. The depicted sequences in the “Deamidation column” correspond to SEQ ID NOs: 4-24 (top to bottom). The depicted sequences in the “Isomerization column” correspond to SEQ ID NOs: 25-45 (top to bottom). The depicted sequences in the “Glycosylation column” correspond to SEQ ID NOs: 46-66 (top to bottom). The depicted sequences in the “Unpaired Cysteine column” correspond to SEQ ID NOs: 67-87 (top to bottom).

Figure 6 is a diagram showing presence of cysteine residues in heavy chain CDR3.

Figure 7 shows presence of cysteine residues in heavy chain CDR3. Sequences correspond to SEQ ID NOs: 88-102 (from top to bottom).

Figure 8 is a diagram illustrating paired cysteine residues in CDRs.

Figure 9 includes graphs showing the presence of liabilities in heavy chain CDR1 (left panel) and CDR2 (right panel) regions before and after bioinformatic filtration.

Figure 10 includes graphs showing the presence of liabilities in VK CDR1 (left panel) and CDR2 (right panel) regions before and after bioinformatic filtration.

Figure 11 includes graphs showing the presence of liabilities in V λ CDR1 (left panel) and CDR2 (right panel) before and after bioinformatic filtration.

Figure 12 includes graphs showing the presence of liabilities in VK CDR3 (left panel) and V λ CDR3 (right panel) before and after bioinformatic filtration.

Figure 13 includes a graph showing presence of liabilities in HC CDR3 before and after bioinformatic filtration.

Figure 14 includes charts showing the percentage of CDRs in each library that are excluded at different abundance threshold. Single thread of 4 reads is used across different libraries to exclude CDRs that arise by sequencing error.

Figure 15 includes a chart showing the number of CDRs remaining for all pooled CDRs except heavy chain CDR3 after exclusion at different threshold reads.

Figure 16 includes graphs showing the length distribution of CDRs in Library 1 (using scaffold derived from abrilumab). The germline length for each CDR is indicated under each graph and the CDRs outside of the highlighted region are eliminated for anomalous length.

Figure 17 includes graphs showing the length distribution of CDRs in Library 2 (using scaffold derived from mepolizumab). The germline length for each CDR is indicated

under each graph and the CDRs outside of the highlighted region are eliminated for anomalous length.

Figure 18 includes graphs showing the length distribution of CDRs in Library 3 (using scaffold derived from crenezumab). The germline length for each CDR is indicated under each graph and the CDRs outside of the highlighted region are eliminated for anomalous length.

Figure 19 includes graphs showing the length distribution of CDRs in Library 4 (using scaffold derived from necitumumab). The germline length for each CDR is indicated under each graph and the CDRs outside of the highlighted region are eliminated for anomalous length.

Figure 20 includes graphs showing the length distribution of CDRs in Library 5 (using scaffold derived from anifrolumab). The germline length for each CDR is indicated under each graph and the CDRs outside of the highlighted region are eliminated for anomalous length.

Figure 21 includes graphs showing the length distribution of CDRs in Library 6 (using scaffold derived from evoculumab). The germline length for each CDR is indicated under each graph and the CDRs outside of the highlighted region are eliminated for anomalous length.

Figures 22A-22F includes graphs showing natural distribution (blue) and replicated natural designed distributions before (red) and after (green) filtering for well folded sequences. Data is shown for 6 different scaffolds/germlines at each of LCDR1-3 and HCDR1-2. The results are based on sequencing the libraries illustrated in figure 25B-G.

Figures 22G-22L includes cumulative plots for the diversity at each CDR position assessed for a natural distribution (natural library – blue), and replicated natural designed distributions before (red) and after (green) filtering for well folded sequences. The results are based on sequencing the libraries illustrated in Figure 25B-25G.

Figure 23 is a diagram illustrating an exemplary process for selecting for functional CDRs, for example, well folded LC CDR3s.

Figure 24 is a diagram showing an exemplary process of inserting oligonucleotides encoding LC CDR1s into a yeast display vector. Nucleic acid sequences representing part of a yeast display vector to be cleaved with BsaI and SfiI correspond to SEQ ID NOs: 103-104 (from top to bottom). Sequences of the cleaved vector (middle) correspond to SEQ ID NOs:

105-108 (labeled from top to bottom then left to right). Sequences of the amplified LCDR1 oligo (middle) correspond to SEQ ID NOs: 109-112 (labeled top to bottom then left to right). Homologous recombination produces the amplified LCDR1 inserted into the yeast display vector (bottom), which corresponds to SEQ ID NOs: 113-116 (labeled top to bottom then left to right).

Figure 25A is a diagram illustrating an exemplary process of isolating well expressed antibody CDRs by sorting yeast displaying single CDR loop libraries after cloning, using VK chains comprising functional KCDR1 as an example.

Figures 25B-25G illustrate the results of applying the exemplary process of isolating displayed single CDR loop libraries for each of the CDR (HCDR1-1, LCDR1-3) libraries displayed in yeast. The X axis indicates the level of antibody display, while the Y axis indicates the number of clones at each particular display level. For all histograms the display level of the clinical candidate is shown as a dotted blue line. For LCDR1-2 and HCDR1-2, fluorescence activated cell sorting was used to sort the most fluorescent 2% of yeast, corresponding to yeast displaying the most highly expressed antibodies. The display levels for the SV5 sorted (blue plot) and the most fluorescent 2% (red plot) are shown. For LCDR3, magnetic activated cell sorting was used to sort yeast displaying the most highly expressed antibodies. The transformed yeast clones (blue plot) are compared to the magnetic activated cell sorted yeast clones (red).

Figure 26 is a diagram illustrating the pairing between a constant VH chain and VL chains with diversity in LC CDRs to select functional antibodies.

Figure 27 is a diagram illustrating pairing between a constant VL chain and VH chains with diversity in VH CDRs for selection of functional antibodies.

Figure 28 is a diagram illustrating construction of an exemplary antibody library comprising VL chains having well-folded LC CDRs selected via yeast display and unmodified VH domains.

Figure 29 is a diagram illustrating an exemplary process of amplifying and assembling VH and VL CDRs into a pre-selected VH or VL scaffold.

Figure 30 is a diagram illustrating construction of an exemplary Fab antibody library via combining a library of VH domains and a library of VL domains.

Figure 31 is a diagram illustrating construction of an exemplary scFv antibody library via combining a library of VH domains and a library of VL domains.

Figure 32 is a flow plot of scFvs corresponding to the clinical candidates used as scaffolds compared to the corresponding libraries. The double peak for library 5 is thought to be due to the presence of a smaller truncated product, which was eliminated in the final phage display libraries.

5 **Figure 33** shows a western blot of libraries 1-3 with the pIII and the scFv-pIII bands indicated.

Figure 34 shows binding of antibodies displayed on yeast binding to targets against which they were selected at different concentrations. Display on yeast followed two rounds of selection by phage display, and two rounds of yeast display. Library 1 and Library 2
10 represent two different libraries created using the same LCDR1-3 and HCDR1-2 diversity, and HCDR3 diversity from different donors.

Figure 35 is a diagram illustrating an exemplary affinity maturation approach via VL shuffling.

15 **Figure 36** is a diagram illustrating further HC CDR1 and HC CDR2 shuffling, following the VL shuffling depicted in Figure 32.

Figure 37 includes diagrams showing clones binding to the indicated antigens at various concentrations (100 nM, 10 nM, 1nM, or 0nM) isolated from Library A or Library B.

Figures 38A-38C include diagrams showing isolation of high affinity antibodies (having binding affinity at the subnanomolar level) from the libraries disclosed herein.

20 **Figure 39** is a diagram showing isolation of antibodies with high binding affinity to additional antigens, PDGF, TGFBR2, and TGFBR3, from the libraries disclosed herein.

DETAILED DESCRIPTION OF THE INVENTION

It is generally accepted that the larger or more diverse an antibody library, measured
25 in terms of the number of different antibodies, the better the antibodies that can be selected from it. Griffiths *et al.*, *EMBOJ* 13(14):3245-3260, 1994 and Perelson *et al.*, *J Theor Biol.*, 81(4):645-70, 1979. The diversity of most antibody libraries has been estimated by counting the number of transformants, assuming that each colony represents a different antibody.

It was generally assumed that the VH gene diversity was the same as the number of
30 colonies obtained, *e.g.*, $\sim 10^8$ for the library described in Sblattero *et al.*, *Nat Biotechnol.*, 18(1):75-80, 2000. However, next-generation sequencing (NGS) showed that the VH clonal diversity (unique HCDR3 amino acid sequences) was actually ~ 30 fold lower (3×10^6).

D'Angelo *et al.*, *MAbs.*, 6(1):160-72, 2014. Notwithstanding this apparent low diversity, many antibodies have been selected from this library. See, *e.g.*, Sblattero *et al.*, *Nat Biotechnol.*, 18(1):75-80, 2000; Glanville *et al.*, *Curr Opin Struct Biol.*, 33:146-60, 2015; Lou *et al.*, *Journal of immunological methods*; 253(1-2):233-42, 2001; Kehoe *et al.*, *Mol Cell*
 5 *Proteomics*, 5(12):2350-63, 2006; Ayriss *et al.*, *J Proteome Res.* 6(3):1072-82, 2007; Velappan *et al.*, *Journal of immunological methods*, 321(1-2):60-9, 2007; Lillo *et al.*, *PLoS One*, 6(12):e27756, 2011; Ferrara *et al.*, *PLoS One*, 7(11):e49535, 2012; Close *et al.*, *BMC Microbiol.* 13:270, 2013; and Ferrara *et al.*, *MAbs*, 7(1):32-41, 2015.

NGS sequencing of another natural antibody library showed an even lower measured
 10 VH diversity (2×10^5), even though the number of donors used (654) was extremely high, and the estimated number of colonies was 3×10^{10} . Glanville *et al.*, *Proceedings of the National Academy of Sciences of the United States of America*, 106(48):20216-21, 2009. Further, Fantini *et al.* *PLoS One*. 12(5):e0177574, 2017 described three libraries with maximal diversities (numbers of colonies) $6\text{--}16 \times 10^6$, and estimated NGS diversities of $3\text{--}9 \times 10^6$.

15 While genetic diversity is essential, effective functional diversity is even more important: a high genetic diversity is of no utility if the encoded antibodies are non-functional and unable to fold properly. Indeed, a single amino acid change in an antibody can result in dramatic changes in expression levels and stability. Some publications have shown the superiority of natural antibody libraries over synthetic ones. Hugo *et al.*, *Protein Eng.*,
 20 16(5):381-6, 2003; Wang *et al.*, *Proteins*, 82(10):2620-30, 2014; and Chan *et al.*, *Journal of immunological methods*, 373(1-2):79-88, 2011. Natural diversity has the advantage that it has been prescreened for functionality by the immune system. However, it has the disadvantage that some antibodies are poorly expressed and folded in *in vitro* display systems, and that diversity can be dominated by a small number of clones.

25 The present disclosure aims, at least in part, at constructing antibody libraries comprising natural diversity such that the members of the libraries would be prescreened by the immune system for functionality, while excluding members that contain potential liabilities, would be poorly expressed, aggregating and/or poorly folded in a common screening system (*e.g.*, yeast display, phage display, or a folding reporter such as β -
 30 lactamase; see, *e.g.*, Saunders *et al.*, *Nat. Che Biol.*, 12:94-101; 1988; and D'Angelo *et al.*, *BMC genomics* 12, suppl. 1, S1-S5; 2011; or green fluorescent protein; see *e.g.* Waldo, *et al.*, *Nat. Biotechnol.*, 17: 691-5; 1999; Cabantous, *et al.*, *PLoS ONE.*, 3:e2387; 2008; and

Cabantous, *et al.*, *J Struct Funct Genomics*, 6:113-9; 2005). The present disclosure thus features, in some embodiments, a method to create extremely diverse, highly functional antibody libraries by combining naturally occurring CDRs, including naturally occurring CDRs containing somatic mutations generated *in vivo*, within antibody scaffolds such that members of the antibody libraries are expected to be well expressed and/or folded, and lacking liabilities.

As used herein, the term “liability” refers to a motif in an antibody (*e.g.*, located in a heavy chain or light chain CDR region) that would negatively affect one or more desired features of the antibody (*e.g.*, stability, good expression in an expression or display system, proper folding, no or reduced aggregation, solubility, no or reduced integrin binding, no or reduced glycosylation, no or reduced deamidation, no or reduced isomerization, no unpaired cysteine, or no or reduced protease sensitivity, *etc.*). By virtue of being comprised of highly functional members, such an antibody library would be expected to be functionally much larger than libraries of similar genetic size, in which antibodies are present that contain any of these liabilities. In other words, the antibody libraries disclosed herein would have a much larger effective diversity.

I. Antibody Libraries and Methods of Construction

Provided herein are antibody libraries comprising the heavy chain and/or light chain CDR populations as described herein, wherein the heavy chain CDRs and/or light chain CDRs are inserted into a pre-selected heavy chain variable domain gene and/or a pre-selected light chain variable domain gene as also described herein, as well as methods of producing such antibody libraries. The heavy chain CDR1s, CDR2s, and/or CDR3s, and the pre-selected heavy chain variable domain may be of a mammalian species, for example, human, mouse, rat, rabbit, dog, pig, or camelid such as camel or llama. In some instances, the heavy chain CDR1s, CDR2s, and CDR3s may be derived from antibodies of the same mammalian species (*e.g.*, human or camelid). Optionally, the pre-selected heavy chain variable domain gene may be from the same mammalian species. Alternatively, the heavy chain CDR1s, CDR2s, and/or CDR3s, and optionally the pre-selected heavy chain variable domain gene may be derived from naturally-occurring antibodies of different mammalian species.

Similarly, the light chain CDR1s, CDR2s, and CDR3s, as well as the pre-selected light chain variable domain gene may be of a mammalian species such as those described

herein. In some instances, the light chain CDR1s, CDR2s, and CDR3s may be derived from antibodies of the same mammalian species (*e.g.*, human or camelid). Optionally, the pre-selected light chain variable domain gene may be from the same mammalian species.

Alternatively, the light chain CDR1s, CDR2s, and/or CDR3s, and optionally the pre-selected light chain variable domain gene may be derived from naturally-occurring antibodies of different mammalian species.

In some embodiments, the heavy chain CDRs and the pre-selected variable domain gene, and the light chain CDRs and the pre-selected variable domain gene are all of the same mammal species, for example, human.

An antibody (interchangeably used in plural form) is an immunoglobulin molecule capable of specific binding to a target, such as a carbohydrate, polynucleotide, lipid, polypeptide, *etc.*, through at least one antigen recognition site, located in the variable region of the immunoglobulin molecule. As used herein, the term “antibody encompasses not only intact (*e.g.*, full-length) antibodies, but also antigen-binding fragments thereof (such as Fab, Fab', F(ab')₂, Fv), single-chain antibody (scFv), fusion proteins comprising an antibody portion, diabodies, nanobodies, single domain antibodies (also known as nanobodies, *e.g.*, a V_H only antibody such as the V_HH antibodies found in camelids), or multispecific antibodies (*e.g.*, bispecific antibodies).

A typical antibody molecule comprises a heavy chain variable region (V_H) and a light chain variable region (V_L), which are usually involved in antigen binding. The V_H and V_L regions can be further subdivided into regions of hypervariability, also known as “complementarity determining regions” (“CDR”), interspersed with regions that are more conserved, which are known as “framework regions” (“FR”). Each V_H and V_L is typically composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-

terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. The extent of the framework region and CDRs can be precisely identified using methodology known in the art, for example, by the Kabat definition, the Chothia definition, the AbM definition, and/or the contact definition, all of which are well known in the art. See, *e.g.*, Kabat, E.A., *et al.* (1991) *Sequences of Proteins of Immunological Interest, Fifth Edition*, U.S.

Department of Health and Human Services, NIH Publication No. 91-3242, Chothia *et al.*, (1989) *Nature* 342:877; Chothia, C. *et al.* (1987) *J. Mol. Biol.* 196:901-917, Al-lazikani *et*

al (1997) *J. Molec. Biol.* 273:927-948; and Almagro, *J. Mol. Recognit.* 17:132-143 (2004). See also hgmp.mrc.ac.uk and bioinf.org.uk/abs).

Single-domain antibodies, also known as nanobodies, are also within the scope of the present disclosure. In some embodiments, a single-domain antibody contains only a heavy chain (VHH). Heavy chain only antibodies (HcAb) are naturally produced by camelids and sharks. The antigen binding portion of the HcAb is comprised of the VHH fragment. Vincke *et al.*, *Methods Mol Biol.* 911:15-26 (2012).

The antibody libraries disclosed herein may contain a population of antibodies of any suitable format. In some embodiments, the antibody library disclosed herein comprise a population of full-length antibodies, which may be of any suitable family (*e.g.*, IgG, or IgA). In other embodiments, the antibody library disclosed herein comprise a population of antigen-binding fragments, for example Fab fragments. In yet other embodiments, the antibody library disclosed herein comprise a population of single-chain antibodies. Alternatively, the antibody library disclosed herein may comprise a population of single-domain antibodies such as VHH fragments.

Exemplary steps for constructing the antibody libraries described herein may include:

- (1) Identifying suitable VH/VL pairs for use as variable gene scaffolds;
- (2) Generating vectors based on the scaffolds containing a single insertion site for each CDR, the remaining CDRs may remain unchanged;
- (3) Identifying naturally occurring CDRs by analysis of a database of naturally occurring antibody sequences (which may be obtained from sequencing members of a natural antibody library);
- (4) Eliminating from the database of naturally occurring CDR sequences those sequences likely to encode liabilities;
- (5) Synthesize the remaining set of CDRs as oligonucleotides;
- (6) Inserting the CDRs at their appropriate sites within the previously modified scaffolds, each scaffold containing CDRs at only one site (*e.g.*, the identified collection of HCDR1s is inserted at the HCDR1 site of the modified scaffold).

In some embodiments, the CDRs (*e.g.*, CDR1, CDR2, CDR3, or a combination thereof) identified as described herein may be experimentally screened or selected for good folding and/or expression and screened or selected against liabilities such as poor folding, poor expression, polyreactivity or aggregation. The selected CDRs may be inserted into

complete V domains within the context of the scaffolds. The resultant complete V domains could be further screened and selected for good folding and/or expression, and/or screened and selected against liabilities such as poor folding or expression, polyreactivity or aggregation. The selected VH/VL complete scaffold pairs can be assembled and cloned into an appropriate display vector (e.g., phage or yeast) for screening of antibodies having desired binding specificity.

A. Selection of Heavy Chain and Light Chain Variable Domain Framework Scaffolds

In some embodiments, the heavy chain and/or light chain framework scaffolds used in constructing the antibody libraries described here may be derived from commercially available therapeutic antibodies (e.g., those whose marketing authorization has been approved by the US Food and Drug Administration or/or the European Medicines Agency) or therapeutic antibodies that are currently in clinical trials, for example, in phase II or phase III trials. As used herein, a therapeutic antibody refers to the antibody molecule of an approved drug product (e.g., in the US, in EP, or in other jurisdictions such as CA or JP), or an antibody molecule that has been or is currently in a clinical trial in a suitable jurisdiction, for example, in the US or in Europe.

The germline heavy chain variable domain and light chain variable domain genes used in such therapeutic antibodies can be examined for features such as aggregation, hydrophobic interaction, polyspecificity, monomericity, level of expression in mammalian host cells (e.g., in HEK cells or CHO cells), T_m of its Fab form, and purification characteristics. See Table 1. Those having desired features, for example, high expression levels in mammalian cells (e.g., ≥ 50 mg/L in HEK cells), high Fab T_m (e.g., > 64 °C), low slope for accelerated stability (e.g., < 0.09), etc. can be selected as framework scaffolds for library constructions. Additional features and selection criterion are provided in Table 1, which shows as exemplary examples those therapeutic antibodies with the best properties, as well as three additional antibodies with poor therapeutic properties. This data is derived from Jain, T. *et al.* Biophysical properties of the clinical-stage antibody landscape. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 944-949, doi:10.1073/pnas.1616408114 (2017).

For each characteristic being evaluated, the worst 10% of the evaluated therapeutic antibodies can be assigned with a flag. See Figure 1. In some instances, those therapeutic antibodies having less than 2 flags (*e.g.*, having 1 flag or none) may be selected for use as the heavy chain and/or light chain framework scaffold.

5 In some embodiments, the heavy chain variable domain gene for use as the heavy chain framework scaffold can be VH1-24, VH2-70, VH3-7, VH4-30-4, VH5-51, VH1-18, VH1-69, VH3-23, VH5-10-1, VH3-9, or VH3-11. Alternatively, or in addition, the light chain variable domain gene for use as the light chain framework scaffold can be K1-12, K4-1, K2D-29, K3-11, K3-20, or L2-14. Such heavy chain and/or light chain framework
10 scaffolds may be germline VH and/or VL genes. Alternatively, the heavy chain and/or light chain framework scaffolds may contain one or more mutations in one or more framework regions (*e.g.*, FR1, FR2, FR3, or FR4) as compared with the germline gene counterpart. Such mutations may be present within the therapeutic antibody, or may be introduced to avoid specific liabilities, *e.g.*, methionine oxidation, aggregation, integrin binding, glycosylation,
15 deamidation, isomerization, unpaired cysteine, or protease sensitivity. In specific examples, the antibody library described herein uses the following VH and VL framework scaffold pairs: VH1-24/VK1-12, VH2-70/VK4-1, VH3-7/VK2D-29, VH4-30-4/VK3-11, VH5-51/VK3-20, or VH1-18/VL2-14.

In some specific examples, the VH and/or VL framework scaffolds used in the
20 antibody library described herein are derived from abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, or evoculumab, the characteristics of each of which are provided in Table 2. As used herein, “derived from” refers to the use of the VH and/or VL genes of any of these therapeutic antibodies, either with no modification, or with one or more mutations introduced into one or more of the framework regions, for example, up to 5 amino
25 acid substitutions (*e.g.*, up to 4, 3, 2, or 1 amino acid substitutions) in the VH gene (*e.g.*, in one or more of the framework regions) and/or in the VL gene (*e.g.*, in one or more of the framework regions).

In some instances, the mutations introduced into a germline VH and/or VL gene or introduced into the VH and/or VL gene of a reference therapeutic antibody (*e.g.*, those listed
30 in Table 2) may be conservative substitutions. As used herein, a “conservative amino acid substitution” refers to an amino acid substitution that does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be

prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Conservative substitutions of amino acids include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

The amino acid sequences of exemplary VH and VL framework scaffolds are provided below (with CDRs that are modified indicated in boldface and mutations relative to parent therapeutic antibodies listed in Table 2 underlined):

Scaffold derived from abrilumab:

- CDRs are bold and underlined
- Mutations from the original antibody are italicized (all Jk have been replaced for Jk4).

VL:

DIQMTQSPSSVSASVGDRTITCRAS**QGISSW**LAWYQQKPGKAPKLLIY**GASNLES**GVPSRFSGSGSGTDFTLTI
SSLQPEDFANYCY**QQANSFPWT**FGGGTKVEIK (SEQ ID NO: 121)

Linker:

SGGSTITSYNVYYTKLSSSGT (SEQ ID NO: 122)

VH:

QVQLVQSGAEVKKPGASVKVSCVSG**GYTSLDLS**IHWRQAPGKGLEWMGG**FDPQDGETI**YAKFKQGRVTMTEDTS
TDTAYMELSSSLKSEDTAVYYC**ATGSSSSWFDP**WGQGLTVTVSS (SEQ ID NO: 123)

Scaffold derived from mepolizumab:

VL:

DIVMTQSPDSLAVSLGERATINCKSS**QSLNLSGNQKNY**LAWYQQKPGQPPKLLIY**GASTRES**GVDPDRFSGSGSGT
DFTLTISLQAEDVAVYYC**QNVHSFPFT**FGGGTKVEIK (SEQ ID NO: 124)

Linker:

SGGSTITSYNVYYTKLSSSGT (SEQ ID NO: 125)

VH:

QVTLRESGPALVKPTQTLTLTCTVSG**GFSL**SAYSVNWIRQPPGKALEWLAM**IWGDGK**IVYNSALKSRLTISKDTSK
NQVVLTMNTMDPVDATYYC**AGDGYYPYAMDN**WGQGLTVTVSS (SEQ ID NO: 126)

Scaffold derived from crenezumab:

VL:

DIVMTQSPLSLPVTGPGEPAISCRSS**QSLVYSNGDTY**LHWYLQKPGQSPQLLIY**KVSNRFS**GVDPDRFSGSGSGTD
FTLKISRVEAEDVGVYYC**SQSTHVPWT**FGGGTKVEIK (SEQ ID NO: 127)

Linker:

SGGSTITSYNVYYTKLSSSGT (SEQ ID NO: 128)

5 VH:

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYGMSWVRQAPGKGLVVASINSNGGSTYYPDSVKGRFTISRDN
KNSLYLQMNSLRAEDTAVYYCASGDYWGQGTTVTVSS (SEQ ID NO: 129)

Scaffold derived from necitumumab:

10

VL:

EIVMTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLT
SSLEPEDFAVYYCHQYGSTPLTFGGGTKVEIK (SEQ ID NO: 130)

15 Linker:

SGGSTITSYNVYYTKLSSSGT (SEQ ID NO: 131)

VH:

20 QVQLQESGPGLVKPSQTLSTCTVSGGSISSGDYYWSWIRQPPGKGLEWIGYIYYSGSTDYNPSLKSRVTMSVDT
SKNQFSLKVNSTAAADTAVYYCARVSIFGVGTFDYWGQGTTLTVTVSS (SEQ ID NO: 132)

Scaffold derived from anifrolumab:

VL:

25 EIVLTQSPGTLSLSPGERATLSCRASQSVSSSFLAWYQQKPGQAPRLLIYCASSRATGIPDRLSGSGSGTDFTLT
ITRLEPEDFAVYYCQQYDSSAITFGGGTKVEIK (SEQ ID NO: 133)

Linker:

SGGSTITSYNVYYTKLSSSGT (SEQ ID NO: 134)

30

VH:

EVQLVQSGAEVKKPGESLKISCKGSGYIFTNYWIAWVRQVPGKGLSMGIIYPGDSDIYRSPSFQGQVTISADKS
ITTAYLQWSSLKASDTAIFYCARHDIEGFDYWGRTTLTVTVSS (SEQ ID NO: 135)

35 Scaffold derived from evolocumab:

VL:

ESALTQPASVSGSPGQSITISCTGTSSDVGGYNSVSQYQQHPGKAPKLMIIYEVSNRPSGVSNRFGSGSGNTASL
40 TISGLQAEDEADYYCNSYTSTSMVFGGGTKLTVL (SEQ ID NO: 136)

Linker:

SGGSTITSYNVYYTKLSSSGT (SEQ ID NO: 137)

VH:

45 EVQLVQSGAEVKKPGASVKVSCASGYTLTSYGISWVRQAPGQGLEWMGWVSFYNGNTNYAQKLQGRGTMTTDPS
TSTAYMELRSLRSDDTAVYYCARGYGMDVWGQGTTLTVTVSS (SEQ ID NO: 138)

Any of the VH and/or VL framework scaffolds described herein can be used to construct a cassette that allows for cloning of one or more of pluralities of nucleic acids each

encoding a diverse population of a heavy chain CDR or a light chain CDR into the corresponding framework scaffold at the corresponding CDR position.

In some instances, restriction sites can be introduced into a heavy chain scaffold flanking the CDR1 region, the CDR2 region, or the CDR3 region for cloning a plurality of nucleic acids encoding a diverse population of heavy chain CDR1s, heavy chain CDR2s, or heavy chain CDR3s, respectively. In some instances, restriction sites can be introduced into a heavy chain framework scaffold flanking at least two or the CDR1, CDR2, and CDR3 (e.g., CDR1+CDR2, CDR1+CDR3, or CDR2+CDR3) for cloning a combination of the corresponding CDR regions into the framework scaffold. In one example, restriction sites can be introduced into a heavy chain framework scaffold flanking all of the CDR1, CDR2, and CDR3 regions for cloning diverse heavy chain CDR1s, CDR2s, and CDR3s at the corresponding locations.

In some instances, restriction sites can be introduced into a light chain scaffold flanking the CDR1 region, the CDR2 region, or the CDR3 region for cloning a plurality of nucleic acids encoding a diverse population of light chain CDR1s, heavy chain CDR2s, or heavy chain CDR3s, respectively. In some instances, restriction sites can be introduced into a light chain framework scaffold flanking at least two or the CDR1, CDR2, and CDR3 (e.g., CDR1+CDR2, CDR1+CDR3, or CDR2+CDR3) for cloning a combination of the corresponding CDR regions into the framework scaffold. In one example, restriction sites can be introduced into a light chain framework scaffold flanking all of the CDR1, CDR2, and CDR3 regions for cloning diverse light chain CDR1s, CDR2s, and CDR3s at the corresponding locations.

Figures 2 and 3 illustrate an exemplary scheme for construction of cassettes for introducing heavy chain CDR1, CDR2, or CDR3 diversities into a heavy chain scaffold and/or cassette for introducing light chain CDR1, CDR2, or CDR3 diversities into a light chain scaffold.

The resultant cassettes can be located in a suitable expression vector for producing the encoded antibodies in a suitable expression, display or folding reporter system.

B. Heavy Chain and Light Chain CDR Populations

The heavy chain and/or light chain CDR1, CDR2, and/or CDR3 populations in the antibody libraries can be derived from naturally-occurring human antibodies. Such CDR

sequences can be obtained by sequencing naturally-occurring antibodies (*e.g.*, human antibodies) in existing natural antibody libraries and analyzing the heavy chain and light chain sequences thus obtained by conventional methods to identify heavy chain and/or light chain CDR sequences. Alternatively, or in addition, naturally-occurring antibody CDR sequences can be obtained by analyzing sequences of such antibodies in publicly available databases of naturally-occurring antibody sequences (*e.g.*, human antibody sequences or camelid VHH antibody sequences), *e.g.*, the NCBI database, the IMGT database, sequences from Jackson et al., *J. Immunol. Methods*, 324:26, 2007, and/or the sequences from Lee et al., *Immunogenetics*, 57:917, 2006, The Observed Antibody Space (antibodymap.org) described in Kovaltsuk, A. *et al.* Observed Antibody Space: A Resource for Data Mining Next-Generation Sequencing of Antibody Repertoires. *Journal of Immunology*, doi:10.4049/jimmunol.1800708 (2018), and/or the iReceptor database (ireceptor.irmacs.sfu.ca) described in Corrie, B. D. *et al.* iReceptor: A platform for querying and analyzing antibody/B-cell and T-cell receptor repertoire data across federated repositories. *Immunol Rev* **284**, 24-41, doi:10.1111/imr.12666 (2018), and/or the sequence database described in Briney, B. et al., Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature*, doi:10.1038/s41586-019-0879-y (2019).

The extent of the framework region and CDRs can be precisely identified using methodology known in the art, for example, by the Kabat definition, the IMGT definition, the Chothia definition, the AbM definition, and/or the contact definition, all of which are well known in the art. See, *e.g.* Kabat, E.A., et al. (1991) Sequences of Proteins of Immunological Interest, Fifth Edition, U.S. Department of Health and Human Services, NIH Publication No. 91-3242, Chothia et al., (1989) *Nature* 342:877; Chothia, C. et al. (1987) *J. Mol. Biol.* 196:901-917, Al-lazikani et al (1997) *J. Molec. Biol.* 273:927-948; Lefranc, M. P. *et al.* IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol* **27**, 55-77 (2003) and Almagro, J. *Mol. Recognit.* 17:132-143 (2004). See also hgmp.mrc.ac.uk, IMGT.org and bioinf.org.uk/abs.

The heavy chain and/or light chain CDR1, CDR2, and/or CDR3 sequences thus obtained may be further analyzed to remove those that comprise a liability, *e.g.*, those listed in Table 4. In some instances, heavy chain and/or light chain CDR1, CDR2, and/or CDR3 sequences comprising one of the liabilities listed in Table 4 (*e.g.*, a glycosylation site, a deamidation site, an isomerization site, an unpaired cysteine, a net charge greater than 1 (*e.g.*,

in LCDR1-2 and/or HC CDR1-2), a tripeptide motif containing at least two aromatic residues (which may affect viscosity), a motif that promotes aggregation, (viii) a polyspecificity site such as those containing a motif of GG, GGG, RR, VG, W, WV, WW, WWW, YY, or, WXW, in which X represents any amino acid residue; a protease sensitive site (fragmentation sensitive site), or an integrin binding site) and/or Figure 5 (using HC CDR2 as an example) can be removed such that the resultant antibody library is free (substantially free or completely free) of members comprising the excluded liability.

Alternatively or in addition, potential glycation sites such as lysine glycation sites may be removed. A glycation site refers to a site in a protein molecule that can be linked to a sugar molecule via a nonenzymatic process. Exemplary glycation sites include, but are not limited to, KE, EK, and ED. Additional liabilities include metal catalyzed fragmentation site (*e.g.*, HS, SH, KT, HXS, or SXH, in which X represents any amino acid residue), polyspecificity aggregation site (*e.g.*, having a motif of $X_1X_2X_3$, in which each of X_1 , X_2 , and X_3 is independently F, I, L, V, W, or Y), and streptavidin binding motif (*e.g.*, HPQ, EPDW (SEQ ID NO: 117), PWXWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), and PWPWL (SEQ ID NO: 120)).

Substantially free means that the number of a heavy or light chain CDR comprising the liability is less than 20% in the library, *e.g.*, less than 15% or less than 10%.

In some examples, heavy chain and/or light chain CDR1, CDR2, and/or CDR3 sequences comprising two or more (*e.g.*, 3, 4, 5, 6, 7, or more) of the liabilities noted above can be removed such that the resultant library is free of (substantially free of or completely free of) members comprising the excluded liabilities. In one example, heavy chain and/or light chain CDR1, CDR2, and/or CDR3 sequences comprising all of the liabilities listed in Table 4 can be removed such that the resultant library is free of (substantially free of or completely free of) members comprising any of the liabilities. Alternatively or in addition, heavy chain and/or light chain CDR1, CDR2, and/or CDR3 sequences comprising all of the liabilities listed in Figure 5 can be removed such that the resultant library is free of (substantially free of or completely free of) members comprising any of the liabilities. In one specific example, heavy chain and/or light chain CDR1, CDR2, and/or CDR3 sequences comprising all of the liabilities disclosed herein can be removed such that the resultant library is free of (substantially free of or completely free of) members comprising any of the liabilities.

In some examples, heavy chain CDR1 and CDR2 sequences comprising one or more of liabilities, *e.g.*, those listed in Table 4 and/or Figure 5, can be removed, while heavy chain CDR3 sequences can be derived from naturally-occurring human antibodies without removal of members having the liabilities. Alternatively, heavy chain CDR3 sequences comprising one or more liabilities can also be removed. Alternatively, or in addition, light chain CDR1, CDR2, and CDR3 sequences comprising one or more of liabilities, *e.g.*, those listed in Table 4 and illustrated by way of example in Figure 5, can be removed.

In some examples, heavy and/or light chain CDR1, CDR2, and/or CDR3 sequences having anomalous lengths can also be excluded. For example, light chain CDR1 and/or CDR2 having a length that is beyond the scope of germline length ± 2 -aa could be excluded. See Figures 16-21.

In some examples, heavy chain CDR1 and CDR2 members containing deamidation sites (*e.g.*, NG, NS, NT, NN, GNF, GNY, GNT, GNG), isomerization sites (*e.g.*, DG, DS, DD), aggregation site (FHW); motifs affecting viscosity (*e.g.*, HYF and HWH), motifs indicating poor developability (*e.g.*, net charge $\geq +1$ in LCDR1-2 and/or HCDR1-2), unpaired cysteine, polyspecificity site (*e.g.*, GGG, RR, VG, VV, VVV, WW, WWW, YY, WXW, X referring to any amino acid residue, and GG), and glycosylation sites (*e.g.*, NXS or NXT, in which X is any amino acid residue except for proline) can be excluded. In some examples, one or more of the following liabilities in heavy chain CDR1 and heavy chain CDR2 members can also be excluded: additional glycosylation sites (*e.g.*, NXC, X being any amino acid residue except for proline), additional deamidation sites (*e.g.*, NA, NH, and/or ND), additional isomerization sites (*e.g.*, DT and/or DH), lysine glycation sites (*e.g.*, KE, EK, and ED), integrin binding sites (*e.g.*, RGD, RYD, LDV, and KGD), protease sensitive sites (fragmentation site) (*e.g.*, DP, DG, DS, DV, DY, DF, DQ, DK, DL, and DD), metal catalyzed fragmentation sites (*e.g.*, HS, SH, KT, HXS, and SXH, in which X represents any amino acid residue), polyspecificity aggregating sites (*e.g.*, having a motif of $X_1X_2X_3$, in which each of X_1 , X_2 , and X_3 independently is F, I, L, V, W, or Y), and/or streptavidin binding sites (*e.g.*, HPQ, EPDW (SEQ ID NO: 117), PXXWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), and PWPWL (SEQ ID NO: 120)).

In some examples, the heavy chain CDR3 members having the one or more liabilities described herein can also be excluded. Alternatively, the heavy chain CDR3 members may

include those derived from naturally-occurring antibodies directly without removal of the one or more liabilities described herein.

Alternatively or in addition, light chain CDR1, CDR2, and/or CDR3 members containing deamidation sites (*e.g.*, NG, NS, NT, NN, GNF, GNY, GNT, GNG),
 5 isomerization sites (*e.g.*, DG, DS, DD), aggregation site (FHW); motifs affecting viscosity (*e.g.*, HYF and HWH), motifs indicating poor developability (*e.g.*, net charge $\geq +1$ in LCDR1-2, HCDR1-2), unpaired cysteine, polyspecificity site (*e.g.*, GGG, RR, VG, VV, VVV, WW, WWW, YY, WXW, X referring to any amino acid residue, and GG), and glycosylation sites (*e.g.*, NXS or NXT, in which X is any amino acid residue except for proline) can be
 10 excluded. In some examples, one or more of the following liabilities in light chain CDR1, CDR2, and/or CDR3 members can also be excluded: additional glycosylation sites (*e.g.*, NXC, X being any amino acid residue except for proline), additional deamidation sites (*e.g.*, NA, NH, and/or ND), additional isomerization sites (*e.g.*, DT and/or DH), lysine glycation sites (*e.g.*, KE, EK, and ED), integrin binding sites (*e.g.*, RGD, RYD, LDV, and KGD),
 15 protease sensitive sites (fragmentation site) (*e.g.*, DP, DG, DS, DV, DY, DF, DQ, DK, DL, and DD), metal catalyzed fragmentation sites (*e.g.*, HS, SH, KT, HXS, and SXH, in which X represents any amino acid residue), polyspecificity aggregatin sites (*e.g.*, having a motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is F, I, L, V, W, or Y), and/or streptavidin binding sites (*e.g.*, HPQ, EPDW (SEQ ID NO: 117), PWXWL (SEQ ID NO:
 20 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), and PWPWL (SEQ ID NO: 120)).

In some examples, the one or more liabilities described herein may be excluded from all of the light chain CDR1, CDR2, and CDR3 members.

The resultant heavy chain and/or light chain CDR1, CDR2, and/or CDR3 sequences
 25 obtained from naturally-occurring antibodies, either excluding sequences comprising one or more liabilities or maintaining all sequences, can be used as templates to synthesis nucleic acids encoding, and replicating, the CDR sequences. Such nucleic acids can be inserted into the corresponding CDR position in the VH and/or VL scaffolds disclosed herein, and are termed “replicated natural CDRs”.

30 When desired, expression vectors carrying the VH and/or VL scaffolds with one or more heavy chain and or light chain CDRs inserted can be introduced into a suitable expression/display system for isolating functional members. Functional members include

those having one or more superior features, for example, good expression and display in a suitable display system, improved folding, reduced aggregation or polyreactivity, and/or greater T_m. Such functional members can be identified by collecting host cells displaying antibodies produced from the expression vectors, and sequencing the corresponding heavy and/or light chain CDR sequences encoded by the expression vectors in the collected host cells.

For example, an initial antibody library may also be sorted for yeast displaying antibodies that have been stained with conformational probes that detect correct antibody folding. Traxlmayr et al., Arch Biochem Biophys. 526(2):174-80, 2012. Examples of such conformational probes include protein A (Hillson et al., The Journal of experimental medicine. 178(1):331-6, 1993; Akerstrom et al., 1994; J. Imm Methods, 177(1-2):151-63, 1994; and Roben et al., J. Immunology 154(12):6437-45, 1995) or protein L (Charbit et al., Gene, 70(1):181-9, 1988; Graille et al., Structure, 9(8):679-87, 2001; and Enever et al., Journal of molecular biology, 347(1):107-20, 2005), that are able to bind to VH3 and VK domains respectively, and derivatives of indole 3-butyric acid (Alves et al., Langmuir, 28(25):9640-8, 2012; Alves et al., Anal Chem., 84(18):7721-8, 2012; Alves et al., Bioconjug Chem., 25(7):1198-202, 2014; and Mustafaoglu et al., Biotechnol Bioeng., 112(7):1327-34, 2015) that binds to the "nucleotide binding site" found in all antibodies (Rajagopalan et al., Proceedings of the National Academy of Sciences of the United States of America, 93(12):6019-24, 1993).

The previous use of conformational probes has been shown to predict high expression and thermostability (Traxlmayr et al., 2012; Shusta et al., J Mol Biol. 292(5):949-56, 1999; Traxlmayr et al., Biochim Biophys Acta., 1824(4):542-9, 2012; Traxlmayr et al., Protein Eng Des Sel., 26(4):255-65, 2013; and Hasenhindl et al., Protein Eng Des Sel., 26(10):675-82, 2013) in yeast display. This approach selects for antibody fragments that are well expressed and well folded. Rather than positive selection for good display, each individual CDR library can be depleted of CDRs that contain liabilities. For example, adapting screens used for antibody screening (Yang et al., MAbs., 5(5):787-94, 2013; Kelly et al., MAbs, 7(4):770-7, 2015; Kohli et al., MAbs. 7(4):752-8, 2015; Obrezanova et al., MAbs., 7(2):352-63, 2015; Wu et al, Protein Eng Des Sel., 28(10):403-14, 2015; Yang et al., MAbs., 9(4):646-53, 2017; Xu et al., Protein Eng Des Sel., 26(10):663-70, 2013; and Kelly et al., MAbs., 9(7):1036-40, 2017) to yeast display sorting, and isolating those yeast displaying antibodies that correspond

to the more “developable” phenotype selects for suitable CDRs that can then be combined to create highly functional libraries.

Examples of such selections include polyspecificity reagents, heparin or chaperones and only retaining those antibodies that do not bind such substances. Further stability
5 increases can be generated by applying a heat shock step (Traxlmayr et al., 2012; Shusta et al., J Mol Biol. 292(5):949-56, 1999; Traxlmayr et al., Biochim Biophys Acta., 1824(4):542-9, 2012; Traxlmayr et al., Protein Eng Des Sel., 26(4):255-65, 2013; and Hasenhindl et al., Protein Eng Des Sel., 26(10):675-82, 2013). See also Figure 23.

The sequences encoding functional members of the heavy and/or light CDR1, CDR2,
10 and/or CDR3 can be used as templates for synthesizing nucleic acids coding for such functional members, or used directly. The resultant nucleic acids can then be inserted into the VH and/or VL scaffold as described herein to produce antibody libraries as also described herein. In some embodiments, the antibody library disclosed herein is substantially free of non-functional members, *e.g.*, having less than 10% (*e.g.*, less than 8%, less than 5%, less
15 than 3%, less than 1%, or lower) non-functional members.

C. Antibody Libraries

The antibody libraries described herein may comprise a plurality of nucleic acids encoding a plurality of antibody heavy chain and/or antibody light chain variable domains,
20 which collectively comprise a common VH and/or VL framework scaffold (*e.g.*, those described herein) with a diverse population of heavy or light chain CDR1s, a diverse population of heavy or light chain CDR2s, and/or a diverse population of heavy or light chain CDRs inserted at the corresponding CDR positions.

In some embodiments, the antibody library described herein is a heavy chain library
25 comprising a plurality of nucleic acids encoding a plurality of antibody heavy chain variable domains. In some examples, the heavy chain library may comprise at least 10^2 diversity of heavy chain CDR1s (having at least 10^2 unique heavy chain CDR1 sequences), for example, at least 10^3 , 10^4 , 10^5 diversity or at least 10^6 diversity. Alternatively, or in addition, the heavy chain library may comprise at least 10^2 diversity of heavy chain CDR2s (having at least 10^2
30 unique heavy chain CDR2 sequences), for example, at least 10^3 , 10^4 , 10^5 diversity or at least 10^6 diversity. In other examples, the heavy chain library may comprise at least 10^2 diversity of heavy chain CDR3s (having at least 10^2 unique heavy chain CDR3 sequences), for

example, at least 10^3 , 10^4 , 10^5 diversity, at least 10^6 diversity, at least 10^7 diversity, or at least 10^8 diversity.

In some examples, the heavy chain library may comprise diversity only in the heavy chain CDR1s, the heavy chain CDR2s, or the heavy chain CDR3s. In other examples, the heavy chain library may comprise diversity in at least two of the heavy chain CDR1, CDR2, and CDR3 regions (*e.g.*, CDR1+CDR2, CDR1+CDR3, or CDR2+CDR3). In one specific example, the heavy chain library comprises diversity in all of the heavy chain CDR1, CDR2, and CDR3 regions.

In some embodiments, the heavy chain library is a secondary library generated for affinity maturation of a pre-selected antibody (the parent antibody) with binding activity to a target antigen. Such a secondary library may comprise diversity in one or two of the heavy chain CDR regions, while keeping the other CDR sequence(s) of the parent antibody. For example, the secondary library may comprise the same heavy CDR1 and CDR2 sequences as the parent antibody, and a diverse population of heavy chain CDR3 sequences. Alternatively, the secondary library may comprise the same heavy CDR3 sequence as the parent antibody and a diverse population of heavy chain CDR1 and/or CDR2 sequences.

Any of the heavy chain libraries disclosed herein may be paired with a common light chain variable region. Alternatively, it may be paired with any of the light chain antibody libraries as also described herein.

Also provided herein are antibody light chain libraries that comprise a plurality of nucleic acids encoding a plurality of antibody light chain variable domains. In some examples, the light chain library may comprise at least 10^2 diversity of light chain CDR1s (having at least 10^2 unique light chain CDR1 sequences), for example, at least 10^3 , 10^4 , 10^5 diversity or at least 10^6 diversity. Alternatively, or in addition, the light chain library may comprise at least 10^2 diversity of light chain CDR2s (having at least 10^2 unique light chain CDR2 sequences), for example, at least 10^3 , 10^4 , 10^5 diversity or at least 10^6 diversity. In other examples, the light chain library may comprise at least 10^2 diversity of light chain CDR3s (having at least 10^2 unique light chain CDR3 sequences), for example, at least 10^3 , 10^4 , 10^5 diversity, at least 10^6 diversity, at least 10^7 diversity, or at least 10^8 diversity.

In some examples, the light chain library may comprise diversity only in the light chain CDR1s, the light chain CDR2s, or the light chain CDR3s. In other examples, the light chain library may comprise diversity in at least two of the light chain CDR1, CDR2, and

CDR3 regions (e.g., CDR1+CDR2, CDR1+CDR3, or CDR2+CDR3). In one specific example, the light chain library comprises diversity in all of the light chain CDR1, CDR2, and CDR3 regions.

In some embodiments, the light chain library is a secondary library generated for affinity maturation of a pre-selected antibody (the parent antibody) with binding activity to a target antigen. Such a secondary library may comprise diversity in one or two of the light chain CDR regions, while keeping the other CDR sequence(s) of the parent antibody. For example, the secondary library may comprise the same light CDR1 and CDR2 sequences as the parent antibody, and a diverse population of light chain CDR3 sequences. Alternatively, the secondary library may comprise the same light CDR3 sequence as the parent antibody and a diverse population of light chain CDR1 and/or CDR2 sequences.

As illustrated in Figure 35 and Figure 36, a secondary library may be generated via VL shuffling and/or VH CDR1 and/or CDR2 shuffling.

Any of the light chain libraries disclosed herein may be paired with a common heavy chain variable region. See, *e.g.*, Figure 26. Alternatively, it may be paired with any of the heavy chain antibody libraries as also described herein. See, *e.g.*, Figure 27.

II. Antibody Library Screening

Any of the antibody libraries described herein may be used to screen for antibodies having binding specificity to an antigen of interest. Antibodies encoded by the nucleic acids in the library can be expressed and displayed using a suitable expressing/display system, for example, a cell-free display system (*e.g.*, ribosome display), a phage display system, a prokaryotic cell-based display system (*e.g.*, bacterial display), or a eukaryotic cell-based display system (*e.g.*, yeast display or mammalian cell display). In certain embodiments, the antibody libraries are expressed and displayed on yeast cells. In other embodiments, the antibody libraries are expressed and displayed on phage particles (phage display). In other embodiments two or more display systems are used, *e.g.* phage display followed by yeast display.

The library of antibodies may be expressed/displayed in a suitable system, *e.g.*, those described herein, in any format. Examples include intact antibodies (full-length antibodies), antigen-binding fragments thereof (*e.g.*, Fab), or single chain antibodies (scFv).

Phage display is a protein display format using bacteriophages (*e.g.*, phage f1, fd, and M13). In this system, at least one antibody chain (*e.g.*, the heavy chain and/or the light chain) is typically covalently linked to a bacteriophage coat protein, for example, a gene III protein, a gene VIII protein, or a major coat protein (see, *e.g.*, WO 00/71694).

5 Phage display is described, for example, in U.S. 5,223,409; Smith (1985) *Science* 228:1315-1317; WO 92/18619; WO 91/17271; WO 92/20791; WO 92/15679; WO 93/01288; WO 92/01047; WO 92/09690; WO 90/02809; de Haard *et al.* (1999) *J. Biol. Chem* 274:18218-30; Hoogenboom *et al.* (1998) *Immunotechnology* 4:1-20; Hoogenboom *et al.* (2000) *Immunol Today* 2:371-8; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372;
10 Hay *et al.* (1992) *Hum Antibod Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J Mol Biol* 226:889-896; Clackson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *PNAS* 89:3576-3580; Garrard *et al.* (1991) *Bio/Technology* 9:1373-1377; and Hoogenboom *et al.* (1991) *Nuc Acid Res* 19:4133-4137.

15 Bacteriophage displaying the protein component can be grown and harvested using standard phage preparatory methods, *e.g.*, PEG precipitation from growth media. After selection of individual display phages, the nucleic acid encoding the selected protein components can be isolated from cells infected with the selected phages or from the phage themselves, after amplification. Individual colonies or plaques can be picked, the nucleic
20 acid isolated and sequenced.

In other embodiments, a eukaryotic expression/display system, *e.g.*, yeast cells or mammalian cells, can be used for expressing and displaying the library of antibodies as described herein. Yeast display is a protein display format, in which a protein component (*e.g.*, an antibody component) is linked to a yeast cell wall protein (*e.g.*, Aga1p or Aga2p)
25 directly or indirectly. In some instances, one chain of an antibody can be covalently fused to the yeast cell wall protein for direct display. In other instances, the association between an antibody component and a yeast cell wall component can be mediated by an intermediate agent. Yeast display is described in, *e.g.*, Cho *et al.*, *J. Immunol. Methods*, 220(1-2):179-188, 1998; Boder *et al.*, *Methods Enzymol.* 192(2):243-248, 2000; van den
30 Beucken *et al.*, *FEBS Lett* 546(2-3):288-294, 2003; and Boder *et al.*, *Arch Biochem Biophys* 526(2):99-106, 2012.

To screen an antibody library as described herein for isolating antibodies capable of binding to a target antigen, the library of antibodies can be in contact with the target antigen under suitable conditions allowing for antibody-antigen binding. Phage particles or host cells displaying antibodies binding to the target antigen can be isolated, for example, by retention on a support member on which the target antigen is immobilized, amplified if needed, and the nucleic acids coding for the displayed antibodies can be determined. The screening process can be repeated multiple times, and display systems can be used in combination. When needed different antigens can be used for selecting antibody members having desired binding specificity or for negative selection to exclude antibody members having binding activity to a non-target antigen.

The screening of the antibodies derived from the libraries described herein can be carried out by any appropriate means. For example, binding activity can be evaluated by standard immunoassay and/or affinity chromatography. Determining the ability of candidate antibodies to bind therapeutic targets can be assayed *in vitro* using, *e.g.*, a BIACORE™ instrument, which measures binding rates of an antibody to a given target antigen based on surface plasmon resonance. *In vivo* assays can be conducted using any of a number of animal models and then subsequently tested, as appropriate, in humans. Cell-based biological assays are also contemplated.

A lead antibody identified from antibody library screening may be subject to affinity maturation as described herein. A secondary library resulting from affinity maturation may be screened for binders having desired features, *e.g.*, high binding affinity and/or binding specificity, following routine practice and/or disclosures provided herein.

General techniques

The practice of the present disclosure will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, such as *Molecular Cloning: A Laboratory Manual*, second edition (Sambrook, et al., 1989) Cold Spring Harbor Press; *Oligonucleotide Synthesis* (M. J. Gait, ed. 1984); *Methods in Molecular Biology*, Humana Press; *Cell Biology: A Laboratory Notebook* (J. E. Cellis, ed., 1989) Academic Press; *Animal Cell Culture* (R. I. Freshney, ed. 1987); *Introduction to Cell and Tissue Culture* (J.

P. Mather and P. E. Roberts, 1998) Plenum Press; Cell and Tissue Culture: Laboratory Procedures (A. Doyle, J. B. Griffiths, and D. G. Newell, eds. 1993-8) J. Wiley and Sons; Methods in Enzymology (Academic Press, Inc.); Handbook of Experimental Immunology (D. M. Weir and C. C. Blackwell, eds.); Gene Transfer Vectors for Mammalian Cells (J. M. Miller and M. P. Calos, eds., 1987); Current Protocols in Molecular Biology (F. M. Ausubel, et al. eds. 1987); PCR: The Polymerase Chain Reaction, (Mullis, et al., eds. 1994); Current Protocols in Immunology (J. E. Coligan et al., eds., 1991); Short Protocols in Molecular Biology (Wiley and Sons, 1999); Immunobiology (C. A. Janeway and P. Travers, 1997); Antibodies (P. Finch, 1997); Antibodies: a practice approach (D. Catty., ed., IRL Press, 1988-1989); Monoclonal antibodies: a practical approach (P. Shepherd and C. Dean, eds., Oxford University Press, 2000); Using antibodies: a laboratory manual (E. Harlow and D. Lane (Cold Spring Harbor Laboratory Press, 1999); The Antibodies (M. Zanetti and J. D. Capra, eds. Harwood Academic Publishers, 1995); *DNA Cloning: A practical Approach*, Volumes I and II (D.N. Glover ed. 1985); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds.(1985»; *Transcription and Translation* (B.D. Hames & S.J. Higgins, eds. (1984»; *Animal Cell Culture* (R.I. Freshney, ed. (1986»; *Immobilized Cells and Enzymes* (IRL Press, (1986»; and B. Perbal, *A practical Guide To Molecular Cloning* (1984); F.M. Ausubel *et al.* (eds.).

Without further elaboration, it is believed that one skilled in the art can, based on the above description, utilize the present invention to its fullest extent. The following specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. All publications cited herein are incorporated by reference for the purposes or subject matter referenced herein.

Example 1: Identifying suitable VH VL pairs for use as scaffolds

Suitable VH/VL pairs as scaffold for CDR insertions are crucial for creating highly diverse, highly functional antibody libraries. The usual rational for selecting scaffolds are: (i) the scaffolds are commonly used in nature (e.g., VH1-69, VH3-23); (ii) the scaffolds are chosen from known antibodies; (iii) the scaffolds are widely used by others; (iv) the scaffolds contains consensus germline sequences; and (v) the VH and VL pairs in the scaffold are considered to be stable, well expressed and non-aggregating.

Provided herein is an alternative approach for identifying suitable VH/VL pairs for use as a scaffold for antibody library construction. A recent publication describing certain approved antibodies and antibodies that are currently approved, or in Phase 2 or Phase 3 clinical trials (Jain, T. *et al.* Biophysical properties of the clinical-stage antibody landscape. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 944-949, doi:10.1073/pnas.1616408114 (2017)) was analyzed. The VH and VL germline genes of these antibodies were determined. Based on the developability data (*e.g.*, aggregation, hydrophobic interaction, polyspecificity, monomericity, expression level in HEK cells, and Fab Tm) provided for each clinical antibody in the paper, the frequency and developability of these germline genes in clinical antibodies was assessed. The antibodies displaying the value in the worst 10% for each of the standard tested was flagged (highlighted in bold and italics in Table 1) and only antibodies with ≤ 1 flag were considered developable. In Table 1 all the clinical antibodies containing ≤ 1 flag are indicated in rows 2-69. In rows 70-72 examples of antibodies with > 2 flags are indicated (Table 1). The antibodies chosen as exemplary examples are highlighted in bold. The correlation between the clinical development stage of the antibodies and the percentage of flags in these antibodies is shown in Figure 1.

Six therapeutic antibodies (abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, and evocelumab) are selected as our final scaffold choices (Highlighted in bold and italics in Table 1, and further described in Table 2). They are well expressed, showing no or minimal liabilities, containing few framework mutations and having low immunogenicity indicating the presence of suitable antibody scaffolds in these antibodies.

For each of the therapeutic antibodies listed in Table 2, seven vectors were designed and synthesized, as described in Example 2. The diagrammatic representation of vectors encoding the original scaffold, the heavy chain CDR1, heavy chain CDR2, heavy CDR3, light chain CDR1, light chain CDR2, and light chain CDR3 scaffolds derived from the six original scaffolds listed in Table 2 is shown in Figure 2. The scaffolds represent five VH gene families (VH1-5) and five VL families (VK1-4; V λ 2). Table 2. Naturally occurring replicated CDRs can be inserted into these selected exemplary antibody scaffolds, which were identified as being well expressed, folded and lacking liabilities, for further testing.

Table 1. Analysis of Therapeutic Antibodies.

Row	Name	HEK Titer (mg/L) <50	Fab T _m by DSF (°C) ≤64	SGAC- SINS AS100 (NH ₄) ₂ SO ₄ mM ≤300	HIC Retention Time (Min) ≥11.8	SMAC Retention Time (Min) ≥13.0	Slope for Accelerated Stability ≥0.09	Poly- Specificity Reagent (PSR) SMP (0- 1) ≥0.27	Affinity - Capture Self- Interaction Nanoparticle Spectroscopy (AC-SINS) Δλ _{max} (nm) Average ≥13.1	CIC Retention Time (Min) >10	CSI-BLI Delta Response (nm) ≥0.02	ELISA >2.0	BVP ELISA ≥4.35
2	abituzumab	89.6	75.5	900.0	9.2	8.7	0.06	0.17	1.5	8.6	0.00	1.14	2.72
3	<i>abrilumab</i>	100.2	71.0	900.0	9.4	8.7	0.03	0.00	-0.9	8.4	-0.02	1.12	1.82
4	adalimumab	134.9	71.0	900.0	8.8	8.7	0.05	0.00	1.1	8.9	-0.01	1.08	1.49
5	alemtuzumab	144.7	74.5	1000.0	8.8	8.7	0.06	0.00	-0.8	8.5	-0.02	1.16	1.46
6	alirocumab	69.2	71.5	900.0	9.0	8.7	0.03	0.00	1.2	8.8	-0.01	1.20	2.18
7	<i>anifrolumab</i>	82.0	62.5	700.0	8.8	8.6	0.07	0.00	-0.6	8.5	-0.02	1.16	1.62
8	bapineuzumab	151.1	73.0	1000.0	8.9	8.7	0.07	0.00	-0.7	8.6	0.06	1.21	3.55

9	benralizumab	146.7	76.0	800.0	9.5	9.1	0.02	0.35	6.0	9.6	-0.01	1.23	1.42
10	brodalumab	150.9	74.5	900.0	9.1	8.7	0.02	0.27	11.2	9.0	-0.01	1.48	2.93
11	canakinumab	45.7	72.0	800.0	9.3	8.7	0.04	0.00	0.7	8.6	0.00	1.20	2.55
12	certolizumab	186.7	81.5	500.0	11.5	10.8	0.04	0.00	0.2	9.3	-0.01	1.14	1.65
13	clazakizumab	113.5	69.5	800.0	9.6	8.9	0.05	0.00	0.9	8.7	-0.03	1.28	4.12
14	<i>crenezumab</i>	149.3	72.0	700.0	10.0	8.7	0.05	0.10	6.4	8.9	0.00	1.13	2.78
15	dacetuzumab	128.5	68.0	1000.0	8.5	8.6	0.00	0.00	0.0	8.5	-0.04	1.08	1.34
16	dacilizumab	245.1	74.0	900.0	9.3	8.8	0.03	0.00	-0.1	8.5	-0.02	1.18	1.41
17	daratumumab	233.3	71.0	800.0	9.5	8.9	0.06	0.00	1.8	8.9	0.00	1.21	3.15
18	eculizumab	226.5	66.0	700.0	10.4	9.3	0.01	0.00	0.0	8.5	-0.04	0.96	3.00
19	efalizumab	167.0	72.5	900.0	8.7	8.6	0.00	0.00	0.7	8.5	-0.02	0.98	1.24

20	elotuzumab	213.2	83.5	700.0	10.3	9.3	0.00	0.00	-0.2	8.5	-0.03	0.98	1.26
21	epratuzumab	78.2	65.0	900.0	9.2	8.6	0.03	0.13	3.0	8.7	-0.01	1.47	2.34
22	<i>evolocumab</i>	260.7	65.0	700.0	10.4	9.1	0.03	0.20	2.2	9.3	-0.01	1.75	1.75
23	farletuzumab	220.8	75.5	800.0	9.5	9.1	0.01	0.00	-0.5	8.7	-0.01	1.07	1.32
24	fasinunab	110.4	71.0	900.0	10.0	8.7	0.07	0.00	-0.7	8.4	-0.02	1.16	2.53
25	ficlatuzumab	249.0	75.0	900.0	9.4	8.9	0.05	0.00	-0.9	8.5	-0.02	1.13	1.23
26	fletikumab	220.4	71.5	700.0	11.0	9.1	0.02	0.00	-0.1	8.5	-0.02	1.02	1.41
27	fresolimumab	166.0	74.0	700.0	10.9	9.1	0.06	0.00	-0.5	8.5	-0.02	1.30	3.51
28	fulranumab	142.0	68.5	900.0	9.3	9.3	0.07	0.19	11.6	9.3	0.00	1.85	6.92
29	gevokizumab	136.4	71.5	1000.0	8.8	8.6	0.07	0.00	-0.5	8.6	-0.03	1.18	1.93
30	ibalizumab	133.3	72.0	800.0	10.2	9.8	0.04	0.00	-0.3	8.7	-0.03	0.99	1.13

31	lintuzumab	230.0	75.5	700.0	10.9	9.4	0.05	0.00	0.9	8.9	-0.02	1.05	1.25
32	matuzumab	224.3	72.0	900.0	9.8	8.8	0.02	0.00	-0.9	8.6	-0.03	1.06	1.03
33	mavrilimumab	150.5	68.5	700.0	10.3	8.7	0.05	0.00	-0.8	8.5	-0.01	1.21	2.16
34	<i>mepolizumab</i>	221.5	78.5	900.0	9.2	8.8	0.04	0.00	-1.0	8.4	-0.04	1.13	1.05
35	mogamulizumab	89.8	68.5	800.0	9.6	8.8	0.04	0.00	-0.5	8.6	-0.02	1.12	2.17
36	motavizumab	133.6	86.0	800.0	9.7	8.8	0.04	0.00	2.5	8.7	-0.01	1.23	5.36
37	natalizumab	251.7	79.5	900.0	9.7	8.8	0.02	0.00	0.8	8.8	-0.01	1.06	1.52
38	<i>necitumumab</i>	198.6	76.5	600.0	10.8	9.9	0.02	0.00	1.3	8.8	-0.02	1.05	1.31
39	nivolumab	178.8	66.0	900.0	9.0	8.7	0.03	0.14	2.4	8.9	-0.01	1.15	1.32
40	obinutuzumab	176.4	73.0	600.0	10.6	9.0	0.01	0.11	1.8	8.8	-0.01	0.95	1.63
41	ofatumumab	249.8	68.0	800.0	9.7	9.5	0.03	0.00	1.2	9.2	-0.02	1.12	1.18

42	olokizumab	115.3	69.0	700.0	9.9	9.0	0.04	0.00	-0.5	8.7	-0.03	1.11	1.23
43	omalizumab	150.4	77.5	800.0	9.5	8.7	0.05	0.00	-0.4	8.5	-0.02	1.12	1.17
44	onartuzumab	147.9	80.0	800.0	9.9	8.9	0.04	0.00	0.0	8.9	-0.02	1.12	1.19
45	otelixizumab	152.1	75.5	1000.0	9.1	8.7	0.09	0.00	4.4	8.7	-0.02	1.13	1.40
46	otlertuzumab	149.6	68.5	600.0	11.0	10.3	0.07	0.00	2.3	9.5	-0.03	1.17	1.78
47	palivizumab	243.1	79.5	900.0	9.3	8.7	0.04	0.00	-0.9	8.5	-0.03	1.12	2.88
48	panitumumab	179.6	78.5	900.0	9.5	8.8	0.04	0.00	-1.1	8.4	-0.03	1.06	1.18
49	panobacumab	107.6	69.0	900.0	9.8	8.9	0.02	0.00	-0.4	9.0	-0.01	1.21	1.90
50	pertuzumab	31.4	78.5	700.0	10.1	8.9	0.04	0.00	-0.2	8.6	-0.04	1.21	1.69
51	pinatuzumab	130.6	79.0	800.0	9.2	8.8	0.07	0.01	0.6	8.8	-0.02	1.27	2.49
52	polatuzumab	225.1	74.0	1000.0	8.8	8.7	0.06	0.00	-1.0	8.3	-0.05	1.36	3.62

53	radretumab	151.2	77.0	900.0	9.5	8.7	0.00	0.13	3.4	8.9	0.00	1.26	3.29
54	ramucirumab	90.7	66.0	900.0	9.4	8.7	0.02	0.00	0.0	8.6	-0.02	1.05	1.25
55	reslizumab	191.6	75.5	700.0	9.8	8.9	0.06	0.23	1.7	8.9	0.00	1.25	2.02
56	romosozumab	227.7	76.0	1000.0	9.2	8.6	0.03	0.00	-1.0	8.4	-0.03	1.01	1.47
57	sarilumab	181.8	64.0	900.0	9.0	8.7	0.05	0.00	1.1	8.7	-0.01	1.19	2.17
58	secukinumab	149.0	72.0	800.0	11.4	8.9	0.05	0.00	-0.6	8.4	-0.04	1.09	1.69
59	sifalimumab	158.6	67.0	800.0	9.7	8.8	0.01	0.06	2.1	9.0	-0.02	2.60	2.50
60	tabalumab	121.6	64.0	700.0	10.8	9.9	0.06	0.00	2.0	9.1	-0.01	1.26	3.68
61	tigatuzumab	179.0	64.5	700.0	10.0	8.7	0.00	0.13	5.5	8.7	-0.01	1.17	1.70
62	tildrakizumab	181.9	77.5	600.0	11.1	9.9	-0.01	0.00	0.8	8.7	-0.01	1.19	1.77
63	tocilizumab	139.6	91.5	900.0	9.1	8.8	0.05	0.00	1.3	8.9	0.00	1.14	2.81

64	tovetumab	277.2	63.5	900.0	8.7	8.6	0.01	0.00	2.2	8.8	-0.01	1.35	2.95
65	trastuzumab	159.5	78.5	800.0	9.7	8.8	0.04	0.00	2.0	8.8	-0.02	1.06	1.34
66	vedolizumab	221.8	80.5	600.0	10.9	12.3	0.07	0.00	0.4	9.0	-0.02	1.15	1.58
67	velutuzumab	225.0	70.0	700.0	11.1	9.7	0.04	0.00	4.8	8.8	-0.02	0.89	1.21
68	zalutunumab	200.5	72.5	900.0	9.3	8.7	0.05	0.00	-0.8	8.4	-0.03	1.28	2.90
69	zanolimumab	116.4	80.5	700.0	9.6	8.8	0.03	0.13	1.5	8.6	-0.01	1.10	1.46
70	atezolizumab	164.1	73.5	300.0	13.4	19.3	0.06	0.07	15.0	10.8	0.06	1.29	6.20
71	belimumab	10.5	60.0	800.0	10.5	9.3	0.13	0.00	0.8	8.6	-0.03	3.61	12.23
72	bevacizumab	50.0	63.5	700.0	11.8	11.1	0.22	0.00	0.8	9.8	-0.02	1.29	2.78

Table 2. Summary of Final Scaffold Choices.

Therapeutic	Target	Phase	Type	VH gene	VK gene	Framework VH mutations	Framework VL mutations	% Immunogenicity*	Fab T _m °C	Phage display used
Abrilumab	$\alpha 4-\beta 7$ integrin	Phase 2	Human	1-24	1-12	1	1	0	71.0	Other H1-24 & K1-12
Mepolizumab	IL-5	Approved	Humanized	2-70	4-1	4	0	6	78.5	Other H2-70 & K4-1
Crenezumab	A β	Phase 3	Humanized	3-7	2D-29	2	3	ND	72.0	None
Necitumumab	EGFR	Approved	Human	4-30-4	3-11	3	1	4.1	76.5	For Necitumumab
Anifrolumab	Interferon receptor	Phase 3	Human	5-51	3-20	1 [^]	2	3.3	62.5	Other H5-51 & K3-20
Evoculumab	PCSK9	Approved	Human	1-18	12-14	2	1	0.3	65	Other H1-18

*These identified CDRs are identical, but synthesized with different flanking regions corresponding to different VH genes (VH1-18 and VH1-24).

[^] These identified CDRs are identical, but synthesized with different flanking regions corresponding to different VK3 genes (VK3-20 and VK3-11).

Example 2: Creating vectors for experimental CDR screening

For each of the six libraries created using the six scaffolds shown in Table 2, seven polynucleotides encoding single-chain variable fragment (scFv) corresponding to each of the scaffolds were synthesized. One of the seven synthesized polynucleotides encodes for the non-modified scFv, and the other six polynucleotides were modified to have one of the original CDRs replaced by a combination of restriction sites including two inverted BsaI sites (a type IIs enzyme that cuts outside of its recognition sequence), an additional SfiI site to ensure cleavage of the vector and serve as a spacer between the BsaI sites, a frameshift and an ochre stop codon to prevent expression of background sequence (Figures 2-3). Each of these modified polynucleotides encoding the scaffolds was cloned into a yeast display vector, and the presence of the stop codon in this sequence prevented the expression of the scaffold on the yeast surface until the modified CDR is replaced with a functional CDR.

Example 3: Generating a database of CDR sequences and informatic elimination of CDRs demonstrating potential liabilities

The process taken to identify suitable CDRs for use in the libraries exemplified herein is illustrated in Figure 4. To generate a database of naturally occurring CDRs (CDRs found in naturally-occurring antibodies such as human antibodies), next generation sequencing (NGS) of the variable genes derived from a total of 40 donors was carried out, comprising a total of >140 million reads. NovaSeq analysis was applied to LCDR3 sequencing data and MiSeq analysis was applied to heavy chain and light chain CDR1 and CDR2 sequencing data. Analysis of the variable gene sequences allowed identification of the numbers of CDRs shown in Table 3.

Altogether, unique CDRs corresponding to the following heavy chain and light chain CDRs were identified showing both unique CDRs and CDRs including potential flanking scaffold oligos (in parentheses). The number of oligos is greater than the number of CDRs because of the need to synthesize some CDRs with different flanking sequences corresponding to different scaffolds.:

- ~ 52,675 (66,020) LCDR1s;
- ~ 19,550 (23,854) LCDR2s;
- ~ 2,180,922 (2,617,051) LCDR3s;
- ~ 147,741 (167,376) HCDR1s;

- ~ 170,758 (202,170) HCDR2s; and
- ~ 13,588,754 HCDR3s

Liabilities that were reduced to short sequences were used to identify CDRs containing them. For example, HCDR2 sequences containing liabilities such as deamidation, isomerization, glycosylation or unpaired cysteines are highlighted in pink (in Figure 5). To underscore the importance of identifying liabilities and eliminating CDRs containing liabilities from the libraries, unpaired cysteines in HCDR3 were illustrated in Figures 6-8 as an example. Cysteines comprise up to 4% of HCDR3 amino acid and they need to be paired and structurally positioned, otherwise the presence of unpaired cysteines or poorly positioned cysteines would introduce undesirable structure or chemical reactivity into the CDR thus rendering the CDR non-functional or non-developable. The list of exemplary liabilities identified is described in Table 4, and it is clear that additional sequence-based liabilities can be similarly screened. The list of unique CDRs previously identified in Table 3 was examined for occurrence of the listed liabilities, and all CDRs containing a liability were computationally eliminated from the list of unique CDRs. Figures 9-13 reflects the extent of elimination of liabilities from the different CDR populations.

Attorney Docket No.: 103073-629979 (S2061.70000WO00)

Table 3. Unique CDRs Identified and the Remaining Unique CDRs after Elimination of liabilities.

Library	VH	VL	unique LCDR1		unique LCDR2		unique LCDR3		unique HCDR1		unique HCDR2		Unique HCDR3		Theoretical diversity	
			total	final	total	final	total	final	total	final	total	final	No HCDR3	with 1e8 HCDR3's		
1 Abritumab	VH1-24	VK1-12	8,838	1,717	6,483	1,406	385,836	74,091	19,635*	2,860	31,412	2,171	13,588,754	1,791,801	1.11E+18	1.11E+26
2 Mepolizumab	VH2-70	VK4-1	6,039	103	685	140	130628	17,917	31,486	2,296	15,550	1,253			7.43E+14	7.43E+22
3 Crenezumab	VH3-7	VK2D-29	4,668	50	1591	229	122745	32,092	34,575	5,920	82,817	4,565			9.93E+15	9.93E+23
4 Necitumumab	VH4-30-4	VK3-11	13,345	1,910	685	972	436129	79,038^	50,335	1,285	28,267	2,739			5.16E+17	5.16E+
5 Amifrolumab	VH5-51	VK3-20	13,345	1,910	685	972	436129	79,038^	11,710	1,979	12,712	669	13,588,754	1,791,801	1.94E+17	1.94E+
6 Evolocumab	VH1-18	V12-14	19,785	1,696	10,792	1,197	1,105,584	94,371	19,635*	2,860	31,412	2,171			1.19E+18	1.19E+
		Total	66,020	5,476	20,921	3,944	2,617,051	297,509	167,376	17,200	202,170	13,568			3,00E+18	3,00E+
Sum of final CDRs (LCDR1-3, HCDR1-2): 337,697.																

*These identified CDRs are identical, but synthesized with different flanking regions corresponding to different VH1 genes (VH1-18 and VH1-24).

^ These identified CDRs are identical, but synthesized with different flanking regions corresponding to different VK3 genes (VK3-20 and VK3-11).

Attorney Docket No.: 103073-629979 (S2061.70000WO00)

Table 4. Liabilities to be Removed.

Type of Liabilities	Exemplary Motif
Glycosylation - impacts stability, solubility, half-life, heterogeneity, and effector function.	NXS, → X = Any Amino Acid but Proline NXT, → X = Any Amino Acid but Proline NXC → X = Any Amino Acid but Proline
Deamidation - Therapeutic antibodies may undergo deamidation during manufacture and storage leading to protein structural changes, aggregation, change in pharmacokinetics, loss of activity and immunogenicity.	NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG
Isomerization - Asp residues can undergo isomerization and reported in CDRs. Known to increase charge heterogeneity	DT, DH, DG, DS, DD
Based on creation of synthetic library, selection against polyspecificity (PSR) and sequencing	GG, GGG, RR, VG, VV, VVV, WW, WWW, YY, WXX (X represents any amino acid residue)
Single cluster in IL-13 human mAb HC3R3 which highly aggregating, alanine mutations increase solubility (aggregation)	FHW
Two aromatic tripeptides in HC3R3 mutated improve viscosity. Compatible with idea that 3 consecutive aromatics is bad news and should be eliminated	HYF, HWH
Positive charge associated with poor developability properties.	Net Charge (+1) in LCDR1-3, HC3R1-2
Unpaired cysteine can impact protein folding, function and stability. These reactive centers lead to formation of covalent aggregates and reduce protein stability	Unpaired Cysteine
Protease sensitivity (fragmentation)	DP, DG, DS, DV, DY, DF, DQ, DK, DL, DD

Attorney Docket No.: 103073-629979 (S2061.70000WO00)

Type of Liabilities	Exemplary Motif
Integrin binding site	RGD, RYD, LDV, KGD
Lysine glycation site	KE, EK, or ED
Metal catalyzed fragmentation	HS, SH, KT, HXS, SXH (X represents any amino acid residue)
Polyspecificity, aggregation	X ₁ X ₂ X ₃ , in which each of X ₁ , X ₂ , and X ₃ independent is F, I, L, V, W, or Y
Streptavidin binding motifs	HPQ, EPDW (SEQ ID NO: 117), PWXXWL (SEQ ID NO: 118), GDWVFI (SEQ ID NO: 119), PWPWL G (SEQ ID NO: 120) (X represents any amino acid residue)

Example 4: Elimination of CDRs arising from sequencing errors

Following the removal of CDRs containing potential liabilities disclosed in Example 3 above, CDRs that may have arisen as a result of sequencing errors were also computationally eliminated. Sequencing errors are more likely when the CDRs sequenced are
5 oversampled. In general, the more copies of a particular CDR, the more likely that it is real, and not the result of a sequencing error.

The abundance of each unique CDR in the dataset was assessed after those containing liabilities had been removed. For each of the different libraries and individual CDRs, the percentage of sequences retained at different abundances (number of reads) was assessed.

10 The more information retained, the more likely that rare CDRs are the result of sequencing errors, and not naturally occurring CDRs. The percentage of sequences eliminated for different CDRs except for HCDR3 at the application of different threshold numbers in each library is represented in Figure 14. A threshold of 4 or more reads was evaluated for each unique CDR to represent the best balance between the number of unique CDRs and the
15 retained sequence information. It is clear that different threshold numbers can be chosen depending upon the number of total reads, and the number of total unique CDRs identified. While different thresholds for each individual CDR for each library could be used, the single threshold of 4 reads was chosen to be consistent. CDRs with less than 4 reads were removed from the library.

20 The use of different thresholds would be expected to yield different numbers of CDRs with different sequences as shown in Figure 15.

Example 5: Elimination of CDRs based on anomalous length

The lengths of germline CDRs are conserved, although they may vary during affinity
25 maturation. CDRs of anomalous length are expected to reduce folding and expression of antibodies that contain them. CDRs with anomalous lengths were also computationally eliminated under the rules described in Figure 4 and CDRs with a length outside of the pink shaded area were eliminated as shown Figures 16-21. The effects on the number of retained CDRs are indicated in Table 5 and Table 6. It is clear that alternative rules may be applied to
30 the identification of naturally occurring CDRs that lack liabilities, and that as new liabilities are identified, they can be similarly eliminated. Furthermore, by carrying out next generation sequencing on CDRs experimentally selected for the presence of liabilities, such as

polyreactivity or aggregation behavior, additional sequence liabilities can be identified and can be eliminated in other libraries created using this approach. As the numbers of CDRs sequenced increases, it is expected that the number of both unique CDRs, as well as the number of unique CDRs lacking liabilities will increase, so increasing the potential library diversity.

After the completion of these different informatic operations, the final number of unique CDRs identified as containing no liabilities, for each scaffold for this exemplary library is indicated in Table 3 showing both unique CDRs and CDRs including flanking scaffold oligos (in parentheses):

- ~ 3,566 (5,476) LCDR1s;
- ~ 2,972 (3,944) LCDR2s;
- ~ 218,471 (297,509) LCDR3s;
- ~ 14,340 (17,200) HCDR1s;
- ~ 11,397 (13,568) HCDR2s; and
- ~ 1,791,801 HCDR3s

Attorney Docket No.: 103073-629979 (S2061.70000WO00)

Table 5. Number of CDRs Retained after Defined Length Restriction.

Library 1	No length restriction	Length Restriction	Difference	Library 2	No length restriction	Length Restriction	Difference	Library 3	No length restriction	Length Restriction	Difference
LCDR1	1,719	1,717	-2	LCDR1	122	103	-19	LCDR1	57	50	-7
LCDR2	1,409	1,406	-3	LCDR2	141	140	-1	LCDR2	231	229	-2
LCDR3	74,134	74,091	-43	LCDR3	17,920	17,917	-3	LCDR3	32,111	32,092	-19
HCDR1	2,878	2,860	-18	HCDR1	2,335	2,296	-39	HCDR1	5,956	5,920	-36
HCDR2	2,189	2,171	-18	HCDR2	1,262	1,253	-9	HCDR2	4,569	4,565	-4
Total	82,329	82,245	-84	Total	21,780	21,709	-71	Total	42,924	42,856	-68
Library 4	No length restriction	Length Restriction	Difference	Library 5	No length restriction	Length Restriction	Difference	Library 6	No length restriction	Length Restriction	Difference
LCDR1	1,917	1,910	-7	HCDR1	1,999	1,979	-20	LCDR1	1,697	1,696	-1
LCDR2	979	972	-7	HCDR2	673	669	-4	LCDR2	1,207	1,197	-10
LCDR3	79,141	79,038	-103	Total	24,452	24,357	-95	LCDR3	94,383	94,371	-12
HCDR1	1,293	1,285	-8					HCDR1	2,878	2,860	-18
HCDR2	2,773	2,739	-34					HCDR2	2,189	2,171	-18
Total	86,103	85,944	-159					Total	102,354	102,295	-59

Table 6. Total Number of CDRs.

	No length restriction	Length Restriction	Difference
Total	338,162	337,697	-465

Example 6: Synthesis and amplification of oligonucleotides corresponding to final CDRs

5 Oligonucleotides corresponding to those identified for HCDR1-2 and LCDR1-3 after the elimination steps as described in the above Examples were synthesized (Twist, Inc., San Francisco, CA), resulting in a total of 337,697 oligonucleotides coding for the selected CDRs. The CDR coding sequence in these oligonucleotides was flanked by 5' and 3' sequences homologous to the framework vectors, into which the CDR coding sequences were cloned.

10 The homologous sequences were used for both amplification and insertion of the oligonucleotides into the yeast display vectors.

The combined pool of replicated natural CDRs amplified using primer pairs specific for each library scaffold and CDR position, cloned into the yeast display vectors described in Example 2 by homologous recombination, resulted in 30 different single CDR loop libraries

15 (6 libraries, LCDR1-3, HCDR1-2). These were sorted for display, using a monoclonal antibody recognizing the SV5 tag by fluorescence activated cell sorting. For each of these libraries, this represents the diversity of replicated natural CDRs that are amplified using the specific primer pairs used (see below) that allow any level of display and are indicated as “Replicated natural before filtering” in figure 22.

20 Figure 22A-F illustrates exemplary advantage of using synthetic oligonucleotides to encode HCDR1-2 and LCDR1-3 replicated natural diversity. Particularly for CDR1-2, the difference in abundance between the most and least abundant CDRs can be >300,000 -fold, with the germline CDR1-2 sequences being by far the most abundant. When natural CDRs are synthesized, the distribution is always far flatter, as shown in figures 22A-F, with the

25 difference in abundance between the most and least abundant CDRs ranging from 10-200 fold, depending upon the CDR and library. The improvement in the diversity at each CDR using the approach described here is further illustrated in figures 22G-L, where the cumulative distribution is indicated for each library and CDR position. The figures for each plot indicate the D50, the number of clones comprising the most abundant 50% of clones,

which for all CDRs in all libraries is significantly higher for the replicated natural diversity, than for the natural diversity. In the case of HCDR3, the VDJ recombinatorial process (including addition of removal of nucleotides at the VD and DJ junctions) results in less variability in abundance between different HCDR3s.

5 The pool of oligonucleotides was subjected to amplification using the following primers:

For L1-CDR1: F-L1-L1-CDR1 to F-L6-L1-CDR1 and R-L1-L1-CDR1 to R-L6-L1-CDR1

For L2-CDR1: F-L1-L2-CDR1 to F-L6-L2-CDR1 and R-L1-L2-CDR1 to R-L6-L2-CDR1

For L3-CDR1: F-L1-L3-CDR1 to F-L6-L3-CDR1 and R-L1-L3-CDR1 to R-L6-L3-CDR1

10 For HCDR1: F-L1-HCDR1 to F-L1-HCDR1 and R-L1-HCDR1 to R-L1-HCDR1

For HCDR2: F-L1-HCDR2 to F-L1-HCDR2 and R-L1-HCDR2 to R-L1-HCDR2

The exemplary amplification primer sequences and assembly primer sequences are provided in Table 7 and Table 8, respectively.

15 **Table 7. Exemplary Amplification Primers.**

	Name	Sequence	T _m (°C)	SEQ ID NO
0	F-L1-L1-CDR1	GCTTCTGTTGGTGATCGTGTACTATTACCTGTCGTGCTTCT	77.9	181
1	F-L2-L1-CDR1	GTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCT	77.9	182
2	F-L3-L1-CDR1	GTTACCCAGGTGAACAGCTTCTATTTCTTGTCGTCTTCT	77.9	183
3	F-L4-L1-CDR1	CTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCT	81.8	184
4	F-L5-L1-CDR1	CTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCT	81.8	185
5	F-L6-L1-CDR1	GGTTCTCCAGGTGAGTCTATCACCATCTCTTGACCGGTACC	80.9	186
6	R-L1-L1-CDR1	CAGTTTTGGAGCTTTACCTGGTTTCTGCTGGTACCAAGCCAG	79.9	187
7	R-L2-L1-CDR1	CAGTTTTGGTGGCTGACCTGGTTTCTGCTGGTACCAAGCCAG	81.8	188
8	R-L3-L1-CDR1	CAGCTGTGGAGACTGACCTGGTTTCTGCAGGTACCAAGTGCAG	82.8	189
9	R-L4-L1-CDR1	CAGACGTGGAGCCTGACCTGGTTTCTGCTGGTACCAAGCCAG	83.8	190
10	R-L5-L1-CDR1	CAGACGTGGAGCCTGACCTGGTTTCTGCTGGTACCAAGCCAG	83.8	191
11	R-L6-L1-CDR1	CAGTTTTGGAGCTTTACCTGGGTGCTGCTGGTACCAAGAAAC	79.9	192
12	F-L1-L2-CDR2	TACCAGCAGAAACCAGGTAAAGCTCCAAAACCTGCTGATCTAC	77.9	193
13	F-L2-L2-CDR2	TACCAGCAGAAACCAGGTGAGCCACCAAAACCTGCTGATCTAC	79.9	194
14	F-L3-L2-CDR2	TACCTGCAGAAACCAGGTGAGTCTCCACAGCTGCTGATCTAC	80.9	195
15	F-L4-L2-CDR2	TACCAGCAGAAACCAGGTGAGGCTCCACGTCTGCTGATCTAC	81.8	196
16	F-L5-L2-CDR2	TACCAGCAGAAACCAGGTGAGGCTCCACGTCTGCTGATCTAC	81.8	197
17	F-L6-L2-CDR2	TACCAGCAGACCCAGGTAAAGCTCCAAAACCTGATGATCTAC	78.9	198
18	R-L1-L2-CDR2	ATCGGTACCAGAACCAGAACCAGAGAAACGAGATGGAACACC	79.9	199
19	R-L2-L2-CDR2	ATCGGTACCAGAACCAGAACCAGAGAAACGATCTGGAACACC	79.9	200
20	R-L3-L2-CDR2	ATCGGTACCAGAACCAGAACCAGAGAAACGATCTGGAACACC	79.9	201
21	R-L4-L2-CDR2	ATCGGTACCAGAACCAGAACCAGAGAAACGAGCTGGGATACC	80.9	202
22	R-L5-L2-CDR2	ATCGGTACCAGAACCAGAACCAGAGAAACGATCTGGGATACC	80.9	203
23	R-L6-L2-CDR2	GGTGTACCAGATTTAGAACAGAGAAACGGTTAGAAACACC	77.0	204
24	F-L1-L3-CDR3	ATCTCTTCTCTGCAGCCAGAAGATTTGCTAACTACTACTGT	77.0	205
25	F-L2-L3-CDR3	ATCTCTTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGT	77.0	206

26	F-L3-LCDR3	ATCTCTCGTGTTGAAGCTGAAGATGTTGGTGTTTACTACTGT	76.0	207
27	F-L4-LCDR3	ATCTCTTCTCTGGAACCAGAAGATTTGCTGTTTACTACTGT	76.0	208
28	F-L5-LCDR3	ATCACCCGTCTGGAACCAGAAGATTTGCTGTTTACTACTGT	77.9	209
29	F-L6-LCDR3	ATCTCTGGTCTGCAGGCTGAAGATGAAGCTGATTACTACTGT	77.9	210
30	R-L1-LCDR3	CGACCCCTCCGGATTTGATTTCAACTTTGGTACCACCACCGAA	79.9	211
31	R-L6-LCDR3	CGACCCCTCCGGACAGAACGGTCAGTTTGGTACCACCACCGAA	83.8	212
32	F-L1-HCDR1	AAAAAACCCAGGTGCTTCTGTAAAGTTTCTTGTAAGTTTCT	72.1	213
33	F-L2-HCDR1	GTAAACCAACCCAGACCTTGACCTGACCTGTACCGTTTCT	80.9	214
34	F-L3-HCDR1	GTTCAAGCCAGGTGGTCTCTGCGTCTGTCTTGCTGCTGCTTCT	81.8	215
35	F-L4-HCDR1	GTAAACCATCTCAGACCTGTCTGACCTGTACCGTTTCT	78.9	216
36	F-L5-HCDR1	AAAAAACCCAGGTGAATCTCTGAAATCTCTTGTAAGTTTCT	73.0	217
37	F-L6-HCDR1	AAAAAACCCAGGTGCTTCTGTAAAGTTTCTTGTAAGTTTCT	73.0	218
38	R-L1-HCDR1	CCATTCCAGACCTTTACCTGGAGCCTGACGAACCCAGTGGAT	81.8	219
39	R-L2-HCDR1	CCATTCCAGAGCTTTACCTGGTGGCTGACGGATCCAGTTAAC	80.9	220
40	R-L3-HCDR1	CAGTTCCAGACCTTTACCTGGAGCCTGACGAACCCAAGACAT	80.9	221
41	R-L4-HCDR1	CCATTCCAGACCTTTACCTGGTGGCTGACGGATCCAAGACCA	81.8	222
42	R-L5-HCDR1	AGATTCCAGACCTTTACCTGGAACCTGACGAACCCAAGCGAT	79.9	223
43	R-L6-HCDR1	CCATTCCAGACCTTGACCTGGAGCCTGACGAACCCAAGAGAT	82.8	224
44	F-L1-HCDR2	GTTGCTCAGGCTCCAGGTAAAGGTCTGGAATGGATGGGTGGT	81.8	225
45	F-L2-HCDR2	ATCCGTCAGCCACCAGGTAAAGCTCTGGAATGGCTGGCTATG	81.8	226
46	F-L3-HCDR2	GTTGCTCAGGCTCCAGGTAAAGGTCTGGAATGGTTGCTTCT	80.9	227
47	F-L4-HCDR2	ATCCGTCAGCCACCAGGTAAAGGTCTGGAATGGATCGGTTAC	80.9	228
48	F-L5-HCDR2	GTTGCTCAGGTTCAGGTAAAGGTCTGGAATCTATGGGTATC	78.9	229
49	F-L6-HCDR2	GTTGCTCAGGCTCCAGGTGAGGCTCTGGAATGGATGGGTGG	83.8	230
50	R-L1-HCDR2	TTCGGTCATGGTAACACGACCTGGAATTTCTGAGCGTAGAT	78.9	231
51	R-L2-HCDR2	AGAGATGGTCAGACGAGATTTGAGAGCAGAGTTGTAAACGAT	77.0	232
52	R-L3-HCDR2	ACGAGAGATGGTGAAACGACCTTTAACAGAATCTGGGTAGTA	77.0	233
53	R-L4-HCDR2	AACAGACATGGTAACACGAGATTTGAGAGATGGGTGTAATC	76.0	234
54	R-L5-HCDR2	AGCAGAGATGGTAACCTGACCTGGAAGATGGAGAGTAACG	79.9	235
55	R-L6-HCDR2	GGTGGTCATGGTACCACGACCTGTCAGTTTCTGAGCGTAGTT	81.8	236
56	F-L1-HCDR3	CTGTCTTCTCTGAAATCTGAGGACACGGCCGTGTATTACTGT	78.9	237
57	F-L2-HCDR3	ATGACCAACATGGATCCTGTGGACACAGCCACATATTACTGT	77.9	238
58	F-L3-HCDR3	ATGAACTCTCTGCGTGCCGAGGACACGGCTGTGTATTACTGT	80.9	239
59	F-L4-HCDR3	GTAACTCTGTACCAGCCGCGGACACGGCTGTGTATTACTGT	80.9	240
60	F-L5-HCDR3	TGGTCTTCTCTGAAAGCTCGGACACCGCCATTTATTACTGT	78.9	241
61	F-L6-HCDR3	CTGCGTTCTCTGCGTTCTGACGACACGGCCGTGTATTACTGT	81.8	242
62	R-JH4	GATTGGTTTGCCGCTAGCTGAGGAGACGGTGACCAGGGTTCC	83.8	243
63	R-JH6	GATTGGTTTGCCGCTAGCTGAGGAGACGGTGACCGTGGTCCC	84.8	244

Table 8. Exemplary Assembly Primers.

	Name	Sequences	T _m (°C)	SEQ ID NO
64	R-L1-LCDR1-ASS	GTAGATCAGCAGTTTTGGAGCTTT	61.8	245
65	R-L2-LCDR1-ASS	GTAGATCAGCAGTTTTGGTGG	59.4	246
66	R-L3-LCDR1-ASS	GTAGATCAGCAGCTGTGGAGA	61.3	247
67	R-L4-LCDR1-ASS	GTAGATCAGCAGACGTGGAG	60.5	248
68	R-L5-LCDR1-ASS	GTAGATCAGCAGACGTGGAG	60.5	249
69	R-L6-LCDR1-ASS	GTAGATCATCAGTTTTGGAGCTTTA	60.9	250
70	F-L1-LCDR2-ASS	CTGGCTTGGTACCAGCAGAAA	61.3	251
71	F-L2-LCDR2-ASS	CTGGCTTGGTACCAGCAGAAA	61.3	252
72	F-L3-LCDR2-ASS	CTGCACTGGTACCTGCAGAAA	61.3	253
73	F-L4-LCDR2-ASS	CTGGCTTGGTACCAGCAGAAA	61.3	254

74	F-L5-LCDR2-ASS	CTGGCTTGGTACCAGCAGAAA	61.3	255
75	F-L6-LCDR2-ASS	GTTTCTTGGTACCAGCAGCAC	61.3	256
76	R-L1-LCDR2-ASS	ACAGTAGTAGTTAGCGAAATCTTCT	60.9	257
77	R-L2-LCDR2-ASS	ACAGTAGTAAACAGCAACATCTTCA	60.9	258
78	R-L3-LCDR2-ASS	ACAGTAGTAAACACCAACATCTTCA	60.9	259
79	R-L4-LCDR2-ASS	ACAGTAGTAAACAGCGAAATCTTCT	60.9	260
80	R-L5-LCDR2-ASS	ACAGTAGTAAACAGCGAAATCTTCT	60.9	261
81	R-L6-LCDR2-ASS	ACAGTAGTAATCAGCTTCATCTTCA	60.9	262
82	F-L1-LCDR3-ASS	GGTGTTCATCTCGTTTCTCT	59.4	263
83	F-L2-LCDR3-ASS	GGTGTTCAGATCGTTTCTCT	59.4	264
84	F-L3-LCDR3-ASS	GGTGTTCAGATCGTTTCTCT	59.4	265
85	F-L4-LCDR3-ASS	GGTATCCCAGCTCGTTTCTCT	61.3	266
86	F-L5-LCDR3-ASS	GGTATCCCAGATCGTCTGTCT	61.3	267
87	F-L6-LCDR3-ASS	GGTGTTCCTAACCGTTTCTCTG	60.3	268
90	R-L1-HCDR1-ASS	ACCACCCATCCATTCCAGAC	60.5	269
91	R-L2-HCDR1-ASS	CATAGCCAGCCATTCCAGAG	60.5	270
92	R-L3-HCDR1-ASS	AGAAGCAACCAGTTCAGACC	61.3	271
93	R-L4-HCDR1-ASS	GTAACCGATCCATTCCAGACC	61.3	272
94	R-L5-HCDR1-ASS	GATACCCATAGATTCCAGACCTTT	61.8	273
95	R-L6-HCDR1-ASS	CCAACCCATCCATTCCAGAC	60.5	274
96	F-L1-HCDR2-ASS	ATCCACTGGGTTCGTGAGG	59.5	275
97	F-L2-HCDR2-ASS	GTAACTGGATCCGTGAGCCA	61.3	276
98	F-L3-HCDR2-ASS	ATGTCTTGGGTTCGTGAGGCT	61.3	277
99	F-L4-HCDR2-ASS	TGGTCTTGGATCCGTGAGC	59.5	278
100	F-L5-HCDR2-ASS	ATCGCTTGGGTTCGTGAGGTT	61.3	279
101	F-L6-HCDR2-ASS	ATCTCTTGGGTTCGTGAGGCT	61.3	280
102	R-L1-HCDR2-ASS	ACAGTAATACACGGCCGTGTC	61.3	281
103	R-L2-HCDR2-ASS	ACAGTAATATGTGGCTGTGTCCA	61.1	282
104	R-L3-HCDR2-ASS	ACAGTAATACACAGCCGTGTC	59.4	283
105	R-L4-HCDR2-ASS	ACAGTAATACACAGCCGTGTC	59.4	284
106	R-L5-HCDR2-ASS	ACAGTAATAAATGGCGGTGTCC	60.3	285
107	R-L6-HCDR2-ASS	ACAGTAATACACGGCCGTGTC	61.3	286
108	F-L1-HCDR3-ASS	ATCTACGCTCAGAAATTCAGG	60.3	287
109	F-L2-HCDR3-ASS	GTTTACAACCTCTGCTCTGAAATCT	60.1	288
110	F-L3-HCDR3-ASS	TACTACCCAGATTCTGTAAAGGT	60.1	289
111	F-L4-HCDR3-ASS	GATTACAACCCATCTCTGAAATCT	60.1	290
112	F-L5-HCDR3-ASS	CGTTACTCTCCATCTTTCCAG	59.4	291
113	F-L6-HCDR3-ASS	AACTACGCTCAGAACTGCAG	59.4	292
114	F-scfv-ASS	CGGATTGTCTTCAACCAACACAA	61.1	293
115	R-scfv-ASS	CTCCTCCTGTTGAATCCAGG	60.5	294
116	F-scfv	CAGTTAGATAAAAGAGGCGCG	59.4	295
117	R-scfv	GCCCAGCAGTGGGTTTGG	60.7	296
88	F-linker-ASS	TCCGGAGGGTCGACCATAA	59.5	297
89	R-linker-ASS	GGTACCGCTCGAGGATAACTT	61.3	298

While the diversity found in HCDR1-2 and LCDR1-3 can be covered by array-based oligonucleotide synthesis relatively easily, this may not be the case for HCDR3 in some instances, where the original diversity can easily exceed 10^8 different HCDR3s. Even after liabilities and CDRs found fewer than 4 times may be eliminated, the number of different

HCDR3s can exceed 10^7 if NovaSeq (3×10^9 reads) is used to assess diversity. This can be addressed either by limiting synthetic HCDR3 diversity to $<10^6$ sequences, which is tractable by array-based synthesis; or by combining synthetic HCDR1-2 and LCDR1-3 diversity with naturally diverse HCDR3 amplified from donor lymphocytes.

RNA from B lymphocytes from Leuko Paks from ten donors, comprising a total of $>10^9$ B cells, was isolated using the Miltenyi StraightFrom LeukoPak CD19 kit. cDNA was prepared using a primer annealing in the IgM constant region. HCDR3s were amplified from the cDNA using all possible combinations of the six forward primers (F-L1-HCDR3 to F-L6-HCDR3) and the two reverse primers (R-JH4 and RJH6) described Table 7 and Table 8. This amplification appends sequences to the 5' and 3' ends homologous to the framework vectors, into which the HCDR3s are to be cloned.

Example 7: Cloning of CDRs into single site CDR vectors and selection for functional CDRs

Although natural replicated CDRs are synthesized based on criteria that should ensure their functionality (*e.g.*, removal of liabilities), oligonucleotide synthesis may not be 100% accurate. In addition to the problem of incorrect sequences, other unidentified liabilities causing poor expression or polyreactivity may be encoded by the synthesized oligonucleotides. Selection of the functional CDRs can be an option to address this issue.

Each of the CDRs remaining from the elimination steps described above was cloned into the appropriate yeast display scaffold vector. The coding sequences of the exemplary scaffold are provided below:

```
> 1 - abrilumab
cagttagataaaagagggcgcgccagcaagcggcgcgcatgccGATATCCAGATGACCCAGTCTCCATCTT
25 CTGTTTCTGCTTCTGTTGGTGATCGTGTTACTATTACCTGTCGTGCTTCTCAGGGTATCTCTTCTTGGCT
GGCTTGGTACCAGCAGAAACCAGGTAAAGCTCCAAAACCTGCTGATCTACGGTGCTTCTAACCTGGAATCT
GGTGTTCCATCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCTCTTCTCTGCAGC
CAGAAGATTTTCGCTAACTACTACTGTGTCAGCAGGCTAACTCTTTCCCATGGACCTTCGGTGGTGGTACCAA
30 AGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATAACGAAGTTATCCTCGAGC
GGTACCCAGGTTTCAGCTGGTTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGCTTCTGTTAAAGTTTCTT
GTAAAGTTTCTGGTTACACCCTGTCTGATCTGTCTATCCACTGGGTTTCGTCAGGCTCCAGGTAAAGGTCT
GGAATGGATGGGTGGTTTCGATCCACAGGATGGTGAAACCATCTACGCTCAGAAATTCCAGGGTCGTGTT
ACCATGACCGAAGATACCTCTACCGATAACCGTTACATGGAACCTGTCTTCTCTGAAATCTGAGGACACGG
CCGTGTATTACTGTGCTACCGGTTCTTCTTCTTCTTGGTTTCGATCCATGGGGTCAGGGAACCTGGTCAC
35 CGTCTCCTCAgctagcggaacccaatcccaaacccactgctgggc (SEQ ID NO: 139)
```

```
> 1a - abrilumab dLCDR1
cagttagataaaagagggcgcgccagcaagcggcgcgcatgccGATATCCAGATGACCCAGTCTCCATCTT
CTGTTTCTGCTTCTGTTGGTGATCGTGTTACTATTACCTGTCGTGCTTCTagagacccatggccagtaagg
```

```
> 1d - abrilumab dHCDR1
cagttagataaaaagaggcgcggcagcaagcgggcgcgcatgccGATATCCAGATGACCCAGTCTCCATCTT
CTGTTTCTGCTTCTGTTGGTGATCGTGTTACTATTACCTGTCGTGCTTCTCAGGGTATCTCTTCTTGGCT
45 GGCTTGGTACCAGCAGAAACCAGGTAAAGCTCCAAAAC TGCTGATCTACGGTGCTTCTAACCTGGAATCT
GGTGTTCCATCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCTGACCATCTCTTCTCTGCAGC
CAGAAGATTTTCGCTAACTACTACTGTCTCAGCAGGCTAACTCTTTCCCATGGACCTTCGGTGGTGGTACCAA
AGTTGAAATCAAATCCGGAGGGTGCACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCGAGC
GGTACCCAGGTTTCAGCTGGTTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGCTTCTGTAAAGTTTCTT
50 GTAAAGTTTCTagagaccatggccagtaagggcgggtctctATCCACTGGGTTCTGTCAGGCTCCAGGTAAA
GGTCTGGAATGGATGGGTGGTTTTCATCCACAGGATGGTGAAACCATCTACGCTCAGAAATTCAGGGTC
GTGTTACCATGACCGAAGATACCTCTACCGATACCGCTTACATGGAAC TGCTTCTCTGAAATCTGAGGA
CACGGCCGTGTATTACTGTGCTACCGGTTCTTCTTCTTCTTGGTTCGATCCATGGGGTCAGGGAACCCCTG
```

GTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 143)
 > 1e - abrilumab dHCDR2
 cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCCAGATGACCCAGTCTCCATCTT
 CTGTTTCTGCTTCTGTTGGTGATCGTGTTACTATTACCTGTCGTGCTTCTCAGGGTATCTCTTCTTGGCT
 5 GGCTTGGTACCAGCAGAAACCAGGTAAAGCTCCAAAACCTGCTGATCTACGGTGCTTCTAACCTGGAATCT
 GGTGTTCCATCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCTGACCATCTCTTCTCTGCAGC
 CAGAAGATTTTCGCTAACTACTACTGTCTGAGCAGGCTAACTCTTTCCCATGGACCTTCGGTGGTGGTACCAA
 AGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATAACGAAGTTATCCTCGAGC
 GGTACCCAGGTTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGCTTCTGTTAAAGTTTCTT
 10 GTAAAGTTTCTGGTTACACCCTGTCTGATCTGTCTATCCACTGGGTTTCGTACGGCTCCAGGTAAAGGTCT
 GGAATGGATGGGTGGTtagagaccatggccagtaaggccgggtctctATCTACGCTCAGAAATTCAGGGTC
 GTGTTACCATGACCGAAGATACCTCTACCGATACCGCTTACATGGAAGTGTCTTCTCTGAAATCTGAGGA
 CACGGCCGTGTATTACTGTGCTACCGGTTCTTCTTCTTCTTGGTTTCGATCCATGGGGTCAGGGAACCCCTG
 GTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 144)
 15 > 1f - abrilumab dHCDR3
 cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCCAGATGACCCAGTCTCCATCTT
 CTGTTTCTGCTTCTGTTGGTGATCGTGTTACTATTACCTGTCGTGCTTCTCAGGGTATCTCTTCTTGGCT
 GGCTTGGTACCAGCAGAAACCAGGTAAAGCTCCAAAACCTGCTGATCTACGGTGCTTCTAACCTGGAATCT
 20 GGTGTTCCATCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCTGACCATCTTCTCTCTGCAGC
 CAGAAGATTTTCGCTAACTACTACTGTCTGAGCAGGCTAACTCTTTCCCATGGACCTTCGGTGGTGGTACCAA
 AGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATAACGAAGTTATCCTCGAGC
 GGTACCCAGGTTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGCTTCTGTTAAAGTTTCTT
 GTAAAGTTTCTGGTTACACCCTGTCTGATCTGTCTATCCACTGGGTTTCGTACGGCTCCAGGTAAAGGTCT
 25 GGAATGGATGGGTGGTTTCGATCCACAGGATGGTGAAACCATCTACGCTCAGAAATTCAGGGTCGTGTT
 ACCATGACCGAAGATACCTCTACCGATACCGCTTACATGGAAGTGTCTTCTCTGAAATCTGAGGACACGG
 CCGTGTATTACTGTtagagaccatggccagtaaggccgggtctctGGAACCCTGGTCACCGTCTCCTCAgct
 agcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 145)
 30 > 2 - mepolizumab
 cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCAGATT
 CTCTGGCTGTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCTCAGTCTCTGCTGAACCTCTGG
 TAACCAGAAAACTACCTGGCTTGGTACCAGCAGAAACCAGGTCAGCCACCAAACTGCTGATCTACGGT
 GCTTCTACCCGTGAATCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCTGA
 35 CCATCTCTTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGTCTGAGAACGTTCACTCTTTCCTTTCAC
 CTTTCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTAT
 ACGAAGTTATCCTCGAGCGGTACCCAGGTTACCCCTGCGTGAATCTGGTCCAGCTCTGGTTAAACCAACCC
 AGACCCCTGACCTGTACCGTTTCTGGTTTCTCTCTGTCTGCTTACTCTGTAACTGGATCCGTCA
 GCCACCAGGTAAAGCTCTGGAATGGCTGGCTATGATCTGGGGTGATGGTAAAATCGTTTACAACTCTGCT
 40 CTGAAATCTCGTCTGACCATCTCTAAAGATACCTCTAAAAACAGGTTGTTCTGACCATGACCAACATGG
 ATCCTGTGGACACAGCCACATATTACTGTGCTGGTGTGGTTACTACCCATACGCTATGGATAACTGGGG
 TCAGGGAACCCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ
 ID NO: 146)
 45 > 2a - mepolizumab dLCDR1
 cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCAGATT
 CTCTGGCTGTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCTtagagaccatggccagtaagg
 ccgggtctctCTGGCTTGGTACCAGCAGAAACCAGGTCAGCCACCAAACTGCTGATCTACGGTGCTTCTA
 CCGTGAATCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCTGACCATCTC
 50 TTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGTCTGAGAACGTTCACTCTTTCCTTTCACCTTCGGT
 GGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATAACGAAGT
 TATCCTCGAGCGGTACCCAGGTTACCCCTGCGTGAATCTGGTCCAGCTCTGGTTAAACCAACCCAGACCCT
 GACCCTGACCTGTACCGTTTCTGGTTTCTCTCTGTCTGCTTACTCTGTAACTGGATCCGTCAGCCACCA

GGTAAAGCTCTGGAATGGCTGGCTATGATCTGGGGTGATGGTAAATCGTTTACAACCTCTGCTCTGAAAT
CTCGTCTGACCATCTCTAAAGATACCTCTAAAAACCAGGTTGTTCTGACCATGACCAACATGGATCCTGT
GGACACAGCCACATATTACTGTGCTGGTGATGGTTACTACCCATACGCTATGGATAACTGGGGTCAGGGA
ACCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO:
5 147)

> 2b - mepolizumab dLCDR2
cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCAGATT
CTCTGGCTGTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCTCAGTCTCTGCTGAACTCTGG
10 TAACCAGAAAACTACCTGGCTTGGTACCAGCAGAAACCAGGTCAGCCACCAAACTGCTGATCTACaga
gaccatggccagtaaggccggtctctGGTGTTCCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTT
CACCTTGACCATCTCTTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGTGAGAAGCTTCACTCTTTC
CCATTACCTTCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATG
TATACTATACGAAGTTATCCTCGAGCGGTACCCAGGTTACCCTGCGTGAATCTGGTCCAGCTCTGGTTAA
15 ACCAACCAGACCCTGACCCTGACCTGTACCGTTTCTGGTTTCTCTCTGTCTGCTTACTCTGTAACTGG
ATCCGTCAGCCACCAGGTAAAGCTCTGGAATGGCTGGCTATGATCTGGGGTGATGGTAAATCGTTTACA
ACTCTGCTCTGAAATCTCGTCTGACCATCTCTAAAGATACCTCTAAAAACCAGGTTGTTCTGACCATGAC
CAACATGGATCCTGTGGACACAGCCACATATTACTGTGCTGGTGATGGTTACTACCCATACGCTATGGAT
AACTGGGGTCAGGGAACCCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgg
20 gc (SEQ ID NO: 148)

> 2c - mepolizumab dLCDR3
cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCAGATT
CTCTGGCTGTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCTCAGTCTCTGCTGAACTCTGG
25 TAACCAGAAAACTACCTGGCTTGGTACCAGCAGAAACCAGGTCAGCCACCAAACTGCTGATCTACGGT
GCTTCTACCCGTGAATCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGA
CCATCTCTTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGTtagagaccatggccagtaaggccggtc
tctTTCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACT
ATACGAAGTTATCCTCGAGCGGTACCCAGGTTACCCTGCGTGAATCTGGTCCAGCTCTGGTTAAACCAAC
30 CCAGACCCTGACCCTGACCTGTACCGTTTCTGGTTTCTCTCTGTCTGCTTACTCTGTAACTGGATCCGT
CAGCCACCAGGTAAAGCTCTGGAATGGCTGGCTATGATCTGGGGTGATGGTAAATCGTTTACAACCTCTG
CTCTGAAATCTCGTCTGACCATCTCTAAAGATACCTCTAAAAACCAGGTTGTTCTGACCATGACCAACAT
GGATCCTGTGGACACAGCCACATATTACTGTGCTGGTGATGGTTACTACCCATACGCTATGGATAACTGG
GGTCAGGGAACCCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc
35 (SEQ ID NO: 149)

> 2d - mepolizumab dHCDR1
cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCAGATT
CTCTGGCTGTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCTCAGTCTCTGCTGAACTCTGG
40 TAACCAGAAAACTACCTGGCTTGGTACCAGCAGAAACCAGGTCAGCCACCAAACTGCTGATCTACGGT
TTCTACCCGTGAATCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGA
CCATCTCTTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGTGAGAAGCTTCACTCTTCCCATTAC
CTTCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTAT
ACGAAGTTATCCTCGAGCGGTACCCAGGTTACCCTGCGTGAATCTGGTCCAGCTCTGGTTAAACCAACCC
45 AGACCCTGACCCTGACCTGTACCGTTTCTtagagaccatggccagtaaggccggtctctGTTAACTGGATC
CGTCAGCCACCAGGTAAAGCTCTGGAATGGCTGGCTATGATCTGGGGTGATGGTAAATCGTTTACAACCT
CTGCTCTGAAATCTCGTCTGACCATCTCTAAAGATACCTCTAAAAACCAGGTTGTTCTGACCATGACCAA
CATGGATCCTGTGGACACAGCCACATATTACTGTGCTGGTGATGGTTACTACCCATACGCTATGGATAAC
TGGGGTCAGGGAACCCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc
50 (SEQ ID NO: 150)

> 2e - mepolizumab dHCDR2
cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCAGATT

CTCTGGCTGTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCTCAGTCTCTGCTGAACTCTGG
 TAACCAGAAAACTACCTGGCTTGGTACCAGCAGAAACCAGGTCAGCCACCAAACTGCTGATCTACGGT
 GCTTCTACCCGTGAATCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGA
 CCATCTCTTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGTGAGAACGTTCACTCTTTCCCATTAC
 5 CTTCCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTAT
 ACGAAGTTATCCTCGAGCGGTACCCAGGTTACCCTGCGTGAATCTGGTCCAGCTCTGGTTAAACCAACCC
 AGACCCCTGACCCTGACCTGTACCGTTTCTGGTTTCTCTCTGTCTGCTTACTCTGTAACTGGATCCGTCA
 GCCACCAGGTAAAGCTCTGGAATGGCTGGCTATGagagaccatggccagtaaggccggtctctATCGTTT
 ACAACTCTGCTCTGAAATCTCGTCTGACCATCTCTAAAGATACCTCTAAAAACCAGGTTGTTCTGACCAT
 10 GACCAACATGGATCCTGTGGACACAGCCACATATTACTGTGCTGGTGTATGGTTACTACCCATACGCTATG
 GATAACTGGGGTCAGGGAACCCCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgc
 tgggc (SEQ ID NO: 151)

> 2f - mepolizumab dHCDR3
 15 cagttagataaaagagggcgccgagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCAGATT
 CTCTGGCTGTTTCTCTGGGTGAACGTGCTACCATCAACTGCAAATCTTCTCAGTCTCTGCTGAACTCTGG
 TAACCAGAAAACTACCTGGCTTGGTACCAGCAGAAACCAGGTCAGCCACCAAACTGCTGATCTACGGT
 GCTTCTACCCGTGAATCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGA
 CCATCTCTTCTCTGCAGGCTGAAGATGTTGCTGTTTACTACTGTGAGAACGTTCACTCTTTCCCATTAC
 20 CTTCCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTAT
 ACGAAGTTATCCTCGAGCGGTACCCAGGTTACCCTGCGTGAATCTGGTCCAGCTCTGGTTAAACCAACCC
 AGACCCCTGACCCTGACCTGTACCGTTTCTGGTTTCTCTCTGTCTGCTTACTCTGTAACTGGATCCGTCA
 GCCACCAGGTAAAGCTCTGGAATGGCTGGCTATGATCTGGGGTGATGGTAAAATCGTTTACAACTCTGCT
 CTGAAAATCTCGTCTGACCATCTCTAAAGATACCTCTAAAAACCAGGTTGTTCTGACCATGACCAACATGG
 25 ATCCTGTGGACACAGCCACATATTACTGTtagagaccatggccagtaaggccggtctctGGAACCCCTGGTC
 ACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 152)

> 3 - crenezumab
 30 cagttagataaaagagggcgccgagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCACTGT
 CTCTGCCAGTTACCCAGGTGAACCAGCTTCTATTTCTTGTCGTTCTTCTCAGTCTCTGGTTTACTCTAA
 CGGTGATACCTACCTGCACTGGTACCTGCAGAAACCAGGTCAGTCTCCACAGCTGCTGATCTACAAAGTT
 TCTAACCGTTTCTCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGAAAA
 TCTCTCGTGTGAAGCTGAAGATGTTGGTGTTTACTACTGTTCTCAGTCTACCCACGTTCCATGGACCTT
 CGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACG
 35 AAGTTATCCTCGAGCGGTACCGAAGTTCACTGGTTGAATCTGGTGGTGGTCTGGTTTCAGCCAGGTGGTT
 CTCTGCGTCTGTCTTGCTGCTTCTGGTTTACCTTCTCTTCTTACGGTATGTCTTGGGTTCTGTCAGGC
 TCCAGGTAAAGTCTGGAAGTGGTTGCTTCTATCAACTCTAACGGTGGTTCTACCTACTACCCAGATTCT
 GTTAAAGTCTGTTTACCATCTCTCGTGATAACGCTAAAACTCTCTGTACCTGCAGATGAACTCTCTGC
 GTGCCGAGGACACGGCTGTGTATTACTGTGCTTCTGGTGATTACTGGGGTCAGGGGACCACGGTCACCGT
 40 CTCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 153)

> 3a - crenezumab dLCDR1
 45 cagttagataaaagagggcgccgagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCACTGT
 CTCTGCCAGTTACCCAGGTGAACCAGCTTCTATTTCTTGTCGTTCTTCTagagaccatggccagtaagg
 ccggtctctCTGCACTGGTACCTGCAGAAACCAGGTCAGTCTCCACAGCTGCTGATCTACAAAGTTTCTA
 ACCGTTTCTCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGAAAATCTC
 TCGTGTGAAGCTGAAGATGTTGGTGTTTACTACTGTTCTCAGTCTACCCACGTTCCATGGACCTTCGGT
 GGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGT
 TATCCTCGAGCGGTACCGAAGTTCACTGGTTGAATCTGGTGGTGGTCTGGTTTCAGCCAGGTGGTTCTCT
 50 GCGTCTGTCTTGCTGCTTCTGGTTTACCTTCTCTTCTTACGGTATGTCTTGGGTTCTGTCAGGCTCCA
 GGTAAAGTCTGGAAGTGGTTGCTTCTATCAACTCTAACGGTGGTTCTACCTACTACCCAGATTCTGTTA
 AAGTCTGTTTACCATCTCTCGTGATAACGCTAAAACTCTCTGTACCTGCAGATGAACTCTCTGCGTGC
 CGAGGACACGGCTGTGTATTACTGTGCTTCTGGTGATTACTGGGGTCAGGGGACCACGGTCACCGTCTCC

TCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 154)

> 3b - crenezumab dLCDR2

cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCACTGT
 5 CTCTGCCAGTTACCCAGGTGAACCAGCTTCTATTTCTTGTCGTTCTTCTCAGTCTCTGGTTTACTCTAA
 CGGTGATACCTACCTGCACTGGTACCTGCAGAAACCAGGTCAGTCTCCACAGCTGCTGATCTACagagac
 catggccagtaaggccgggtctctGGTGTTCCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTCAC
 CCTGAAAATCTCTCGTGTTGAAGCTGAAGATGTTGGTGTCTTACTACTGTTCTCAGTCTACCCACGTTCCA
 TGGACCTTCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTAT
 10 ACTATACGAAGTTATCCTCGAGCGGTACCGAAGTTCAGCTGGTTGAATCTGGTGGTGGTCTGGTTTCAGCC
 AGGTGGTTCTCTGCGTCTGTCTTGTGCTGCTTCTGGTTTCACCTTCTCTTCTTACGGTATGTCTTGGGTT
 CGTCAGGCTCCAGGTAAAGGTCTGGAAGTGGTTGCTTCTATCAACTCTAACGGTGGTTCTACCTACTACC
 CAGATTCTGTTAAAGGTCGTTTCACCATCTCTCGTGATAACGCTAAAACTCTCTGTACCTGCAGATGAA
 CTCTCTGCGTGCCGAGGACACGGCTGTGTATTACTGTGCTTCTGGTGATTACTGGGGTCAGGGGACCACG
 15 GTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 155)

> 3c - crenezumab dLCDR3

cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCACTGT
 CTCTGCCAGTTACCCAGGTGAACCAGCTTCTATTTCTTGTCGTTCTTCTCAGTCTCTGGTTTACTCTAA
 20 CGGTGATACCTACCTGCACTGGTACCTGCAGAAACCAGGTCAGTCTCCACAGCTGCTGATCTACAAAGTT
 TCTAACCGTTTCTCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTCACCTGAAAA
 TCTCTCGTGTTGAAGCTGAAGATGTTGGTGTCTTACTACTGTagagaccatggccagtaaggccgggtctct
 TTCGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATA
 CGAAGTTATCCTCGAGCGGTACCGAAGTTCAGCTGGTTGAATCTGGTGGTGGTCTGGTTTCAGCCAGGTGG
 25 TTCTCTGCGTCTGTCTTGTGCTGCTTCTGGTTTCACCTTCTCTTCTTACGGTATGTCTTGGGTTTCGT
 GCTCCAGGTAAAGGTCTGGAAGTGGTTGCTTCTATCAACTCTAACGGTGGTTCTACCTACTACCCAGATT
 CTGTTAAAGGTCGTTTCACCATCTCTCGTGATAACGCTAAAACTCTCTGTACCTGCAGATGAACTCTCT
 GCGTGCCGAGGACACGGCTGTGTATTACTGTGCTTCTGGTGATTACTGGGGTCAGGGGACCACGGTCACC
 GTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 156)

30

> 3d - crenezumab dHCDR1

cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCACTGT
 CTCTGCCAGTTACCCAGGTGAACCAGCTTCTATTTCTTGTCGTTCTTCTCAGTCTCTGGTTTACTCTAA
 CGGTGATACCTACCTGCACTGGTACCTGCAGAAACCAGGTCAGTCTCCACAGCTGCTGATCTACAAAGTT
 35 TCTAACCGTTTCTCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTCACCTGAAAA
 TCTCTCGTGTTGAAGCTGAAGATGTTGGTGTCTTACTACTGTTCTCAGTCTACCCACGTTCCATGGACCTT
 CGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACG
 AAGTTATCCTCGAGCGGTACCGAAGTTCAGCTGGTTGAATCTGGTGGTGGTCTGGTTTCAGCCAGGTGGTT
 CTCTGCGTCTGTCTTGTGCTGCTTCTagagaccatggccagtaaggccgggtctctATGCTTGGGTTTCGT
 40 CAGGCTCCAGGTAAAGGTCTGGAAGTGGTTGCTTCTATCAACTCTAACGGTGGTTCTACCTACTACCCAG
 ATTCTGTTAAAGGTCGTTTCACCATCTCTCGTGATAACGCTAAAACTCTCTGTACCTGCAGATGAACTC
 TCTGCGTGCCGAGGACACGGCTGTGTATTACTGTGCTTCTGGTGATTACTGGGGTCAGGGGACCACGGTC
 ACCGTCCTCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 157)

45 > 3e - crenezumab dHCDR2

cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCACTGT
 CTCTGCCAGTTACCCAGGTGAACCAGCTTCTATTTCTTGTCGTTCTTCTCAGTCTCTGGTTTACTCTAA
 CGGTGATACCTACCTGCACTGGTACCTGCAGAAACCAGGTCAGTCTCCACAGCTGCTGATCTACAAAGTT
 TCTAACCGTTTCTCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTCACCTGAAAA
 50 TCTCTCGTGTTGAAGCTGAAGATGTTGGTGTCTTACTACTGTTCTCAGTCTACCCACGTTCCATGGACCTT
 CGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACG

AAGTTATCCTCGAGCGGTACCGAAGTTCAGCTGGTTGAATCTGGTGGTGGTCTGGTTCAGCCAGGTGGTT
 CTCTGCGTCTGTCTTGTGCTGCTTCTGGTTTACCTTCTCTTCTTACGGTATGTCTTGGGTTTCGTAGGC
 TCCAGGTAAAGGTCTGGAAGTGGTTGCTTCTAGagaccatggccagtaaggccggtctctTACTACCCAG
 ATTCTGTAAAGGTCTGTTTACCATCTCTCGTGATAACGCTAAAACTCTCTGTACCTGCAGATGAACTC
 5 TCTGCGTGCCGAGGACACGGCTGTGTATTACTGTGCTTCTGGTGATTACTGGGGTCAGGGGACCACGGTC
 ACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 158)

> 3f - crenezumab dHCDR3

cagttagataaaagagggcggcagcaagcggcgcgcatgccGATATCGTTATGACCCAGTCTCCACTGT
 10 CTCTGCCAGTTACCCAGGTGAACAGCTTCTATTTCTTGTCTGTTCTTCTCAGTCTCTGGTTTACTCTAA
 CGGTGATACCTACCTGCACTGGTACCTGCAGAAACCAGGTCAGTCTCCACAGCTGCTGATCTACAAAGTT
 TCTAACCGTTTTCTCTGGTGTTCAGATCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTCACCCTGAAAA
 TCTCTCGTGTGAAGCTGAAGATGTTGGTGTCTTACTACTGTTCTCAGTCTACCCACGTTCATGGACCTT
 CGGTGGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACG
 15 AAGTTATCCTCGAGCGGTACCGAAGTTCAGCTGGTTGAATCTGGTGGTGGTCTGGTTCAGCCAGGTGGTT
 CTCTGCGTCTGTCTTGTGCTGCTTCTGGTTTACCTTCTCTTCTTACGGTATGTCTTGGGTTTCGTAGGC
 TCCAGGTAAAGGTCTGGAAGTGGTTGCTTCTATCAACTCTAACGGTGGTTCTACCTACTACCCAGATTCT
 GTTAAAGGTCTGTTTACCATCTCTCGTGATAACGCTAAAACTCTCTGTACCTGCAGATGAACTCTCTGC
 GTGCCGAGGACACGGCTGTGTATTACTGTAGagaccatggccagtaaggccggtctctGGGACCACGGTC
 20 ACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 159)

> 4 - necitumumab

cagttagataaaagagggcggcagcaagcggcgcgcatgccGAAATCGTTATGACCCAGTCTCCAGCTA
 CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCTGCTTCTCAGTCTGTTTCTTCTTACCT
 25 GGCTTGGTACCAGCAGAAACCAGGTACGGCTCCACGTCTGCTGATCTACGATGCTTCTAACCGTGCTACC
 GGTATCCCAGCTCGTTTTCTCTGGTTCTGGTTCTGGTACCGATTTCACCCTGACCATCTCTTCTCTGGAAC
 CAGAAGATTTCTGCTGTTTACTACTGTACCCAGTACGGTTCTACCCCACTGACCTTCGGTGGTGGTACCAA
 AGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCGAGC
 GGTACCCAGTTTCAGCTGCAGGAATCTGGTCCAGGTCTGGTTAAACCATCTCAGACCCTGTCTCTGACCT
 30 GTACCGTTTTCTGGTGGTTCTATCTCTTCTGGTGATTACTACTGGTCTTGGATCCGTCAGCCACCAGGTAA
 AGGTCTGGAATGGATCGGTTACATCTACTACTCTGGTTCTACCGATTACAACCCATCTCTGAAATCTCGT
 GTTACCATGTCTGTTGATACCTCTAAAAACCAGTTCTCTCTGAAAGTTAACTCTGTTACCGCCGCGGACA
 CGGCTGTGTATTACTGTGCTCGTGTCTTCTATCTTCGGTGTGGTACCTTCGATTACTGGGGTCAGGGAAC
 CCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO:
 35 160)

> 4a - necitumumab dLCDR1

cagttagataaaagagggcggcagcaagcggcgcgcatgccGAAATCGTTATGACCCAGTCTCCAGCTA
 CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCTGCTTCTAGagaccatggccagtaagg
 40 ccggtctctCTGGCTTGGTACCAGCAGAAACCAGGTACGGCTCCACGTCTGCTGATCTACGATGCTTCTA
 ACCGTGCTACCGGTATCCCAGCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTCACCCTGACCATCTC
 TTCTCTGGAACCAGAAGATTTCTGCTGTTTACTACTGTACCCAGTACGGTTCTACCCCACTGACCTTCGGT
 GGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGT
 TATCCTCGAGCGGTACCCAGGTTCAGCTGCAGGAATCTGGTCCAGGTCTGGTTAAACCATCTCAGACCCT
 45 GTCTCTGACCTGTACCGTTTTCTGGTGGTTCTATCTCTTCTGGTGATTACTACTGGTCTTGGATCCGTCAG
 CCACCAGGTAAAGGTCTGGAATGGATCGGTTACATCTACTACTCTGGTTCTACCGATTACAACCCATCTC
 TGAAATCTCGTGTTACCATGTCTGTTGATACCTCTAAAAACCAGTTCTCTCTGAAAGTTAACTCTGTTAC
 CGCCGCGGACACGGCTGTGTATTACTGTGCTCGTGTCTTCTATCTTCGGTGTGGTACCTTCGATTACTGG
 GGTCAGGGAACCGTGGTACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc
 50 (SEQ ID NO: 161)

> 4b - necitumumab dLCDR2

cagttagataaaagagggcgcggcagcaagcgggcgcgcatgccGAAATCGTTATGACCCAGTCTCCAGCTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTACCT
GGCTTGGTACCAGCAGAAACCAGGTGAGGCTCCACGTCTGCTGATCTACagagaccatggccagtaaggc
5 cgggtctctGGTATCCCAGCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCTCTTC
TCTGGAACCAGAAGATTTTCGTGTTTACTACTGTCACCAGTACGGTTCTACCCCACTGACCTTCGGTGGT
GGTACCAAAGTTGAAATCAAATCCGGAGGGTTCGACCATAACTTCGTATAATGTATACTATACGAAGTTAT
CCTCGAGCGGTACCAGGTTTCAGCTGCAGGAATCTGGTCCAGGTCTGGTTAAACCATCTCAGACCCCTGTC
TCTGACCTGTACCGTTTCTGGTGGTTCTATCTCTTCTGGTGATTACTACTGGTCTTGGATCCGTCAGCCA
10 CCAGGTAAAGGTCTGGAATGGATCGGTTACATCTACTACTCTGGTTCTACCGATTACAACCCATCTCTGA
AATCTCGTGTTACCATGTCTGTTGATACCTCTAAAAACCAGTTCTCTCTGAAAGTTAACTCTGTTACCGC
CGCGGACACGGCTGTGTATTACTGTGCTCGTGTTTCTATCTTCGGTGTTGGTACCTTCGATTACTGGGGT
CAGGGAACCCCTGGTACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ
ID NO: 162)

> 4c - necitumumab dLCDR3

cagttagataaaagagggcgcggcagcaagcgggcgcgcatgccGAAATCGTTATGACCCAGTCTCCAGCTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTACCT
GGCTTGGTACCAGCAGAAACCAGGTGAGGCTCCACGTCTGCTGATCTACGATGCTTCTAACCGTGCTACC
20 GGTATCCCAGCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCTCTTCTCTGGAAC
CAGAAGATTTTCGTGTTTACTACTGTtagagaccatggccagtaaggccgggtctctTTCGGTGGTGGTACC
AAAGTTGAAATCAAATCCGGAGGGTTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCGA
GCGGTACCAGGTTTCAGCTGCAGGAATCTGGTCCAGGTCTGGTTAAACCATCTCAGACCCCTGTCTCTGAC
CTGTACCGTTTCTGGTGGTTCTATCTCTTCTGGTGATTACTACTGGTCTTGGATCCGTCAGCCACCAGGT
25 AAAGGTCTGGAATGGATCGGTTACATCTACTACTCTGGTTCTACCGATTACAACCCATCTCTGAAATCTC
GTGTTACCATGTCTGTTGATACCTCTAAAAACCAGTTCTCTCTGAAAGTTAACTCTGTTACCGCCGCGGA
CACGGCTGTGTATTACTGTGCTCGTGTTTCTATCTTCGGTGTTGGTACCTTCGATTACTGGGGTCAGGGA
ACCCCTGGTACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO:
163)

> 4d - necitumumab dHCDR1

cagttagataaaagagggcgcggcagcaagcgggcgcgcatgccGAAATCGTTATGACCCAGTCTCCAGCTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTACCT
GGCTTGGTACCAGCAGAAACCAGGTGAGGCTCCACGTCTGCTGATCTACGATGCTTCTAACCGTGCTACC
35 GGTATCCCAGCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCTCTTCTCTGGAAC
CAGAAGATTTTCGTGTTTACTACTGTGTCACCAGTACGGTTCTACCCCACTGACCTTCGGTGGTGGTACCAA
AGTTGAAATCAAATCCGGAGGGTTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCGAGC
GGTACCAGGTTTCAGCTGCAGGAATCTGGTCCAGGTCTGGTTAAACCATCTCAGACCCCTGTCTCTGACCT
GTACCGTTTCTagagaccatggccagtaaggccgggtctctTGGTCTTGGATCCGTCAGCCACCAGGTAAA
40 GGTCTGGAATGGATCGGTTACATCTACTACTCTGGTTCTACCGATTACAACCCATCTCTGAAATCTCGTG
TTACCATGTCTGTTGATACCTCTAAAAACCAGTTCTCTCTGAAAGTTAACTCTGTTACCGCCGCGGACAC
GGCTGTGTATTACTGTGCTCGTGTTTCTATCTTCGGTGTTGGTACCTTCGATTACTGGGGTCAGGGAACC
CTGGTACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO:
164)

> 4e - necitumumab dHCDR2

cagttagataaaagagggcgcggcagcaagcgggcgcgcatgccGAAATCGTTATGACCCAGTCTCCAGCTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTACCT
GGCTTGGTACCAGCAGAAACCAGGTGAGGCTCCACGTCTGCTGATCTACGATGCTTCTAACCGTGCTACC
50 GGTATCCCAGCTCGTTTCTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCTCTTCTCTGGAAC
CAGAAGATTTTCGTGTTTACTACTGTGTCACCAGTACGGTTCTACCCCACTGACCTTCGGTGGTGGTACCAA
AGTTGAAATCAAATCCGGAGGGTTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCGAGC

GGTACCCAGGTTTCAGCTGCAGGAATCTGGTCCAGGTCTGGTTAAACCATCTCAGACCCTGTCTCTGACCT
GTACCGTTTTCTGGTGGTTCTATCTCTTCTGGTGATTACTACTGGTCTTGGATCCGTCAGCCACCAGGTAA
AGGTCTGGAATGGATCGGTTACagagaccatggccagtaaggccggtctctGATTACAACCCATCTCTGA
AATCTCGTGTTACCATGTCTGTTGATACCTCTAAAAACCAGTTCTCTCTGAAAGTTAACTCTGTTACCGC
5 CGCGGACACGGCTGTGTATTACTGTGCTCGTGTCTTCTATCTTCGGTGTGGTACCTTCGATTACTGGGGT
CAGGGAACCCCTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ
ID NO: 165)

> 4f - necitumumab dHCDR3

10 cagttagataaaaagagggcgcgccagcaagcggcgcgcatgccGAAATCGTTATGACCCAGTCTCCAGCTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCTGCTTCTCAGTCTGTTTCTTCTTACCT
GGCTTGGTACCAGCAGAAACCAGGTACGGCTCCACGTCTGCTGATCTACGATGCTTCTAACCGTGCTACC
GGTATCCCAGCTCGTTTCTCTGGTTCTGGTCTGGTACCGATTTACCCCTGACCATCTCTTCTCTGGAAC
CAGAAGATTTCTGCTGTTTACTACTGTACCACTACGGTTCTACCCACTGACCTTCGGTGGTGGTACCAA
15 AGTTGAAATCAAATCCGGAGGGTTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCGAGC
GGTACCCAGGTTTCAGCTGCAGGAATCTGGTCCAGGTCTGGTTAAACCATCTCAGACCCTGTCTCTGACCT
GTACCGTTTTCTGGTGGTTCTATCTCTTCTGGTGATTACTACTGGTCTTGGATCCGTCAGCCACCAGGTAA
AGGTCTGGAATGGATCGGTTACATCTACTACTCTGGTTCTACCGATTACAACCCATCTCTGAAATCTCGT
GTTACCATGTCTGTTGATACCTCTAAAAACCAGTTCTCTCTGAAAGTTAACTCTGTTACCGCCGCGGACA
20 CGGCTGTGTATTACTGTAGagaccatggccagtaaggccggtctctGGAACCCCTGGTCACCGTCTCCTCA
gctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 166)

> 5 - anifrolumab

cagttagataaaaagagggcgcgccagcaagcggcgcgcatgccGAAATCGTTCTGACCCAGTCTCCAGGTA
25 CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCTGCTTCTCAGTCTGTTTCTTCTTCTTT
CTTGGCTTGGTACCAGCAGAAACCAGGTACGGTCCACGTCTGCTGATCTACGGTGCTTCTTCTCGTGCT
ACCGGTATCCCAGATCGTCTGTCTGGTTCTGGTACCGATTTACCCCTGACCATCACCCGCTCTGG
AACAGAAAGATTTCTGCTGTTTACTACTGTACGAGTACGATTCTTCTGCTATCACCTTCGGTGGTGGTAC
CAAAGTTGAAATCAAATCCGGAGGGTTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCG
30 AGCGGTACCGAAGTTTCAGCTGGTTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGAATCTCTGAAAATCT
CTTGTAAGGTTCTGGTTACATCTTACCAACTACTGGATCGCTTGGGTTTCGTCAGGTTCCAGGTAAAGG
TCTGGAATCTATGGGTATCATCTACCCAGGTGATTCTGATATCCGTTACTCTCCATCTTTCCAGGGTCAG
GTTACCATCTCTGCTGATAAATCTATACCAACCGCTTACCTGCAGTGGTCTTCTCTGAAAGCCTCGGACA
CCGCCATTTATTACTGTGCTCGTCACGATATCGAAGTTTCGATTACTGGGGTCGTGGAACCCCTGGTCAC
35 CGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 167)

> 5a - anifrolumab dLCDR1

cagttagataaaaagagggcgcgccagcaagcggcgcgcatgccGAAATCGTTCTGACCCAGTCTCCAGGTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCTGCTTCTAGagaccatggccagtaagg
40 ccggtctctCTGGCTTGGTACCAGCAGAAACCAGGTACGGTCCACGTCTGCTGATCTACCGTGCTTCTT
CTCGTGCTACCGGTATCCCAGATCGTCTGTCTGGTTCTGGTACCGATTTACCCCTGACCATCAC
CCGTCTGGAACAGAAAGATTTCTGCTGTTTACTACTGTACGAGTACGATTCTTCTGCTATCACCTTCGGT
GGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTTCGACCATAACTTCGTATAATGTATACTATACGAAGT
TATCCTCGAGCGGTACCGAAGTTTCAGCTGGTTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGAATCTCT
45 GAAAAATCTCTTGTAAGGTTCTGGTTACATCTTACCAACTACTGGATCGCTTGGGTTTCGTCAGGTTCCA
GGTAAAGGTTCTGGAATCTATGGGTATCATCTACCCAGGTGATTCTGATATCCGTTACTCTCCATCTTTCC
AGGGTCAGGTTACCATCTCTGCTGATAAATCTATACCAACCGCTTACCTGCAGTGGTCTTCTCTGAAAGC
CTCGGACACCGCCATTTATTACTGTGCTCGTCACGATATCGAAGTTTCGATTACTGGGGTCGTGGAACC
CTGGTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO:
50 168)

> 5b - anifrolumab dLCDR2

cagttagataaaagagggcgcgagcaagcgggcgcgcatgccGAAATCGTTCTGACCCAGTCTCCAGGTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTT
CCTGGCTTGTTACCAGCAGAAACCAGGTCAGGCTCCACGTCTGCTGATCTACagagaccatggccagtaa
5 ggccggtctctGGTATCCCAGATCGTCTGTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCAC
CCGTCTGGAACCAGAAGATTTGCTGTTTACTACTGTCAGCAGTACGATTCTTCTGCTATCACCTTCGGT
GGTGGTACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGT
TATCCTCGAGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGAATCTCT
GAAAAATCTCTTGTAAGGTTCTGGTTACATCTTCACCACTACTGGATCGCTTGGGTTCGTCAGGTTCCA
10 GGTAAAGGTCTGGAATCTATGGGTATCATCTACCCAGGTGATTCTGATATCCGTTACTCTCCATCTTTCC
AGGGTCAGGTTACCATCTCTGCTGATAAATCTATCACCACCGCTTACCTGCAGTGGTCTTCTCTGAAAGC
CTCGGACACCGCCATTTATTACTGTGCTCGTCACGATATCGAAGGTTTCGATTACTGGGGTCGTGGAACC
CTGGTCACCGTCTCCTCAgctagcgggcaaaccaatcccaaaccactgctgggc (SEQ ID NO:
169)

> 5c - anifrolumab dLCDR3

cagttagataaaagagggcgcgagcaagcgggcgcgcatgccGAAATCGTTCTGACCCAGTCTCCAGGTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTT
CCTGGCTTGTTACCAGCAGAAACCAGGTCAGGCTCCACGTCTGCTGATCTACGGTGCTTCTTCTCGTGCT
20 ACCGGTATCCCAGATCGTCTGTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCACCCGTCTGG
AACCAGAAGATTTGCTGTTTACTACTGTagagaccatggccagtaaggccggtctctTTCGGTGGTGGT
ACCAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCT
CGAGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGAATCTCTGAAAT
CTCTTGTAAGGTTCTGGTTACATCTTCACCACTACTGGATCGCTTGGGTTCGTCAGGTTCCAGGTAAA
25 GGTCTGGAATCTATGGGTATCATCTACCCAGGTGATTCTGATATCCGTTACTCTCCATCTTTCCAGGGTC
AGGTTACCATCTCTGCTGATAAATCTATCACCACCGCTTACCTGCAGTGGTCTTCTCTGAAAGCCTCGGA
CACCGCCATTTATTACTGTGCTCGTCACGATATCGAAGGTTTCGATTACTGGGGTCGTGGAACCCTGGTC
ACCGTCTCCTCAgctagcgggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 170)

> 5d - anifrolumab dHCDR1

cagttagataaaagagggcgcgagcaagcgggcgcgcatgccGAAATCGTTCTGACCCAGTCTCCAGGTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTT
CCTGGCTTGTTACCAGCAGAAACCAGGTCAGGCTCCACGTCTGCTGATCTACGGTGCTTCTTCTCGTGCT
ACCGGTATCCCAGATCGTCTGTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCACCCGTCTGG
35 AACCAGAAGATTTGCTGTTTACTACTGTCAGCAGTACGATTCTTCTGCTATCACCTTCGGTGGTGGTAC
CAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCG
AGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACAGGTGAATCTCTGAAATCT
CTTGTAAGGTTCTagagaccatggccagtaaggccggtctctATCGCTTGGGTTCGTCAGGTTCCAGGT
AAAGGTCTGGAATCTATGGGTATCATCTACCCAGGTGATTCTGATATCCGTTACTCTCCATCTTTCCAGG
40 GTCAGGTTACCATCTCTGCTGATAAATCTATCACCACCGCTTACCTGCAGTGGTCTTCTCTGAAAGCCTC
GGACACCGCCATTTATTACTGTGCTCGTCACGATATCGAAGGTTTCGATTACTGGGGTCGTGGAACCCTG
GTCACCGTCTCCTCAgctagcgggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 171)

> 5e - anifrolumab dHCDR2

cagttagataaaagagggcgcgagcaagcgggcgcgcatgccGAAATCGTTCTGACCCAGTCTCCAGGTA
CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCGTGCTTCTCAGTCTGTTTCTTCTTT
CCTGGCTTGTTACCAGCAGAAACCAGGTCAGGCTCCACGTCTGCTGATCTACGGTGCTTCTTCTCGTGCT
ACCGGTATCCCAGATCGTCTGTCTGGTTCTGGTTCTGGTACCGATTTACCCCTGACCATCACCCGTCTGG
AACCAGAAGATTTGCTGTTTACTACTGTCAGCAGTACGATTCTTCTGCTATCACCTTCGGTGGTGGTAC
50 CAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCG

AGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACCAGGTGAATCTCTGAAAATCT
 CTTGTAAAGGTTCTGGTTACATCTTCACCAACTACTGGATCGCTTGGGTTTCGTCAGGTTCCAGGTAAAGG
 TCTGGAATCTATGGGTATCagagaccatggccagtaaggccggtctctCGTTACTCTCCATCTTTCCAGG
 GTCAGGTTACCATCTCTGCTGATAAATCTATCACCACCGCTTACCTGCAGTGGTCTTCTCTGAAAGCCTC
 5 GACACCGCCATTTATTACTGTGCTCGTCACGATATCGAAGGTTTCGATTACTGGGGTCGTGGAACCTG
 GTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 172)

> 5f - anifrolumab dHCDR3

cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGAAATCGTTCTGACCCAGTCTCCAGGTA
 10 CCCTGTCTCTGTCTCCAGGTGAACGTGCCACTCTGTCTTGTCTGCTTCTCAGTCTGTTTCTTCTTT
 CCTGGCTTGGTACCAGCAGAAACCAGGTACAGGCTCCACGTCTGCTGATCTACGGTGCTTCTTCTCGTGCT
 ACCGGTATCCAGATCGTCTGTCTGGTTCGTGTTCTGGTACCGATTTCACCCCTGACCATCACCCGTCTGG
 AACCAGAAAGATTTTCGTTTACTACTGTGTCAGCAGTACGATTCTTCTGCTATCACCTTCGGTGGTGGTAC
 CAAAGTTGAAATCAAATCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCCTCG
 15 AGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACCAGGTGAATCTCTGAAAATCT
 CTTGTAAAGGTTCTGGTTACATCTTCACCAACTACTGGATCGCTTGGGTTTCGTCAGGTTCCAGGTAAAGG
 TCTGGAATCTATGGGTATCATCTACCCAGGTGATTCTGATATCCGTTACTCTCCATCTTTCCAGGGTCAG
 GTTACCATCTCTGCTGATAAATCTATCACCACCGCTTACCTGCAGTGGTCTTCTCTGAAAGCCTCGGACA
 CCGCCATTTATTACTGTtagagaccatggccagtaaggccggtctctGGAACCTGGTCACCGTCTCCTCA
 20 gctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 173)

> 6 - evolocumab

cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGAATCTGCTCTGACCCAGCCAGCTTCTG
 TTTCTGGTTCCTCCAGGTCACTCTATCACCATCTCTTGTACCGGTACCTCTTCTGATGTTGGTGGTTACAA
 25 CTCTGTTTCTTGGTACCAGCAGCACCCAGGTAAAGCTCCAAAAGTATGATCTACGAAGTTTCTAACCGT
 CCATCTGGTGTCTTCTAACCGTTTCTCTGGTCTTAAATCTGGTAACACCGCTTCTCTGACCATCTCTGGTC
 TGCAGGCTGAAGATGAAGCTGATTACTACTGTAACCTCTTACACCTCTACCTCTATGGTTTTTCGGTGGTGG
 TACCAAAGTACCGTTCTGTCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGTTATCC
 TCGAGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACCAGGTGCTTCTGTTAAAG
 30 TTTCTTGTAAGCTTCTGGTTACACCTGACCTCTTACGGTATCTCTTGGGTTTCGTCAGGCTCCAGGTCA
 GGGTCTGGAATGGATGGGTTGGGTTTCTTTCTACAACGGTAACACCAACTACGCTCAGAAAGTGCAGGGT
 CGTGGTACCATGACCACCGATCCATCTACCTCTACCGCTTACATGGAAGTGCCTTCTCTGCGTTCTGACG
 ACACGGCCGTGTATTACTGTGCTCGTGGTTACGGTATGGATGTTTGGGGTCAGGGGACCACGGTCACCGT
 CTCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 174)

> 6a - evolocumab dLCDR1

cagttagataaaagagggcgcggcagcaagcggcgcgcatgccGAATCTGCTCTGACCCAGCCAGCTTCTG
 TTTCTGGTTCCTCCAGGTCACTCTATCACCATCTCTTGTACCGGTACCAgagaccatggccagtaaggccg
 gtctctGTTTCTTGGTACCAGCAGCACCCAGGTAAAGCTCCAAAAGTATGATCTACGAAGTTTCTAACCC
 40 GTCCATCTGGTGTCTTCTAACCGTTTCTCTGGTCTTAAATCTGGTAACACCGCTTCTCTGACCATCTCTGG
 TCTGCAGGCTGAAGATGAAGCTGATTACTACTGTAACCTCTTACACCTCTACCTCTATGGTTTTTCGGTGGT
 GGTACCAAAGTACCGTTCTGTCCGGAGGGTCGACCATAACTTCGTATAATGTATACTATACGAAGTTAT
 CCTCGAGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGCTGAAGTTAAAAAACCAGGTGCTTCTGTTAA
 AGTTTCTTGTAAGCTTCTGGTTACACCTGACCTCTTACGGTATCTCTTGGGTTTCGTCAGGCTCCAGGT
 45 CAGGGTCTGGAATGGATGGGTTGGGTTTCTTTCTACAACGGTAACACCAACTACGCTCAGAAAGTGCAGG
 GTCGTGGTACCATGACCACCGATCCATCTACCTCTACCGCTTACATGGAAGTGCCTTCTCTGCGTTCTGA
 CGACACGGCCGTGTATTACTGTGCTCGTGGTTACGGTATGGATGTTTGGGGTCAGGGGACCACGGTCACC
 GTCTCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 175)

50

> 6b - evolocumab dLCDR2

cagtttagataaaaagagggcgcggcagcaagcggcgcgcatgccGAATCTGCTCTGACCCAGCCAGCTTCTG
 TTTCTGGTTCTCCAGGTCAGTCTATCACCATCTCTTGTACCGGTACCTCTTCTGATGTTGGTGGTTACAA
 CTCTGTTTCTTGGTACCAGCAGCACCCAGGTAAAGCTCCAAAAGCTGATGATCTACagagaccatggccag
 5 taagggcgggtctcttGGTGTCTTCTAACCGTTTCTCTGGTTCTAAATCTGGTAACACCGCTTCTCTGACCAT
 CTCTGGTCTGCAGGCTGAAGATGAAGCTGATTACTACTGTAACCTCTACCTCTATGGTTTTTC
 GGTGGTGGTACCAAAGTACCGTTCTGTCCGGAGGGTCGACCATAAAGTTCGTATAATGTATACTATACGA
 AGTTATCCTCGAGCGGTACCGAAGTTTCAGCTGGTTCAGTCTGGTGTGAAGTTAAAAAACCAGGTGCTTC
 TGTAAAGTTTCTTGTAAAGCTTCTGGTTACACCTGACCTCTTACGGTATCTCTTGGGTTCGTCAGGCT
 10 CCAGGTACAGGTCTGGAATGGATGGGTGGGTCTTTCTTCTACAACGGTAACACCAACTACGCTCAGAAAC
 TGCAGGGTTCGTGGTACCATGACCACCGATCCATCTACCTCTACCGCTTACATGGAAGTTCGCTTCTCTGCG
 TTCTGACGACACGGCCGTGTATTACTGTGCTCGTGGTTACGGTATGGATGTTTGGGGTCAGGGGACCACG
 GTCACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 176)

15 > 6c - evolocumab dLCDR3

cagtttagataaaaagagggcgcggcagcaagcggcgcgcatgccGAATCTGCTCTGACCCAGCCAGCTTCTG
 TTTCTGGTTCTCCAGGTCAGTCTATCACCATCTCTTGTACCGGTACCTCTTCTGATGTTGGTGGTTACAA
 CTCTGTTTCTTGGTACCAGCAGCACCCAGGTAAAGCTCCAAAAGCTGATGATCTACGAAGTTTCTAACCGT
 CCATCTGGTGTCTTCTAACCGTTTCTCTGGTTCTAAATCTGGTAACACCGCTTCTCTGACCATCTCTGGTC
 20 TGCAGGCTGAAGATGAAGCTGATTACTACTGTagagaccatggccagtaagggcgggtctcttTCGGTGGT
 GGTACCAAAGTACCGTTCTGTCCGGAGGGTCGACCATAAAGTTCGTATAATGTATACTATACGAAGTTAT
 CCTCGAGCGGTACCGAAGTTTCAGCTGGTTCAGTCTGGTGTGAAGTTAAAAAACCAGGTGCTTCTGTTAA
 AGTTTCTTGTAAAGCTTCTGGTTACACCTGACCTCTTACGGTATCTCTTGGGTTCGTCAGGCTCCAGGT
 CAGGGTCTGGAATGGATGGGTGGGTCTTTCTTCTACAACGGTAACACCAACTACGCTCAGAACTCAGG
 25 GTCGTGGTACCATGACCACCGATCCATCTACCTCTACCGCTTACATGGAAGTTCGCTTCTCTGCGTTCTGA
 CGACACGGCCGTGTATTACTGTGCTCGTGGTTACGGTATGGATGTTTGGGGTCAGGGGACCACGGTCACC
 GTCTCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 177)

> 6d - evolocumab dHCDR1

30 cagtttagataaaaagagggcgcggcagcaagcggcgcgcatgccGAATCTGCTCTGACCCAGCCAGCTTCTG
 TTTCTGGTTCTCCAGGTCAGTCTATCACCATCTCTTGTACCGGTACCTCTTCTGATGTTGGTGGTTACAA
 CTCTGTTTCTTGGTACCAGCAGCACCCAGGTAAAGCTCCAAAAGCTGATGATCTACGAAGTTTCTAACCGT
 CCATCTGGTGTCTTCTAACCGTTTCTCTGGTTCTAAATCTGGTAACACCGCTTCTCTGACCATCTCTGGTC
 TGCAGGCTGAAGATGAAGCTGATTACTACTGTAACCTCTTACACCTCTACCTCTATGGTTTTTCGGTGGTGG
 35 TACCAAAGTACCGTTCTGTCCGGAGGGTCGACCATAAAGTTCGTATAATGTATACTATACGAAGTTATCC
 TCGAGCGGTACCGAAGTTTCAGCTGGTTCAGTCTGGTGTGAAGTTAAAAAACCAGGTGCTTCTGTTAAAG
 TTTCTTGTAAAGCTTCTTagagaccatggccagtaagggcgggtctcttATCTCTTGGGTTCGTCAGGCTCCA
 GGTCAGGGTCTGGAATGGATGGGTGGGTCTTTCTTCTACAACGGTAACACCAACTACGCTCAGAACTGC
 AGGGTCGTGGTACCATGACCACCGATCCATCTACCTCTACCGCTTACATGGAAGTTCGCTTCTCTGCGTTCT
 40 TGACGACACGGCCGTGTATTACTGTGCTCGTGGTTACGGTATGGATGTTTGGGGTCAGGGGACCACGGTC
 ACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 178)

> 6e - evolocumab dHCDR2

45 cagtttagataaaaagagggcgcggcagcaagcggcgcgcatgccGAATCTGCTCTGACCCAGCCAGCTTCTG
 TTTCTGGTTCTCCAGGTCAGTCTATCACCATCTCTTGTACCGGTACCTCTTCTGATGTTGGTGGTTACAA
 CTCTGTTTCTTGGTACCAGCAGCACCCAGGTAAAGCTCCAAAAGCTGATGATCTACGAAGTTTCTAACCGT
 CCATCTGGTGTCTTCTAACCGTTTCTCTGGTTCTAAATCTGGTAACACCGCTTCTCTGACCATCTCTGGTC
 TGCAGGCTGAAGATGAAGCTGATTACTACTGTAACCTCTTACACCTCTACCTCTATGGTTTTTCGGTGGTGG
 TACCAAAGTACCGTTCTGTCCGGAGGGTCGACCATAAAGTTCGTATAATGTATACTATACGAAGTTATCC
 50 TCGAGCGGTACCGAAGTTTCAGCTGGTTCAGTCTGGTGTGAAGTTAAAAAACCAGGTGCTTCTGTTAAAG
 TTTCTTGTAAAGCTTCTGGTTACACCTGACCTCTTACGGTATCTCTTGGGTTCGTCAGGCTCCAGGTCA

GGGTCTGGAATGGATGGGTTGGagagaccatggccagtaaggccggtctctAACTACGCTCAGAACTGC
 AGGGTCGTGGTACCATGACCACCGATCCATCTACCTCTACCGCTTACATGGAAGTTCGCTTCTGCGTTC
 TGACGACACGGCCGTGTATTACTGTGCTCGTGGTTACGGTATGGATGTTTGGGGTCAGGGGACCACGGTC
 ACCGTCTCCTCAgctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 179)

5

> 6f - evolocumab dHCDR3

cagttagataaaagagggcgccgagcaagcggcgccgcatgccGAATCTGCTCTGACCCAGCCAGCTTCTG
 TTTCTGGTCTCCAGGTCAGTCTATCACCATCTCTTGTACCGGTACCTCTTCTGATGTTGGTGGTTACAA
 CTCTGTTTCTTGGTACCAGCAGCACCCAGGTAAAGCTCCAAAAGTATGATGATCTACGAAGTTTCTAACCCT
 10 CCATCTGGTGTCTTCTAACCCTTCTCTGGTCTAAATCTGGTAACACCGCTTCTCTGACCATCTCTGGTC
 TGCAGGCTGAAGATGAAGCTGATTACTACTGTAACCTTACACCTCTACCTCTATGGTTTTCGGTGGTGG
 TACCAAAGTACCGTTCTGTCCGGAGGGTCGACCATAAATTCGTATAATGTATACTATACGAAGTTATCC
 TCGAGCGGTACCGAAGTTCAGCTGGTTCAGTCTGGTGTGAAGTTAAAAAACCAGGTGCTTCTGTAAAG
 TTTCTTGTAAGCTTCTGGTTACACCTGACCTCTTACGGTATCTCTTGGGTTTCGTAGGCTCCAGGTCA
 15 GGGTCTGGAATGGATGGGTTGGGTTTCTTCTACAACGGTAACACCAACTACGCTCAGAAAGTGCAGGGT
 CGTGGTACCATGACCACCGATCCATCTACCTCTACCGCTTACATGGAAGTGCCTTCTGCGTTCTGACG
 ACACGGCCGTGTATTACTGTtagagaccatggccagtaaggccggtctctGGGACCACGGTCACCGTCTCC
 TCagctagcggcaaaccaatcccaaaccactgctgggc (SEQ ID NO: 180)

20 Sequences 1-6 refer to the coding sequence of the exemplary scaffolds as indicated
 and sequences 1a-1f, 2a-2f, 3a-3f, 4a-4f, 5a-5f, and 6a-6f refer to sequences in which cloning
 sites were inserted flanking the corresponding CDRs of each exemplary scaffold as indicated.

Using LCDR1 of library 1 as an example, the LCDR1 scaffold vector was digested
 with BsaI and SfiI leaving a gap at the site of LCDR1, as shown in figure 24. The cleaved
 25 vector and the collection of LCDR1 oligonucleotides were then transformed into yeast. Inside
 yeast cells, homologous recombination between the vector and the LCDR1 oligonucleotides
 results in insertion of the LCDR1 oligonucleotides into the LCDR1 scaffold vector. The
 entire population of LCDR1 yeast scaffold vectors carrying each of the LCDR1
 oligonucleotides identified above constitutes a LCDR1 yeast display library (Figure 25). For
 30 the LCDR1 scaffold vector, all portions of the VH and VL domains are constant except for
 the LCDR1 that is assessed. Selection for functional LCDR1 was carried out by sorting for
 display (*i.e.*, expression). After the display of the scFv is induced, the yeast cells are stained
 with fluorescent-labelled antibody that detects scFv display (using the mAb recognizing the
 SV5 tag). For CDR1-2 libraries, the populations are analyzed by flow cytometry and sorted
 35 by fluorescence activated cell sorting by gating the top 2% most fluorescent cells among the
 positive population – this ensures the enrichment for CDRs that promote high levels of
 display. In figures 22A-L, the population after this stringent 2% sorting is indicated as
 “Replicated natural after filtering”. At least a 10-fold number of cells is sorted as compared
 to the theoretical diversity to ensure recovery of all possible clones. For LCDR3, after scFv

display induction and staining with fluorescent-labelled antibody that detects the scFv display (SV5), the positive population (scFv displaying) is purified using MACS (magnetic-activated cell sorting) employing magnetic nanoparticles that recognize the primary antibody (SV5) used. The higher the level of the scFv display, the higher the probability of the cell binding the nanoparticles – especially when competition is employed by having a number of cells that far exceeds the binding capacity of the nanoparticles, thus, enriching for well displaying sequences, analogously to fluorescence activated cell sorting of the most fluorescent 2% for the CDR1-2 described above. The MACS technique is preferentially employed for LCDR3 due to the capacity of purifying a large number of cells in a short span of time, since the theoretical diversity of the LCDR3 is several fold higher than CDR1-2.

Although scFvs are used in this example for yeast display, the format of CDR-specific scaffold vectors can be VH, VL, scFv, Fab or full-length immunoglobulin, the requirement being that display can be carried out. By sorting libraries of single CDRs cloned into well folded scaffolds, those CDRs that contain stop codons, frameshifts, or are poorly expressed or polyreactive may be eliminated. Effective display on the yeast surface has been previously correlated with improved stability and folding in diverse proteins (Cherf, G. M. and J. R. Cochran (2015). "Applications of Yeast Surface Display for Protein Engineering." *Methods Mol Biol* **1319**: 155-175; Pavoov, T. V., et al., (2012). "An enhanced approach for engineering thermally stable proteins using yeast display." *Protein engineering, design & selection : PEDS* **25**(10): 625-630; Pepper, L. R., et al., (2008). "A decade of yeast surface display technology: where are we now?" *Comb Chem High Throughput Screen* **11**(2): 127-134; Xu, L., et al., (2013). "Rapid optimization and prototyping for therapeutic antibody-like molecules." *MAbs* **5**(2): 237-254.). After each CDR library is sorted, a collection of well expressed, non-polyreactive CDRs is obtained by isolating DNA from yeast cells expressing well folded CDRs. Schematic illustrations of how functional CDR libraries are cloned and sorted are shown in Figures 23-25. The non-filtered and filtered libraries for the 6 different scaffolds are show in Figures 25B-G: the populations are analyzed by flow cytometry and expression levels (x axis) are represented as a histogram. The analysis shows a clear improvement after enrichment (top 2% by fluorescence activated cell sorting for CDRs 1-2, and magnetic activated cell sorting for LCDR3), with most libraries showing display levels exceeding that of the original clinical candidate from which the scaffolds were generated.

In the examples provided here, we have sorted for improved expression levels. However, a similar approach can be taken using any selective method that distinguishes yeast displaying antibodies with desirable properties (e.g. high expression, low polyreactivity, good developability) from yeast displaying antibodies with undesirable properties (e.g. low expression, high polyreactivity, poor developability). For example, to select antibodies with reduced polyreactivity, clones displaying antibodies (scFvs, Fabs, IgGs, or other antibody fragment) that do not bind to polyspecificity reagents would be selected. Examples of such polyspecificity reagents include those described in (Hotzel, I. *et al.* A strategy for risk mitigation of antibodies with fast clearance. *MAbs* **4**, 753-760, doi:10.4161/mabs.22189 (2012); Xu, Y. *et al.* Addressing polyspecificity of antibodies selected from an in vitro yeast presentation system: a FACS-based, high-throughput selection and analytical tool. *Protein Eng Des Sel* **26**, 663-670, doi:10.1093/protein/gzt047 (2013); Kelly, R. L. *et al.* Chaperone proteins as single component reagents to assess antibody nonspecificity. *MAbs* **9**, 1036-1040, doi:10.1080/19420862.2017.1356529 (2017).)

Example 8: Assembly of full length scFv library

Once each individual CDR library was screened for expression and developability, the VH and VL CDRs were assembled into full length VH or VL domains. Individual CDRs in VL and VH were amplified with the flanking sequence using the primers described in Table 7 and Table 8. This was carried out as illustrated in figures 29-31, with the VH and VL first assembled from their constituent CDR fragments and flanking framework regions, and then combined into complete scFvs or Fabs. An alternative approach would be to assemble each full length VH or VL within the context of their non-modified VL or VH partner chains as illustrated in Figures 26-28. This alternative approach would allow sorting (by FACS or MACS) for VH or VL libraries that are functional within the context of their unmodified partner chains. However, we found that the direct assembly of complete scFvs directly from libraries of filtered CDRs led to highly functional scFv libraries without the need for this intermediate step, as illustrated in figure 32, which shows that the peak display level of scFvs derived from the clinical antibodies used as scaffolds is similar to that for the corresponding libraries, except that the distribution of library display levels is broader than that of the clinical candidate antibodies used as scaffolds, and includes some scFvs that are displayed better than the parental clinical scaffold scFv. The functional antibody library can be

assembled within the context of different display vectors, including phage, yeast or mammalian display vectors.

5

Example 9: Cloning into a phage display vector (pDAN5)

Once the scFv or Fab libraries were assembled, they were ligated into a phage display vector, such as pDAN5 to explore their functionality. This vector contains a cloning site upstream of the g3 of the filamentous phage, comprised by the restriction enzyme sites for BssHII and NheI. The scFv/Fabs created in Example 8 were amplified with flanking primers containing the BssHII restriction site upstream of the light chain and the NheI downstream of the heavy chain. The PCR product was then digested with the same enzymes to generate cohesive ends. The pDAN5 plasmid was cultivated in *E. coli*, extracted by alkaline lysis and purified by cesium chloride/ethidium bromide gradient. The plasmid was digested with the same enzymes and the backbone purified by agarose gel electrophoresis extraction followed by chromatography to remove contaminants. The backbone was ligated to the scFv/Fab library using T4 DNA ligase overnight at 16°C. The ligation was purified and electrotransformed into electrocompetent *E. coli* TG1 cells. The transformed cells were plated out on agar plates containing carbenicillin and glucose to select for bacteria that received the plasmid. Analysis of the scFv libraries by PacBio sequencing revealed >90% open reading frames (Table 9), and essentially no clone duplication (Table 10).

25

Attorney Docket No.: 103073-629979 (S2061.70000WO00)

Table 9. Percentage of open reading frames in scFv libraries as assessed by PacBio sequencing.

Library	sequences analyzed	correct frame	ORFs	ORF %
Lib1	6,510	6,103	6,050	93%
Lib2	5,699	5,403	5,355	94%
Lib3	7,012	6,623	6,583	94%
Lib4	9,168	8,660	8,613	94%
Lib5	7,640	7,282	7,242	95%
Lib6	5,378	5,027	5,005	93%

Table 10. Analysis of library diversity by PacBio.

Library	No. reads	Full-Length		HCDR3+LCDR3		HCDR3		LCDR3		HCDR2		LCDR2		HCDR1		LCDR1	
		No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
		Unique reads															
1	6000	5998	99.97	5998	99.97	5851	97.52	5679	94.65	1786	29.77	2224	37.07	2158	35.97	1328	22.13
2	5265	5258	99.87	5258	99.87	4740	90.03	4675	88.79	1927	36.60	1796	34.11	2267	43.06	481	9.14
3	6496	6493	99.95	6493	99.95	6300	96.98	5828	89.72	3077	47.37	1329	20.46	3676	56.59	438	6.74
4	8423	8420	99.96	8420	99.96	8137	96.60	7965	94.56	2538	30.13	1355	16.09	2525	29.98	1399	16.61
5	7030	7029	99.99	7015	99.79	6457	91.85	6263	89.09	1780	25.32	1873	26.64	2684	38.18	2090	29.73
6	4899	4898	99.98	4898	99.98	4735	96.65	4725	96.45	1639	33.46	1148	23.43	1833	37.42	1545	31.54

Example 10: Creation of bacteriophage particles, including western blot

The transformed bacteria were cultivated in a shaking flask containing liquid 2xYT media + carbenicillin + glucose (the glucose is to inhibit scFv/Fab expression) at 37°C until an OD_{600nm} of 0.5 was reached. The bacteria were superinfected with M13KO7 helper
 5 phage (at a multiplicity of infection of 5) for 30 min at 37°C without shaking. The bacteria were centrifuged, the media removed and replaced with 2xYT media + carbenicillin + kanamycin and cultivated for 16h at 25°C in a shaker incubator.

To recover the phage particles, the cultures were centrifuged to separate the bacteria and the supernatant, where the phage is found. The supernatant was mixed with a 20% PEG
 10 8000 + 2.5 M NaCl solution at a 5:1 ratio. This causes the phage to precipitate, allowing them to be harvested by centrifugation. The supernatant was discarded and the phage pellet was resuspended in a PBS solution. The display of the scFv/Fab by the phage was assessed by SDS-PAGE + western blot using an antibody that specifically recognizes the expression tag (SV5) as show in Figure 33.

Example 11: Antibody selection by combined phage and yeast display using the library

After construction and phage particle production, the library was screened against targets of interest. While selection can be carried out using phage display alone (Sblattero, D. & Bradbury, A. Exploiting recombination in single bacteria to make large phage antibody
 20 libraries. *Nat Biotechnol* **18**, 75-80 (2000)), we prefer to combine phage and yeast display technologies (Ferrara, F. *et al.* Using phage and yeast display to select hundreds of monoclonal antibodies: application to antigen 85, a tuberculosis biomarker. *PLoS One* **7**, e49535 (2012)). 10¹² phage particles displaying the scFv library were used in two rounds of selection against the biotinylated recombinant human antigens of clinical interest –
 25 interleukin 6, interferon alpha, and GM-CSF - using the Kingfisher magnetic bead system: 2x10⁷ streptavidin-conjugated magnetic beads (Dynabeads M-280) coated with the biotinylated proteins (100-400 nM) were washed, coated with the antigen, incubated with the phage particles and washed again to remove non-binders. Phage particles were then eluted by reducing the pH and infecting F⁺ pilus-carrying bacteria (Ominmax-2T1, Thermo Fisher
 30 Scientific). The phages were propagated, and the selection cycle reiterated. After two rounds of phage enrichment, the scFvs were PCR amplified and transferred to an N-terminal yeast display system by homologous recombination (pSpec yeast display vector), in which the scFv

is displayed fused to the N terminus of Aga-2. The transformed yeast was then induced for scFv display by adding galactose to the culture media. The induced yeast minilibraries were then used for another two to three rounds of enrichment against the biotinylated recombinant human antigens by fluorescence activated cell sorting. Up to 10,000 yeast cells with positive antigen binding signal were sorted and propagated in each round. Target concentration in the first sorting round was 100nM, reduced to 10nM and then 1nM. After these rounds of phage and yeast sorting enrichment the recovered populations was analyzed by flow cytometry to test for binding against the antigens in decreasing concentration of antigen and in the absence of the antigen to check for non-specific binding to secondary reagents (Figure 34). Results show that the library can successfully yield high affinity binders to all antigens tested.

Example 12: Affinity Determination of Selected Antibodies

Affinity determination of antibodies selected from the naïve library using the phage + yeast display protocol described in Example 11, was performed following the approaches described herein. Binding affinity of the antibody variants thus obtained to various targets (*e.g.*, GM-CSF, IFN- α 2A and IL-6) was examined using a Cytiva LSA machine. Briefly, supernatants from yeast expressing scFv-Fc fusions from selections against GM-CSF, IFN- α 2A and IL-6 were immobilized on a Cytiva LSA HC200M chip with anti-human Fc. The chips were activated with 1:1:1 100 mM MES pH 5.5, 100 mM S-NHS, 400 mM EDC (all reconstituted in MES 5.5), and 100 μ L of each were mixed in a vial immediately before running the assay. The polyclonal goat anti-human IgG was immobilized for 10-minute at 50 μ g/mL followed by 7-minute deactivation with 1 M Ethanolamine pH 8.5.

The scFv-Fc supernatants were diluted two or three fold into HBSTE buffer and cycled for 12 minutes across the anti-Fc surface. Antigens were tested in a three-fold dilution series starting at 6 nM for IFN-2A and 167 nM for IL-6/GM-SCF. The antigen samples were tested from lowest to highest concentration. Figure 37.

Data was processed using a floated Rmax parameter for the IFN-2A and GM-SCF clones that did not dissociate fully between binding cycles; some of the data were also fit using a bulk shift parameter.

As shown in Figures 38A-38C, the affinities of antibodies selected directly from the library, constructed as described in Examples 1-10, are shown to be extremely potent, with many subnanomolar antibodies having been selected.

Antibodies to additional targets, including PDGF, TGFBR2, and TGFBR3, were explored using this approach and similar results were observed. Figure 39.

Example 13: Antibody Maturation

5 To select an optimized VL and VH pair and assemble the CDRs into a mature antibody, the following approach can be used. First, the VH is kept in unmodified form while the VLs in the LCDR libraries are shuffled. The remaining functional VLs are assembled with the unmodified VH and the formed antibodies are tested for functionality (Figure 35). Likewise, the VL is kept unmodified and the VHs comprising synthetic HCDR1/2 and natural
10 HCDR3 in the VH library is shuffled. Each of the remaining functional VH can be assembled with the unmodified VL and the newly formed antibodies are tested for their functions (Figure 36).

OTHER EMBODIMENTS

15 All of the features disclosed in this specification may be combined in any combination. Each feature disclosed in this specification may be replaced by an alternative feature serving the same, equivalent, or similar purpose. Thus, unless expressly stated otherwise, each feature disclosed is only an example of a generic series of equivalent or similar features.

20 From the above description, one skilled in the art can easily ascertain the essential characteristics of the present invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, other embodiments are also within the claims.

EQUIVALENTS

25 While several inventive embodiments have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and/or structures for performing the function and/or obtaining the results and/or one or more of the advantages described herein, and each of such variations and/or modifications is deemed to
30 be within the scope of the inventive embodiments described herein. More generally, those skilled in the art will readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that the actual parameters,

dimensions, materials, and/or configurations will depend upon the specific application or applications for which the inventive teachings is/are used. Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation, many equivalents to the specific inventive embodiments described herein. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, inventive embodiments may be practiced otherwise than as specifically described and claimed. Inventive embodiments of the present disclosure are directed to each individual feature, system, article, material, kit, and/or method described herein. In addition, any combination of two or more such features, systems, articles, materials, kits, and/or methods, if such features, systems, articles, materials, kits, and/or methods are not mutually inconsistent, is included within the inventive scope of the present disclosure.

All definitions, as defined and used herein, should be understood to control over dictionary definitions, definitions in documents incorporated by reference, and/or ordinary meanings of the defined terms.

All references, patents and patent applications disclosed herein are incorporated by reference with respect to the subject matter for which each is cited, which in some cases may encompass the entirety of the document.

The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.”

The phrase “and/or,” as used herein in the specification and in the claims, should be understood to mean “either or both” of the elements so conjoined, i.e., elements that are conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with “and/or” should be construed in the same fashion, i.e., “one or more” of the elements so conjoined. Other elements may optionally be present other than the elements specifically identified by the “and/or” clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to “A and/or B”, when used in conjunction with open-ended language such as “comprising” can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

As used herein in the specification and in the claims, “or” should be understood to have the same meaning as “and/or” as defined above. For example, when separating items in a list, “or” or “and/or” shall be interpreted as being inclusive, i.e., the inclusion of at least one, but also including more than one, of a number or list of elements, and, optionally,
5 additional unlisted items. Only terms clearly indicated to the contrary, such as “only one of” or “exactly one of,” or, when used in the claims, “consisting of,” will refer to the inclusion of exactly one element of a number or list of elements. In general, the term “or” as used herein shall only be interpreted as indicating exclusive alternatives (i.e. “one or the other but not both”) when preceded by terms of exclusivity, such as “either,” “one of,” “only one of,” or
10 “exactly one of.” “Consisting essentially of,” when used in the claims, shall have its ordinary meaning as used in the field of patent law.

As used herein in the specification and in the claims, the phrase “at least one,” in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily
15 including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This definition also allows that elements may optionally be present other than the elements specifically identified within the list of elements to which the phrase “at least one” refers, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, “at least
20 one of A and B” (or, equivalently, “at least one of A or B,” or, equivalently “at least one of A and/or B”) can refer, in one embodiment, to at least one, optionally including more than one, A, with no B present (and optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one,
25 optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method
30 are recited.

What Is Claimed Is:

1. A method for producing an antibody library, comprising:
providing (a) a first plurality of nucleic acids encoding a population of
5 naturally-occurring antibody heavy chain complementary determining region 1 (CDR1)
fragments, and/or (b) a second plurality of nucleic acids encoding a population of naturally-
occurring antibody heavy chain complementary determining region 2 (CDR2) fragments; and
inserting the first plurality of nucleic acids and/or the second plurality of
nucleic acids into the CDR1 region and/or the CDR2 region, respectively, of an antibody
10 heavy chain variable domain gene, thereby producing an antibody library.

2. The method of claim 1, wherein the antibody library comprises members, in
each of which the heavy chain CDR1 and heavy chain CDR2 are not from the same naturally-
occurring antibody.

3. The method of claim 2, wherein at least 50% of the members in the antibody
library contain heavy chain CDR1 and heavy chain CDR2 that are not from the same
naturally-occurring antibody.

4. The method of any one of claims 1-3, further comprising providing a third
plurality of nucleic acids encoding a population of naturally-occurring heavy chain
complementary determining region 3 (CDR3) fragments, and inserting the third plurality of
nucleic acids into the CDR3 region of the heavy chain variable region gene.

5. The method of any one of claims 1-4, wherein the antibody library comprises
members, each of which comprises a heavy chain CDR1, a heavy chain CDR2, and a heavy
chain CDR3 that are not from the same naturally-occurring antibody.

6. The method of claim 5, wherein at least 50% of the members in the antibody
library contain heavy chain CDR1, heavy chain CDR2, and/or heavy chain CDR3 that are not
from the same naturally-occurring antibody.

7. The method of any one of claims 1-6, wherein the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising one or more of:

- (i) a glycosylation site,
- (ii) a deamidation site,
- (iii) an isomerization site,
- (iv) unpaired cysteine,
- (v) net charge greater than 1,
- (vi) a tripeptide motif containing at least two aromatic residues,
- (vii) a motif that promotes aggregation,
- (viii) a polyspecificity site;
- (ix) a protease sensitive site,
- (x) an integrin binding site,
- (xi) a lysine glycation site,
- (xii) a metal catalyzed fragmentation site,
- (xiii) a polyspecificity aggregation site; and
- (xiv) a streptavidin binding motif.

8. The method of claim 7, wherein the glycosylation site of (i) comprises the motif NXS, NXT, or NXC, in which X represents any naturally-occurring amino acid residue except for proline.

9. The method of claim 7 or claim 8, wherein the deamidation site of (ii) comprises the motif of NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG.

10. The method of any one of claims 7-9, wherein the isomerization site of (iii) comprises the motif of DT, DH, DS, DG, or DD.

11. The method of any one of claims 7-10, wherein the tripeptide of (vi) is HYF or HWH.

12. The method of any one of claims 7-11, wherein the motif that promotes aggregation of (vii) comprises the motif of FHW.

13. The method of any one of claims 7-12, wherein the polyspecificity site of (viii) comprises the motif GG, GGG, RR, VG, W, WV, WW, WWW, YY, or WXW, in which X represents any amino acid residue.

14. The method of any one of claims 7-13, wherein the protease cleavage site of (ix) comprises the motif of DX, in which X is P, G, S, V, Y, F, Q, K, L, or D.

15. The method of any one of claims 7-14, wherein the integrin binding site of (x) comprises RGD, RYD, LDV, or KGD.

16. The method of any one of claims 7-15, wherein the lysine glycation site of (xi) comprises KE, EK, or ED.

17. The method of any one of claims 7-16, wherein the metal catalyzed fragmentation site of (xii) comprises the motif of HS, SH, KT, HXS, or SXH, in which X represents any amino acid residue.

18. The method of any one of claims 7-17, wherein the polyspecificity aggregation site of (xiii) comprises a motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is selected from the group consisting of F, I, L, V, W and Y.

19. The method of any one of claims 7-18, wherein the streptavidin binding motif of (xiv) comprises the motif HPQ, EPDW (SEQ ID NO: 117), PWXWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), or PWPWL (SEQ ID NO: 120).

20. The method of any one of claims 1-19, wherein the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising at least two of (i)-(xiv).

5

21. The method of claim 20, wherein the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising (i)-(ix).

10

22. The method of claim 21, wherein the population of antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragment is further free of members comprising one or more of (x)-(xiv).

15

23. The method of any one of claims 7-22, wherein the first plurality of nucleic acid is produced by a process comprising:

(a) obtaining amino acid sequences of the heavy chain CDR1 regions of a population of naturally-occurring antibodies,

20

(b) excluding from (a) the heavy chain CDR1 amino acid sequences that comprise one or more of (i) to (xiv) to obtain liability-free heavy chain CDR1 sequences, and

(c) synthesizing the first plurality of nucleic acids that encode the liability-free heavy chain CDR1 regions.

25

24. The method of any one of claims 1-23, wherein the second plurality of nucleic acid is produced by a process comprising:

(a) obtaining amino acid sequences of the heavy chain CDR2 regions of a population of naturally-occurring antibodies,

(b) excluding from (a) the heavy chain CDR2 amino acid sequences that comprise one or more of (i) to (xiv) to obtain liability-free heavy chain CDR2 sequences, and

30

(c) synthesizing the second plurality of nucleic acids that encode the liability-free heavy chain CDR2 regions.

25. The method of any one of claims 3-24, wherein the third plurality of nucleic acids is produced by a process comprising:

(a) obtaining amino acid sequences of the heavy chain CDR3 regions of a population of naturally-occurring antibodies,

5 (b) excluding from (a) the heavy chain CDR3 amino acid sequences that comprise one or more of (i) to (xiv) to obtain liability-free heavy chain CDR3 sequences, and

(c) synthesizing the second plurality of nucleic acids that encode the liability-free heavy chain CDR3 regions.

10 26. The method of claims 20-25, wherein the process for producing the first, second, and/or third plurality of nucleic acids further comprises, after step (c), isolating functional members from the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions.

27. The method of claim 26, wherein the functional members of the liability-free heavy chain CDR1, CDR2, and/or CDR3 are isolated by expressing antibodies comprising the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions in host cells in a manner that the antibodies are displayed on surface of the host cells, isolating the antibodies that display on the host cells, and identifying the CDR1, CDR2, and/or CDR3 regions in the displayed antibodies, which are functional members of the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions.

28. The method of claim 27, wherein the isolated antibodies that display on the host cells show reduced binding to polyspecificity reagents.

25 29. The method of claim 26, wherein the functional members of the liability-free heavy chain CDR1, CDR2, and/or CDR3 are isolated by expressing antibodies comprising the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions in fusion with a folding reporter, which optionally is β lactamase or green fluorescent protein, or fragments thereof, to obtain members with improved folding.

30

30. The method of any one of claims 4-29, wherein the population of human antibody heavy chain CDR3 fragments is derived from B cells.

31. The method of any one of claims 1-30, wherein the antibody heavy chain variable region gene is a human heavy chain variable region gene, which optionally is selected from the group consisting of VH1-24, VH2-70, VH3-7, VH4-30-4, VH5-51, VH1-18, VH1-69, VH3-23, VH5-10-1, VH3-9, and VH3-11.

32. The method of claim 31, wherein the human antibody heavy chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, and evoculumab.

33. The method of claim 31, wherein the human antibody heavy chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abritumumab, adalimumab, alemtuzumab, alirocumab, bapineuzumab, benralizumab, brodalumab, canakinumab, certolizumab, clazakizumab, dacetuzumab, daclizumab, daratumumab, eculizumab, efalizumab, elotuzumab, epratuzumab, farletuzumab, fasinumab, ficlatuzumab, fletikumab, fresolimumab, fulranumab, gevokizumab, ibalizumab, lintuzumab, matuzumab, mavrilimumab, mogamulizumab, motavizumab, natalizumab, nivolumab, obinutuzumab, ofatumumab, olokizumab, omalizumab, onartuzumab, otelixizumab, otlertuzumab, palivizumab, panitumumab, panobacumab, pertuzumab, pinatuzumab, polatuzumab, radretumab, ramucirumab, reslizumab, romosozumab, sarilumab, secukinumab, sifalimumab, tabalumab, tigatuzumab, tildrakizumab, tocilizumab, tovetumab, trastuzumab, vedolizumab, veltuzumab, zalutumumab, and zanolimumab.

34. The method of any one of claims 1-33, wherein the method further comprises:
providing a fourth plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 1 (CDR1) fragments, a fifth plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 2 (CDR2) fragments, and/or a sixth plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 3 (CDR3) fragments, and

inserting the fourth plurality of nucleic acids, the fifth plurality of nucleic acids, and/or the sixth plurality of nucleic acids into the CDR1 region, the CDR2 region, and the CDR3 region, respectively, of an antibody light chain variable domain gene.

35. The method of claim 33, wherein the population of naturally-occurring antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising one or more of:

- (i) a glycosylation site,
- (ii) a deamidation site,
- (iii) an isomerization site,
- (iv) unpaired cysteine,
- (v) net charge greater than 1,
- (vi) a tripeptide motif containing at least two aromatic residues,
- (vii) a motif that promotes aggregation,
- (viii) a polyspecificity site;
- (ix) a protease sensitive site,
- (x) an integrin binding site,
- (xi) a lysine glycation site,
- (xii) a metal catalyzed fragmentation site,
- (xiii) a polyspecificity aggregation site; and/or
- (xiv) a streptavidin binding motif.

36. The method of claim 35, wherein the glycosylation site of (i) comprises the motif NXS, NXT, or NXC, in which X represents any naturally-occurring amino acid residue except for proline.,

37. The method of claim 35 or claim 36, wherein the deamidation site of (ii) comprises the motif of NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG.

38. The method of any one of claims 35-37, wherein the isomerization site of (iii) comprises the motif of DT, DH, DS, DG, or DD.

39. The method of any one of claims 35-38, wherein the tripeptide of (vi) is HYF or HWH.

40. The method of any one of claims 35-39, wherein the motif that promotes aggregation of (vii) comprises the motif of FHW.

41. The method of any one of claims 35-40, wherein the polyspecificity site of (viii) comprises the motif GG, GGG, RR, VG, W, WV, WW, WWW, YY, or WXW, in which X represents any amino acid residue.

42. The method of any one of claims 35-41, wherein the protease cleavage site of (ix) comprises the motif of DX, in which X is P, G, S, V, Y, F, Q, K, L, or D.

43. The method of any one of claims 35-42, wherein the integrin binding site of (x) comprises RGD, RYD, LDV, or KGD.

44. The method of any one of claims 35-43, wherein the lysine glycation site of (xi) comprises KE, EK, or ED.

45. The method of any one of claims 35-44, wherein the metal catalyzed fragmentation site of (xii) comprises the motif of HS, SH, KT, HXS, or SXH, in which X represents any amino acid residue.

46. The method of any one of claims 35-45, wherein the polyspecificity aggregation site of (xiii) comprises a motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is selected from the group consisting of F, I, L, V, W and Y.

47. The method of any one of claims 35-46, wherein the streptavidin binding motif of (xiv) comprises the motif HPQ, EPDW (SEQ ID NO: 117), PWXWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), or PWPWL (SEQ ID NO: 120).

48. The method of any one of claims 35-47, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising at least two of (i)-(xiv).

5

49. The method of claim 48, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising (i)-(ix).

10

50. The method of claim 49, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragment is further free of members comprising one or more of (x)-(xiv).

15

51. The method of any one of claims 35-50, wherein the fourth plurality of nucleic acids, the fifth plurality of nucleic acids, and/or the sixth plurality of nucleic acids is produced by a process comprising:

(a) obtaining amino acid sequences of the light chain CDR1, CDR2, and/or CDR3 regions of a population of naturally-occurring antibodies,

20

(b) excluding from (a) the light chain CDR1, CDR2, and/or CDR3 amino acid sequences that comprise one or more of (i) to (xiv) to obtain liability-free light chain CDR1, CDR2, and/or CDR3 sequences, and

(c) synthesizing the fourth plurality of nucleic acids, the fifth plurality of the nucleic acids, and/or the sixth plurality of nucleic acids that encode the liability-free light chain

25

CDR1, CDR2, and/or CDR3 regions.

52. The method of claim 51, wherein the process for producing the fourth, fifth, and/or sixth plurality of nucleic acids further comprises, after step (c), isolating function members from the liability-free light chain CDR1, CDR2, and/or CDR3 regions.

30

53. The method of claim 52, wherein the functional members of the liability-free light chain CDR1, CDR2, and/or CDR3 are isolated by expressing antibodies comprising the

liability-free light chain CDR1, CDR2, and/or CDR3 regions in host cells in a manner that the antibodies are displayed on surface of the host cells, isolating the antibodies that display on the host cells, and identifying the CDR1, CDR2, and/or CDR3 regions in the displayed antibodies, which are functional members of the liability-free light chain CDR1, CDR2,
5 and/or CDR3 regions.

54. The method of claim 53, wherein the functional members show reduced binding to polyspecificity reagents.

10 55. The method of claim 54, wherein the functional members of the liability-free heavy chain CDR1, CDR2, and/or CDR3 are isolated by expressing antibodies comprising the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions in fusion with a folding reporter, which optionally is β lactamase or green fluorescent protein, or fragments thereof, to obtain members with improved folding.

15 56. The method of any one of claims 35-55, wherein the antibody library comprises members, each of which comprises a light chain CDR1, a light chain CDR2, and/or a light chain CDR3 that are not found in the same naturally-occurring antibody.

20 57. The method of claim 56, wherein at least 50% of the members in the antibody library comprise light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not found in the same naturally-occurring antibody.

25 58. The method of any one of claims 35-57, wherein the antibody light chain variable domain gene is a human antibody light chain variable domain gene, which optionally is selected from the group consisting of K1-12, K4-1, K2D-29, K3-11, K3-20, and L2-14.

30 59. The method of claim 58, wherein the human antibody light chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, and evoculumab.

60. The method of claim 58, wherein the human antibody light chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abrituzumab, adalimumab, alemtuzumab, alirocumab, bapineuzumab, benralizumab, brodalumab, canakinumab, certolizumab, clazakizumab, dacetuzumab, daclizumab, 5 daratumumab, eculizumab, efalizumab, elotuzumab, epratuzumab, farletuzumab, fasinumab, ficlatuzumab, fletikumab, fresolimumab, fulranumab, gevokizumab, ibalizumab, lintuzumab, matuzumab, mavrilimumab, mogamulizumab, motavizumab, natalizumab, nivolumab, obinutuzumab, ofatumumab, olokizumab, omalizumab, onartuzumab, otelixizumab, otlertuzumab, palivizumab, panitumumab, panobacumab, pertuzumab, pinatuzumab, 10 polatuzumab, radretumab, ramucirumab, reslizumab, romosozumab, sarilumab, secukinumab, sifalimumab, tabalumab, tigatuzumab, tildrakizumab, tocilizumab, tovetumab, trastuzumab, vedolizumab, veltuzumab, zalutumumab, and zanolimumab.

61. The method of any one of claims 1-60, wherein the heavy chain CDR1, 15 CDR2, and CDR3 fragments, the heavy chain variable domain gene, and optionally the light chain CDR1, CDR2, and CDR3 fragments, and the light chain variable domain gene are derived from naturally-occurring antibodies of a mammalian species.

62. The method of claim 61, wherein the mammalian species is human or camelid. 20

63. The method of any one of claims 1-62, wherein the antibody library is a full-length antibody library, a Fab antibody library, a single-chain antibody library, or a single domain antibody library.

64. The method of any one of claims 1-63, wherein the antibody library is a 25 human antibody library.

65. The method of any one of claims 1-33, wherein the antibody library is a camelid VHH antibody library. 30

66. A method for preparing a human antibody light chain library, the method comprising

providing (a) a first plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 1 (CDR1) fragments, (b) a
5 second plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 2 (CDR2) fragments, and/or (c) a third plurality of nucleic acids encoding a population of naturally-occurring antibody light chain complementary determining region 3 (CDR3) fragments, and

10 inserting the first plurality of nucleic acids, the second plurality of nucleic acids, and/or the third plurality of nucleic acids into the CDR1 region, the CDR2 region, and the CDR3 region, respectively, of an antibody light chain variable domain gene, thereby producing the human antibody light chain library.

67. The method of claim 50, wherein the population of naturally-occurring
15 antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising one or more of:

- (i) a glycosylation site,
- (ii) a deamidation site,
- 20 (iii) an isomerization site,
- (iv) unpaired cysteine,
- (v) net charge greater than 1,
- (vi) a tripeptide motif containing at least two aromatic residues,
- (vii) a motif that promotes aggregation,
- 25 (viii) a polyspecificity site;
- (ix) a protease sensitive site,
- (x) an integrin binding site,
- (xi) a lysine glycation site,
- (xii) a metal catalyzed fragmentation site,
- 30 (xiii) a polyspecificity aggregation site; and
- (xiv) a streptavidin binding motif.

68. The method of claim 67, wherein the glycosylation site of (i) comprises the motif NXS, NXT, or NXC, in which X represents any naturally-occurring amino acid residue except for proline.

5 69. The method of claim 67 or claim 68, wherein the deamidation site of (ii) comprises the motif of NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG.

70. The method of any one of claims 67-69, wherein the isomerization site of (iii) comprises the motif of DT, DH, DS, DG, or DD.

10 71. The method of any one of claims 67-70, wherein the tripeptide of (vi) is HYF or HWH.

72. The method of any one of claims 67-71, wherein the motif that promotes aggregation of (vii) comprises the motif of FHW.

15

73. The method of any one of claims 67-72, wherein the polyspecificity site of (viii) comprises the motif GG, GGG, RR, VG, W, WV, WW, WWW, YY, or WXW, in which X represents any amino acid residue.

20

74. The method of any one of claims 67-73, wherein the protease cleavage site of (ix) comprises the motif of DX, in which X is P, G, S, V, Y, F, Q, K, L, or D.

75. The method of any one of claims 67-74, wherein the integrin binding site of (x) comprises RGD, RYD, LDV, or KGD.

25

76. The method of any one of claims 67-74, wherein the lysine glycation site of (xi) comprises KE, EK, or ED.

77. The method of any one of claims 67-76, wherein the metal catalyzed fragmentation site of (xii) comprises the motif of HS, SH, KT, HXS, or SXH, in which X represents any amino acid residue.

30

78. The method of any one of claims 67-77, wherein the polyspecificity aggregation site of (xiii) comprises a motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is selected from the group consisting of F, I, L, V, W and Y.

5 79. The method of any one of claims 67-78, wherein the streptavidin binding motif of (xiv) comprises the motif HPQ, EPDW (SEQ ID NO: 117), PWWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), or PWPWL (SEQ ID NO: 120).

10 80. The method of any one of claims 67-79, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising at least two of (i)-(xiv).

15 81. The method of claim 80, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising (i)-(ix).

20 82. The method of claim 81, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragment is further free of members comprising one or more of (x)-(xiv).

25 83. The method of any one of claims 66-82, wherein the first plurality of nucleic acids, the second plurality of nucleic acids, and/or the third plurality of nucleic acids is produced by a process comprising:

(a) obtaining amino acid sequences of the light chain CDR1, CDR2, and/or CDR3 regions of a population of naturally-occurring antibodies,

(b) excluding from (a) the light chain CDR1, CDR2, and/or CDR3 amino acid sequences that comprise one or more of (i) to (xiv) to obtain liability-free light chain CDR1, CDR2, and/or CDR3 sequences, and

30

(c) synthesizing the first plurality of nucleic acids, the second plurality of the nucleic acids, and/or the third plurality of nucleic acids that encode the liability-free light chain CDR1, CDR2, and/or CDR3 regions.

5 84. The method of claim 83, wherein the process for producing the first, second, and/or third plurality of nucleic acids further comprises, after step (c), isolating functional members from the liability-free light chain CDR1, CDR2, and/or CDR3 regions.

10 85. The method of claim 84, wherein the functional members of the liability-free light chain CDR1, CDR2, and/or CDR3 are isolated by expressing antibodies comprising the liability-free light chain CDR1, CDR2, and/or CDR3 regions in host cells in a manner that the antibodies are displayed on surface of the host cells, isolating the antibodies that display on the host cells, and identifying the CDR1, CDR2, and/or CDR3 regions in the displayed antibodies, which are functional members of the liability-free light chain CDR1, CDR2,
15 and/or CDR3 regions.

 86. The method of claim 85, wherein the functional members show reduced binding to polyspecificity reagents.

20 87. The method of claim 65, wherein the functional members of the liability-free heavy chain CDR1, CDR2, and/or CDR3 are isolated by expressing antibodies comprising the liability-free heavy chain CDR1, CDR2, and/or CDR3 regions in fusion with a folding reporter, which optionally is β lactamase or green fluorescent protein, or fragments thereof, to obtain members with improved folding.

25 88. The method of any one of claims 66-87, wherein the antibody library comprises members each of which comprises a light chain CDR1, a light chain CDR2, and/or a light chain CDR3 that are not found in the same naturally-occurring antibody.

30 89. The method of claim 88, wherein at least 50% of the members in the antibody library comprise light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not found in the same naturally-occurring antibody.

90. The method of any one of claims 66-89, wherein the antibody light chain variable domain gene is a human antibody light chain variable domain, which optionally is selected from the group consisting of K1-12, K4-1, K2D-29, K3-11, K3-20, and L2-14.

5 91. The method of claim 90, wherein the human antibody light chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, and evoculumab.

92. The method of claim 90, wherein the human antibody light chain variable
 10 region gene is derived from a therapeutic antibody selected from the group consisting of abrituzumab, adalimumab, alemtuzumab, alirocumab, bapineuzumab, benralizumab, brodalumab, canakinumab, certolizumab, clazakizumab, dacetuzumab, daclizumab, daratumumab, eculizumab, efalizumab, elotuzumab, epratuzumab, farletuzumab, fasinumab, ficlatuzumab, fletikumab, fresolimumab, fulranumab, gevokizumab, ibalizumab, lintuzumab,
 15 matuzumab, mavrilimumab, mogamulizumab, motavizumab, natalizumab, nivolumab, obinutuzumab, ofatumumab, olokizumab, omalizumab, onartuzumab, otelixizumab, otlertuzumab, palivizumab, panitumumab, panobacumab, pertuzumab, pinatuzumab, polatuzumab, radretumab, ramucirumab, reslizumab, romosozumab, sarilumab, secukinumab, sifalimumab, tabalumab, tigatuzumab, tildrakizumab, tocilizumab, tovetumab, trastuzumab,
 20 vedolizumab, veltuzumab, zalutumumab, and zanolimumab..

93. The method of any one of claims 66-92, wherein the light chain CDR1, CDR2, and CDR3 fragments are derived from naturally-occurring antibodies of a mammalian species.

25 94. The method of claim 93, wherein the mammalian species is human.

95. An antibody library, which is prepared by a method of any one of claims 1-94.

30 96. An antibody library, comprising a first plurality of nucleic acids or a plurality of genetic packages comprising the nucleic acids, wherein the first plurality of nucleic acids encode a population of antibody heavy chain variable domains, which collectively comprise a

population of heavy chain CDR1s, a population of heavy chain CDR2s, and/or a population of heavy chain CDR3s located at the CDR1 region, the CDR2 region, and the CDR3 region of a common antibody heavy chain variable domain gene, wherein the heavy chain CDR1s, the heavy chain CDR2s, the heavy chain CDR3s, and optionally the common antibody chain variable domain gene are derived from naturally-occurring antibodies of the same mammal species.

97. The antibody library of claim 96, wherein the population of heavy chain CDR1s, the population of heavy chain CDR2s, and/or the population of heavy chain CDR3s is free of members comprising one or more of:

- (i) a glycosylation site,
- (ii) a deamidation site,
- (iii) an isomerization site,
- (iv) unpaired cysteine,
- (v) net charge greater than 1,
- (vi) a tripeptide motif containing at least two aromatic residues,
- (vii) a motif that promotes aggregation,
- (viii) a polyspecificity site;
- (ix) a protease sensitive site,
- (x) an integrin binding site,
- (xi) a lysine glycation site,
- (xii) a metal catalyzed fragmentation site,
- (xiii) a polyspecificity aggregation site; and
- (xiv) a streptavidin binding motif.

98. The antibody library of claim 97, wherein the glycosylation site of (i) comprises the motif NXS, NXT, or NXC, in which X represents any naturally-occurring amino acid residue except for proline.

99. The antibody library of claim 97 or claim 98, wherein the deamidation site of (ii) comprises the motif of NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG.

100. The antibody library of any one of claims 97-99, wherein the isomerization site of (iii) comprises the motif of DT, DH, DS, DG, or DD.

101. The antibody library of any one of claims 97-100, wherein the tripeptide of
5 (vi) is HYF or HWH.

102. The antibody library of any one of claims 97-101, wherein the motif that promotes aggregation of (vii) comprises the motif of FHW.

103. The antibody library of any one of claims 97-102, wherein the polyspecificity site of (viii) comprises the motif GG, GGG, RR, VG, W, WV, WW, WWW, YY, or WXW, in which X represents any amino acid residue.

104. The antibody library of any one of claims 97-103, wherein the protease
15 cleavage site of (ix) comprises the motif of DX, in which X is P, G, S, V, Y, F, Q, K, L, or D.

105. The antibody library of any one of claims 97-104, wherein the integrin binding site of (x) comprises RGD, RYD, LDV, or KGD.

106. The antibody library of any one of claims 97-105, wherein the lysine glycation site of (xi) comprises KE, EK, or ED.

107. The antibody library of any one of claims 97-106, wherein the metal catalyzed fragmentation site of (xii) comprises the motif of HS, SH, KT, HXS, or SXH, in which X
25 represents any amino acid residue.

108. The antibody library of any one of claims 97-107, wherein the polyspecificity aggregation site of (xiii) comprises a motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is selected from the group consisting of F, I, L, V, W and Y.

30

109. The antibody library of any one of claims 97-108, wherein the streptavidin binding motif of (xiv) comprises the motif HPQ, EPDW (SEQ ID NO: 117), PWXWL (SEQ

ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), or PWPWLG (SEQ ID NO: 120).

110. The antibody library of any one of claims 97-19, wherein the population of
5 antibody heavy chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising at least two of (i)-(xiv).

111. The antibody library of claim 110, wherein the population of antibody heavy
10 chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragments is free of members comprising (i)-(ix).

112. The antibody library of claim 111, wherein the population of antibody heavy
15 chain CDR1 fragments, the population of antibody heavy chain CDR2 fragments, and/or the population of antibody heavy chain CDR3 fragment is further free of members comprising one or more of (x)-(xiv).

113. The antibody library of any one of claims 97-112, wherein the population of
20 heavy chain CDR1s, the population of heavy chain CDR2s, and/or the population of heavy chain CDR3s is further substantially free of non-functional members.

114. The antibody library of any one of claims 97-113, wherein the population of
heavy chain CDR1s and/or the population of heavy chain CDR2s is free of members
comprising one or more of (i)-(xiv) and the population of heavy chain CDR3s is derived from
25 naturally-occurring antibodies without removal of liabilities.

115. The antibody library of claim 114, wherein the population of heavy chain
CDR3s is derived from human B cells.

30 116. The antibody library of any one of claims 97-115, which comprises members having heavy chain CDR1, heavy chain CDR2, and/or heavy chain CDR3 that are not from the same naturally-occurring antibody.

117. The antibody library of claim 116, wherein at least 50% of the members in the antibody library have heavy chain CDR1, heavy chain CDR2, and/or heavy chain CDR3 that are not from the same naturally-occurring antibody.

5 118. The antibody library of any one of claims 97-117, wherein the antibody heavy chain variable region gene is a human antibody heavy chain variable region gene, which optionally is selected from the group consisting of VH1-24, VH2-70, VH3-7, VH4-30-4, VH5-51, VH1-18, , VH1-69, VH3-23, VH5-10-1, VH3-9, and VH3-11.

10 119. The antibody library of claim 118, wherein the human antibody heavy chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, and evoculumab.

15 120. The antibody library of claim 118, wherein the human antibody heavy chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abinituzumab, adalimumab, alemtuzumab, alirocumab, bapineuzumab, benralizumab, brodalumab, canakinumab, certolizumab, clazakizumab, dacetuzumab, daclizumab, daratumumab, eculizumab, efalizumab, elotuzumab, epratuzumab, farletuzumab, fasinumab, ficlatuzumab, fletikumab, fresolimumab, fulranumab, gevokizumab, ibalizumab, lintuzumab, matuzumab, mavrilimumab, mogamulizumab, motavizumab, natalizumab, nivolumab, obinutuzumab, ofatumumab, olokizumab, omalizumab, onartuzumab, orelizumab, orelertuzumab, palivizumab, panitumumab, panobacumab, pertuzumab, pinatuzumab, polatuzumab, radretumab, ramucirumab, reslizumab, romosozumab, sarilumab, secukinumab, sifalimumab, tabalumab, tigatuzumab, tildrakizumab, tocilizumab, tovetumab, 25 trastuzumab, vedolizumab, veltuzumab, zalutumumab, and zanolimumab.

121. The antibody library of any one of claims 97-120, further comprising a nucleic acid encoding a common light chain variable domain.

30 122. The antibody library of claim 121, wherein the common light chain variable domain is VK3-20.

123. The antibody library of any one of claims 97-120, further comprising a second plurality of nucleic acids encoding a population of antibody light chain variable domains, which collectively comprise a population of light chain CDR1s, a population of light chain CDR2s, and a population of light chain CDR3s located at the CDR1 region, the CDR2 region, and the CDR3 region of a common antibody light chain variable domain gene, wherein the light chain CDR1s, CDR2s, CDR3s, and optionally the common antibody light chain domain gene are derived from naturally-occurring antibody of the same mammal species.

124. The antibody library of claim 123, wherein the population of light chain CDR1s, the population of light chain CDR2s, and/or the population of light chain CDR3s is free of members comprising one or more of:

- (i) a glycosylation site,
- (ii) a deamidation site,
- (iii) an isomerization site,
- (iv) unpaired cysteine,
- (v) net charge greater than 1,
- (vi) a tripeptide motif containing at least two aromatic residues,
- (vii) a motif that promotes aggregation,
- (viii) a polyspecificity site;
- (ix) a protease sensitive site,
- (x) an integrin binding site,
- (xi) a lysine glycation site,
- (xii) a metal catalyzed fragmentation site,
- (xiii) a polyspecificity aggregation site; and
- (xiv) a streptavidin binding motif.

125. The antibody library of claim 124, wherein the glycosylation site of (i) comprises the motif NXS,NXT, or NXC, in which X represents any naturally-occurring amino acid residue except for proline.

126. The antibody library of claim 124 or claim 125, wherein the deamidation site of (ii) comprises the motif of NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG.

127. The antibody library of any one of claims 124-126, wherein the isomerization
5 site of (iii) comprises the motif of DT, DH, DS, DG, or DD.

128. The antibody library of any one of claims 124-127, wherein the tripeptide of (vi) is HYF or HWH.

10 129. The antibody library of any one of claims 124-128, wherein the motif that promotes aggregation of (vii) comprises the motif of FHW.

130. The antibody library of any one of claims 124-129, wherein the protease sensitive site of (viii) comprises the motif GG, GGG, RR, VG, W, WV, WW, WWW, YY, or
15 WXW, in which X represents any amino acid residue.

131. The antibody library of any one of claims 124-130, wherein the protease cleavage site of (ix) comprises the motif of DX, in which X is P, G, S, V, Y, F, Q, K, L, or D.

20 132. The antibody library of any one of claims 124-131, wherein the integrin binding site of (x) comprises RGD, RYD, LDV, or KGD.

133. The antibody library of any one of claims 124-132, wherein the lysine glycation site of (xi) comprises KE, EK, or ED.

25 134. The antibody library of any one of claims 124-133, wherein the metal catalyzed fragmentation site of (xii) comprises the motif of HS, SH, KT, HXS, or SXH, in which X represents any amino acid residue.

30 135. The antibody library of any one of claims 124-134, wherein the polyspecificity aggregation site of (xiii) comprises a motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is selected from the group consisting of F, I, L, V, W and Y.

136. The antibody library of any one of claims 124-135, wherein the streptavidin binding motif of (xiv) comprises the motif HPQ, EPDW (SEQ ID NO: 117), PWXWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), or PWPWL (SEQ ID NO: 120).

137. The antibody library of any one of claims 124-136, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising at least two of (i)-(xiv).

138. The method of claim 137, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising (i)-(ix).

139. The antibody library of claim 138, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragment is further free of members comprising one or more of (x)-(xiv).

140. The antibody library of any one of claims 123-139, which comprises members having light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not from the same naturally-occurring antibody.

141. The antibody library of claim 140, wherein at least 50% of the members in the antibody library have light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not from the same naturally-occurring antibody.

142. The antibody library of any one of claims 123-141, wherein the antibody light chain variable domain gene is a human light chain variable gene, which optionally is selected from the group consisting of K1-12, K4-1, K2D-29, K3-11, K3-20, and L2-14.

143. The antibody library of claim 142, wherein the human antibody light chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, and evoculumab.

5

144. The antibody library of claim 142-, wherein the human antibody heavy chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abritumumab, adalimumab, alemtuzumab, alirocumab, bapineuzumab, benralizumab, brodalumab, canakinumab, certolizumab, clazakizumab, dacetuzumab, 10 daclizumab, daratumumab, eculizumab, efalizumab, elotuzumab, epratuzumab, farletuzumab, fasinumab, ficlatuzumab, fletikumab, fresolimumab, fulranumab, gevokizumab, ibalizumab, lintuzumab, matuzumab, mavrilimumab, mogamulizumab, motavizumab, natalizumab, nivolumab, obinutuzumab, ofatumumab, olokizumab, omalizumab, onartuzumab, otelixizumab, otlertuzumab, palivizumab, panitumumab, panobacumab, pertuzumab, 15 pinatuzumab, polatuzumab, radretumab, ramucirumab, reslizumab, romosozumab, sarilumab, secukinumab, sifalimumab, tabalumab, tigatuzumab, tildrakizumab, tocilizumab, tovetumab, trastuzumab, vedolizumab, veltuzumab, zalutumumab, and zanolimumab.

145. The antibody library of any one of claims 96-144, wherein the heavy chain 20 CDR1, CDR2, and CDR3 fragments, the heavy chain variable domain gene, the light chain CDR1, CDR2, and CDR3 fragments, and the light chain variable domain gene are derived from naturally-occurring antibodies of a mammalian species.

146. The antibody library of claim 145, wherein the mammalian species is human 25 or camelid.

147. The antibody library of any one of claims 96-146, wherein the antibody library is a full-length antibody library, a Fab antibody library, a single-chain antibody library, or a single domain antibody library.

30

148. The antibody library of any one of claims 96-147, wherein the antibody library is a human antibody library.

149. The antibody library of any one of claims 97-120, wherein the antibody library is a camelid VHH antibody library.

5 150. An antibody light chain library, comprising a plurality of nucleic acids or a plurality of genetic packages comprising the nucleic acids, wherein the plurality of nucleic acids encode a population of antibody light chain variable domains, which collectively comprise a population of light chain CDR1s, a population of light chain CDR2s, and/or a population of light chain CDR3s inserted at the CDR1 region, the CDR2 region, and the
10 CDR3 region of a common antibody light chain variable domain gene, wherein the light chain CDR1s, CDR2s, and CDR3s, and optionally the common antibody light chain variable domain gene are derived from naturally-occurring antibodies of the same mammal species.

151. The antibody library of claim 150, wherein the population of light chain
15 CDR1s, the population of light chain CDR2s, and/or the population of light chain CDR3s is free of members comprising one or more of:

- (i) a glycosylation site,
- (ii) a deamidation site,
- (iii) an isomerization site,
- 20 (iv) unpaired cysteine,
- (v) net charge greater than 1,
- (vi) a tripeptide motif containing at least two aromatic residues,
- (vii) a motif that promotes aggregation,
- (viii) a polyspecificity site;
- 25 (ix) a protease sensitive site,
- (x) an integrin binding site,
- (xi) a lysine glycation site,
- (xii) a metal catalyzed fragmentation site,
- (xiii) a polyspecificity aggregation site; and
- 30 (xiv) a streptavidin binding motif.

152. The antibody library of claim 151, wherein the glycosylation site of (i) comprises the motif NXS, NXT, or NXC, in which X represents any naturally-occurring amino acid residue except for proline.

5 153. The antibody library of claim 151 or claim 152, wherein the deamidation site of (ii) comprises the motif of NG, NS, NT, NN, NA, NH, ND, GNF, GNY, GNT, or GNG.

154. The antibody library of any one of claims 151-153, wherein the isomerization site of (iii) comprises the motif of DT, DH, DS, DG, or DD.

10

155. The antibody library of any one of claims 151-154, wherein the tripeptide of (vi) is HYF or HWH.

156. The antibody library of any one of claims 151-155, wherein the motif that
15 promotes aggregation of (vii) comprises the motif of FHW.

157. The antibody library of any one of claims 151-156, wherein the polyspecificity site of (viii) comprises the motif GG, GGG, RR, VG, W, WV, WW, WWW, YY, or WXW, in which X represents any amino acid residue.

20

158. The antibody library of any one of claims 151-157, wherein the protease cleavage site of (ix) comprises the motif of DX, in which X is P, G, S, V, Y, F, Q, K, L, or D.

159. The antibody library of any one of claims 151-158, wherein the integrin
25 binding site of (x) comprises RGD, RYD, LDV, or KGD.

160. The antibody library of any one of claims 151-159, wherein the lysine glycation site of (xi) comprises KE, EK, or ED.

30 161. The antibody library of any one of claims 151-160, wherein the metal catalyzed fragmentation site of (xii) comprises the motif of HS, SH, KT, HXS, or SXH, in which X represents any amino acid residue.

162. The antibody library of any one of claims 151-161, wherein the polyspecificity aggregation site of (xiii) comprises a motif of $X_1X_2X_3$, wherein each of X_1 , X_2 , and X_3 independently is selected from the group consisting of F, I, L, V, W and Y.

5

163. The antibody library of any one of claims 151-162, wherein the streptavidin binding motif of (xiv) comprises the motif HPQ, EPDW (SEQ ID NO: 117), PWWL (SEQ ID NO: 118), in which X represents any amino acid residue, GDWVFI (SEQ ID NO: 119), or PWPWL (SEQ ID NO: 120).

10

164. The antibody library of any one of claims 151-163, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising at least two of (i)-(xiv).

15

165. The method of claim 164, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragments is free of members comprising (i)-(ix).

20

166. The antibody library of claim 165, wherein the population of antibody light chain CDR1 fragments, the population of antibody light chain CDR2 fragments, and/or the population of antibody light chain CDR3 fragment is further free of members comprising one or more of (x)-(xiv).

25

167. The antibody library of any one of claims 150-166, which comprises members having light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not from the same naturally-occurring antibody.

30

168. The antibody library of claim 167, wherein at least 50% of the members in the antibody library have light chain CDR1, light chain CDR2, and/or light chain CDR3 that are not from the same naturally-occurring antibody.

169. The antibody library of any one of claims 150-168, wherein the antibody light chain variable domain gene is a human antibody light chain variable domain gene, which optionally is selected from the group consisting of K1-12, K4-1, K2D-29, K3-11, K3-20, and L2-14.

5

170. The antibody library of claim 169, wherein the human antibody light chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abrilumab, mepolizumab, crenezumab, necitumumab, anifrolumab, and evoculumab.

10

171. The antibody library of claim 169, wherein the human antibody light chain variable region gene is derived from a therapeutic antibody selected from the group consisting of abritumumab, adalimumab, alemtuzumab, alirocumab, bapineuzumab, benralizumab, brodalumab, canakinumab, certolizumab, clazakizumab, dacetuzumab, daclizumab, daratumumab, eculizumab, efalizumab, elotuzumab, epratuzumab, farletuzumab, fasinumab, ficlatuzumab, fletikumab, fresolimumab, fulranumab, gevokizumab, ibalizumab, lintuzumab, matuzumab, mavrilimumab, mogamulizumab, motavizumab, natalizumab, nivolumab, obinutuzumab, ofatumumab, olokizumab, omalizumab, onartuzumab, orelizumab, orelizumab, palivizumab, panitumumab, panobacumab, pertuzumab, pinatuzumab, polatuzumab, radretumab, ramucirumab, reslizumab, romosozumab, sarilumab, secukinumab, sifalimumab, tabalumab, tigatuzumab, tildrakizumab, tocilizumab, tovetumab, trastuzumab, vedolizumab, veltuzumab, zalutumumab, and zanolimumab.

15

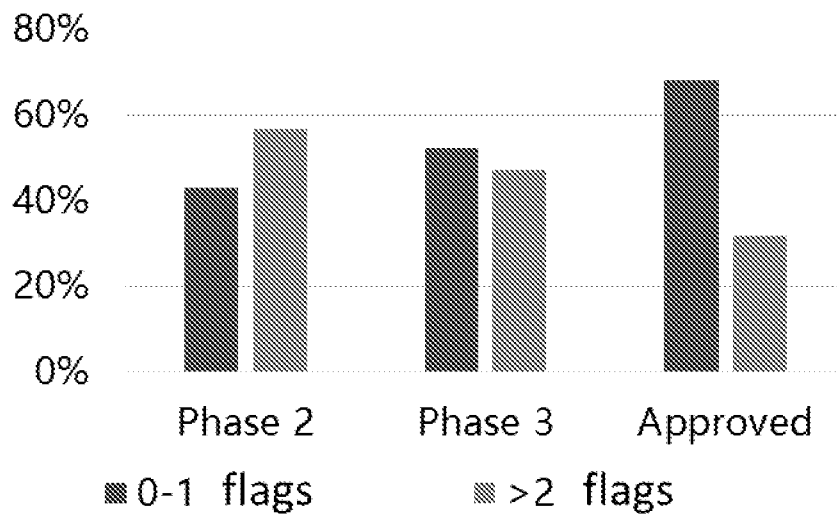
20

25

172. The antibody library of any one of claims 150-171, wherein the mammalian species is human.

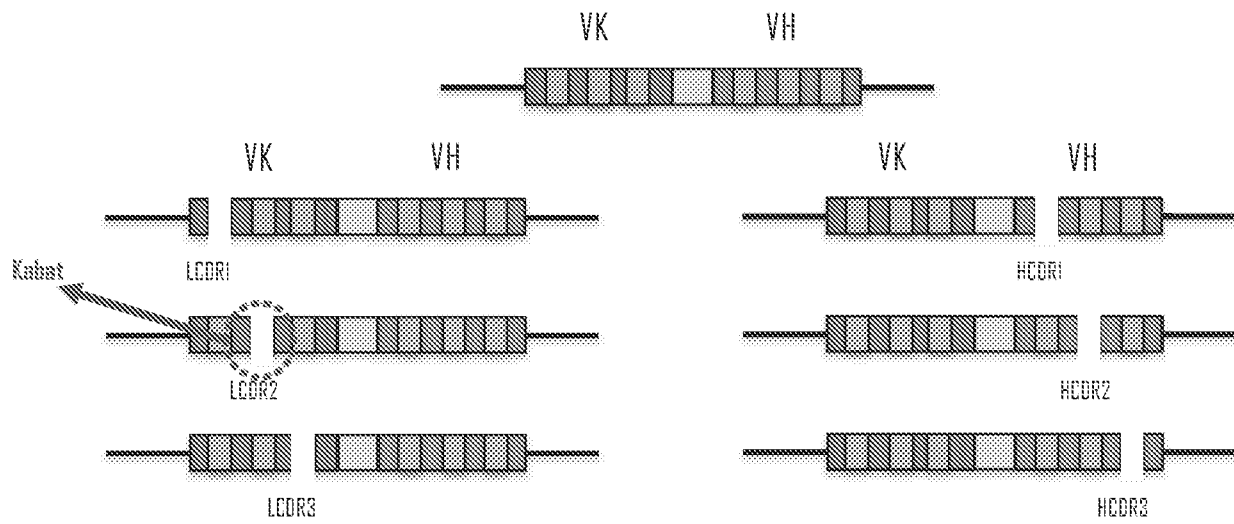
1/58

Figure 1



2/58

Figure 2



3/58

Figure 3

Two Bsal sites in opposite orientations

- Scarless cloning
- Removes the exact portion of the CDR as well as the Bsal site

5' GTGTTACCATCACCTGTCGTGCTTCTagataccatggccagtaaggcccttctctCTGGCTTGGTACCAGCAGAAA
 3' CACAATGGTAGTGGACAGCACGAAGAtctctggtaaccttcattccggccagagaGACCGAACCATGGTCGTCTTT

18 19 20 21 22 23 24 25 26 1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7
 R V T I T C R A S R D H G Q * G R S L A W Y Q Q K

Internal SfiI site

- Minimize uncut vector background
- Serves as spacer between Bsal sites

Ochre stop codon and frame-shift

- Prevent expression of background sequences

4/58

Figure 4

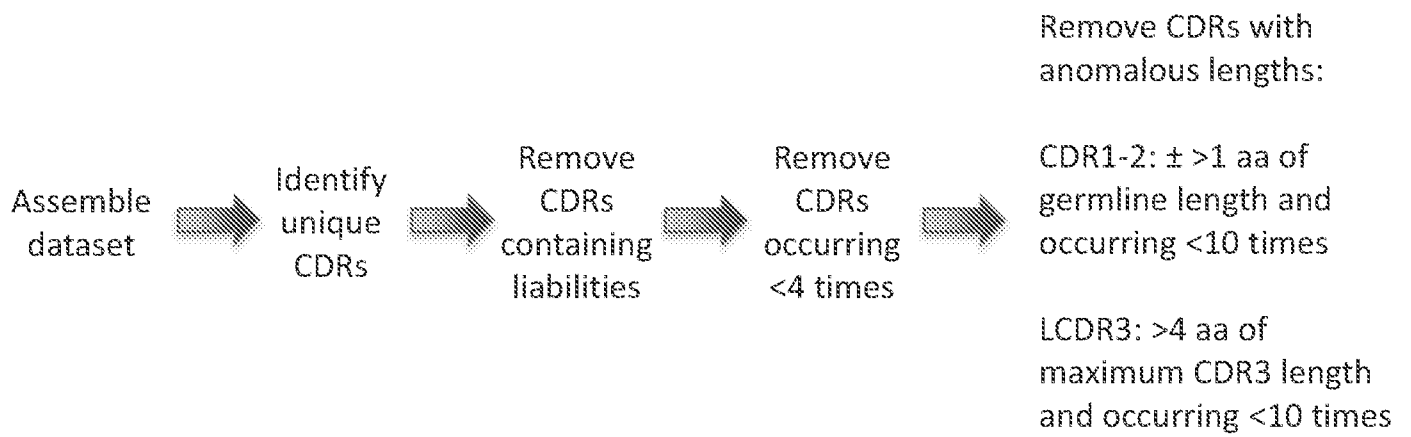
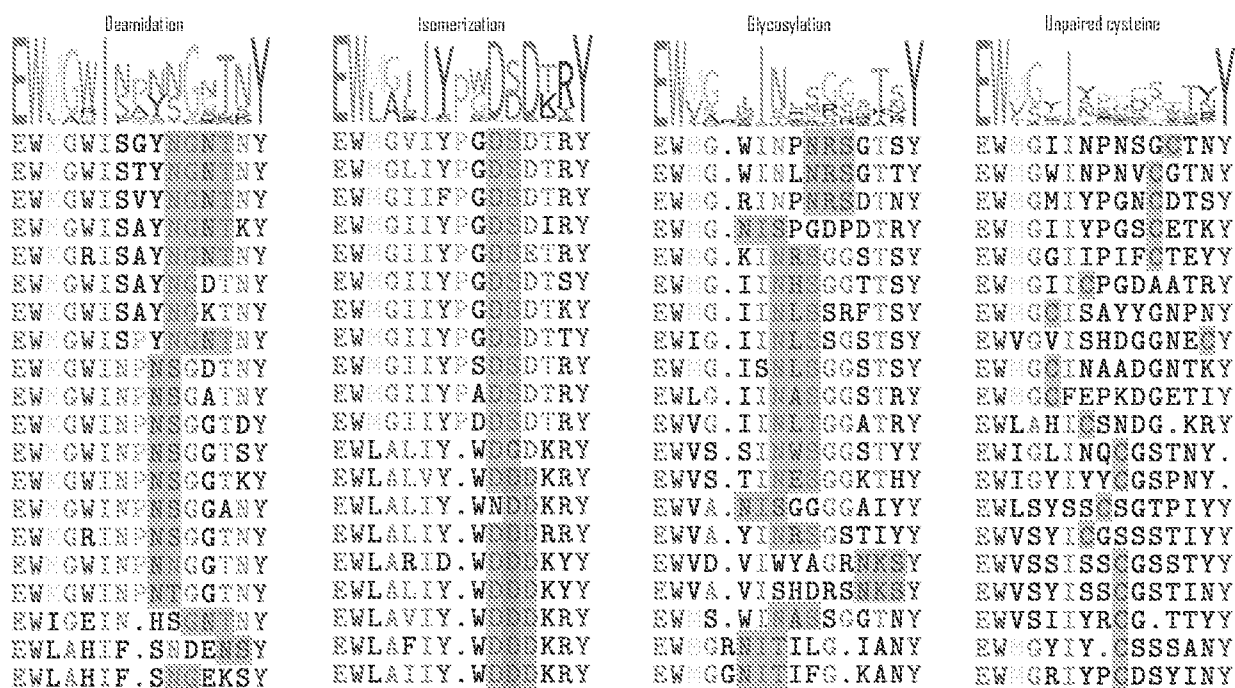
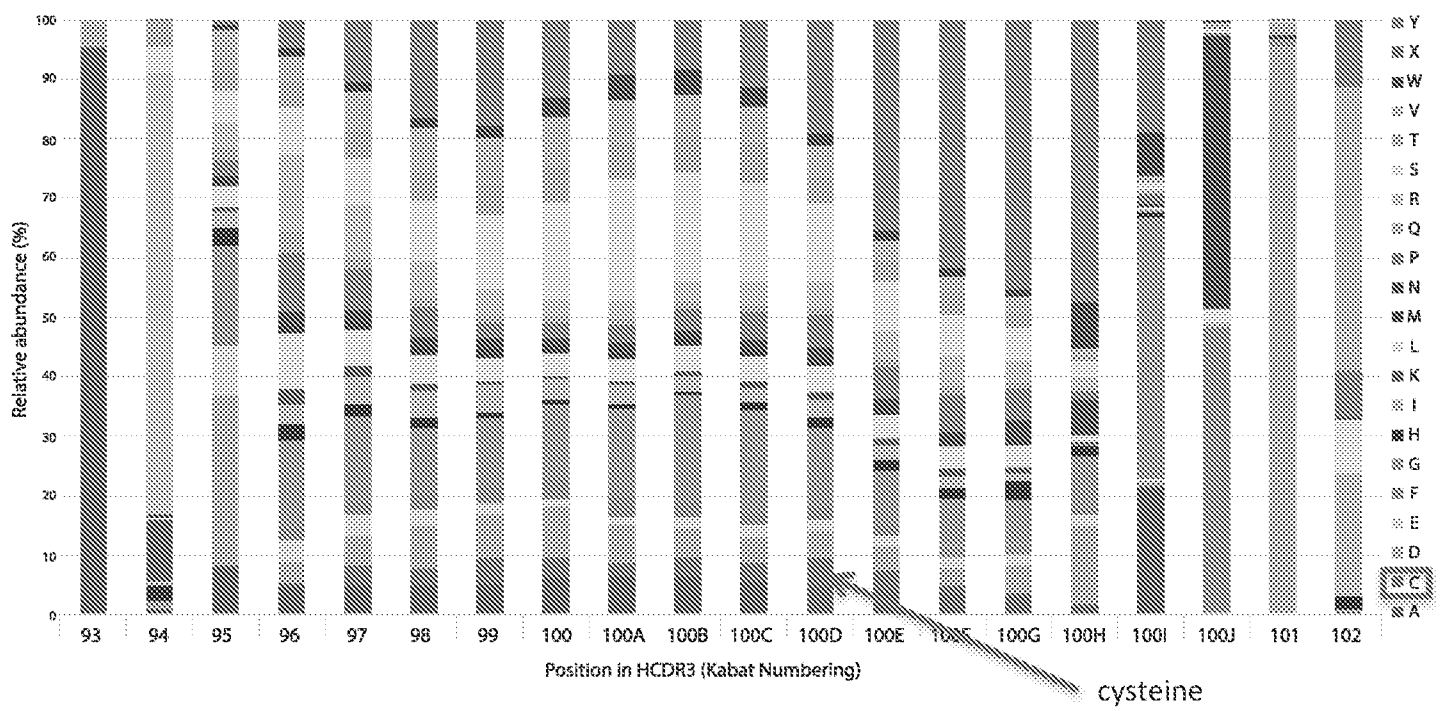


Figure 5





7/58

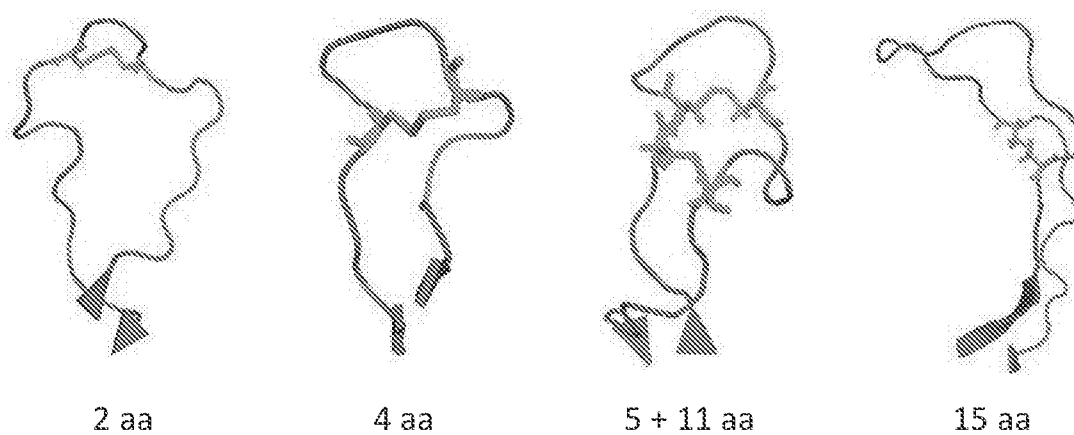
Figure 7

- Cysteines comprise up to 4% of HCDR3 amino acids
 - Depends upon HCDR3 length and position
- 85% of cysteines in HCDR3 pair within the HCDR3
- 10% pair with an additional cysteine elsewhere
- Cysteine-cysteine placement highly structured
 - Generally separated by 2 or 4 amino acids
 - There are exceptions

AGPSITESHYLDAAKDYYYGLDV
 AKDARDLLADWHFDL
 AKFSGKD SGTS RDY
 ARAPDADAD HKGAFGY
 ARDGGHGF SSAS FGPDY
 ARRGSDY GDFPWQY
 ARSPSYISGGTVFDH
 ARVGY SSTS NRGAFDI
 VRGH DGTTSRAX
 VRKGPS PH GDFHWQH
 VRSVTPRY GGGFYGEFDY
 VRTAD ERDP KGWVFPH
 VTLPDL PGDN TYPDAS
 VRGRS GGRRH NGAD FNWDFQH
 AKDLREDE EEWSDYYDFGKQLP RKSRGVAGIFD
 G

8/58

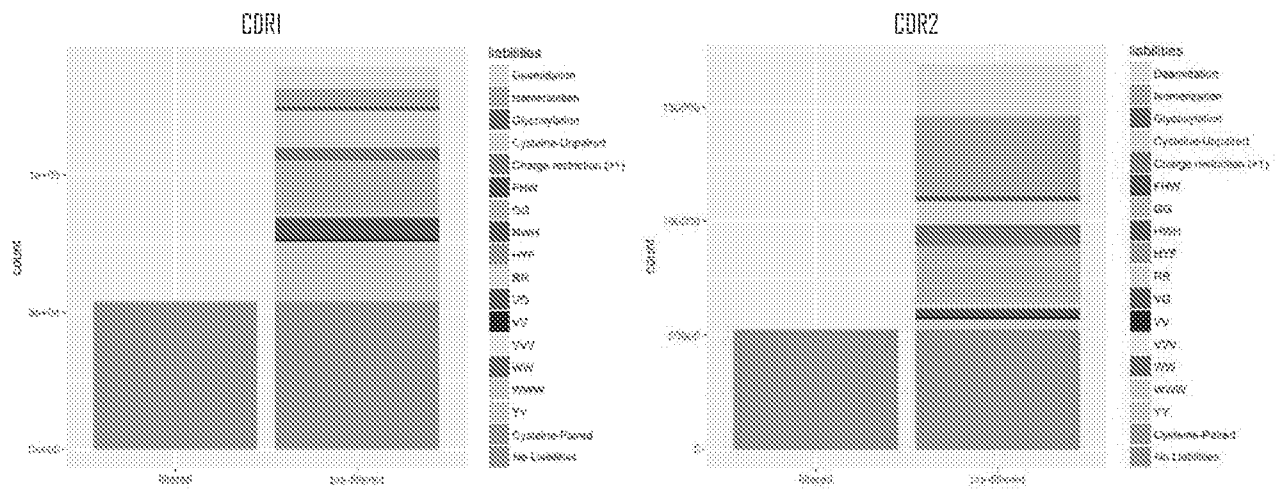
Figure 8



Cysteines in CDRs should be
paired and structurally placed

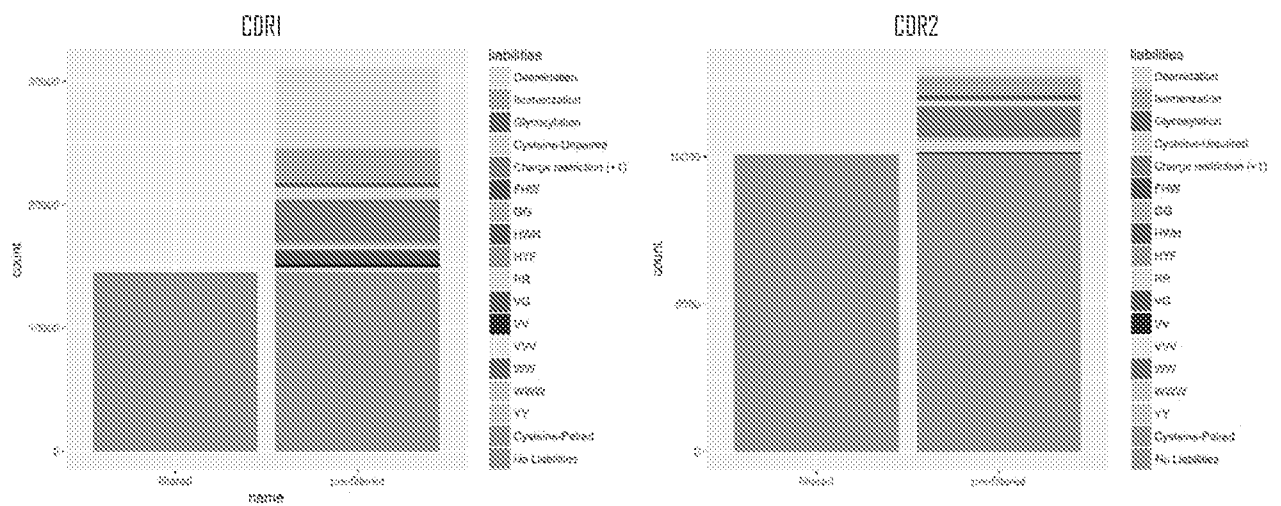
9/58

Figure 9



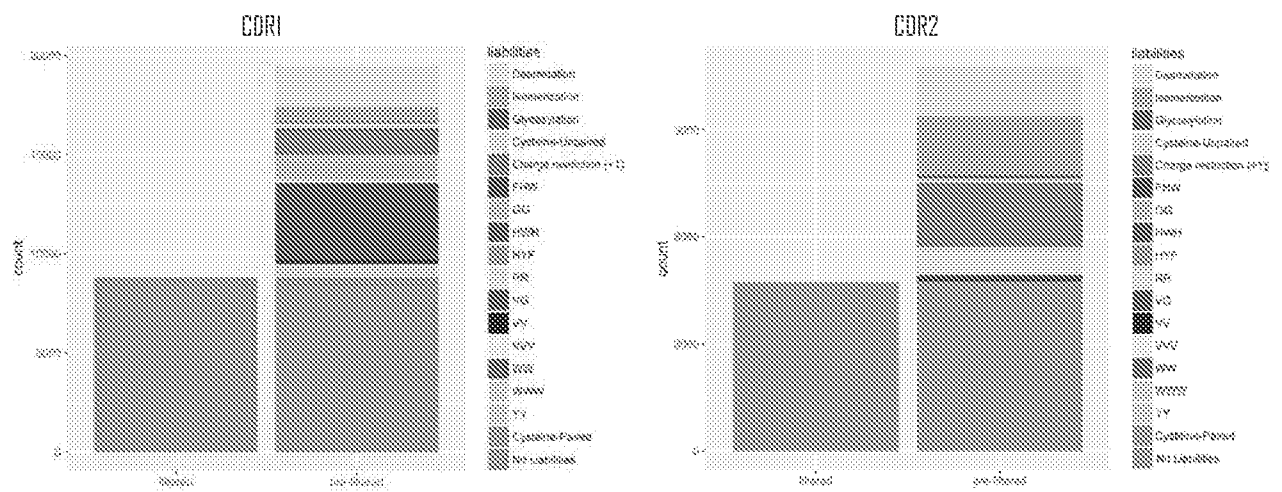
10/58

Figure 10



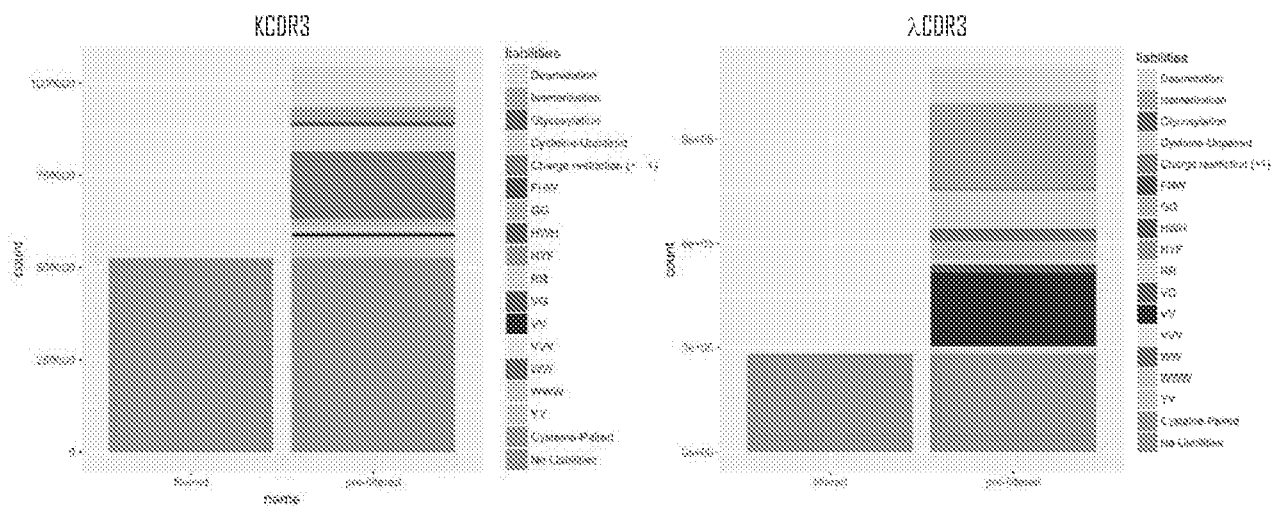
11/58

Figure 11



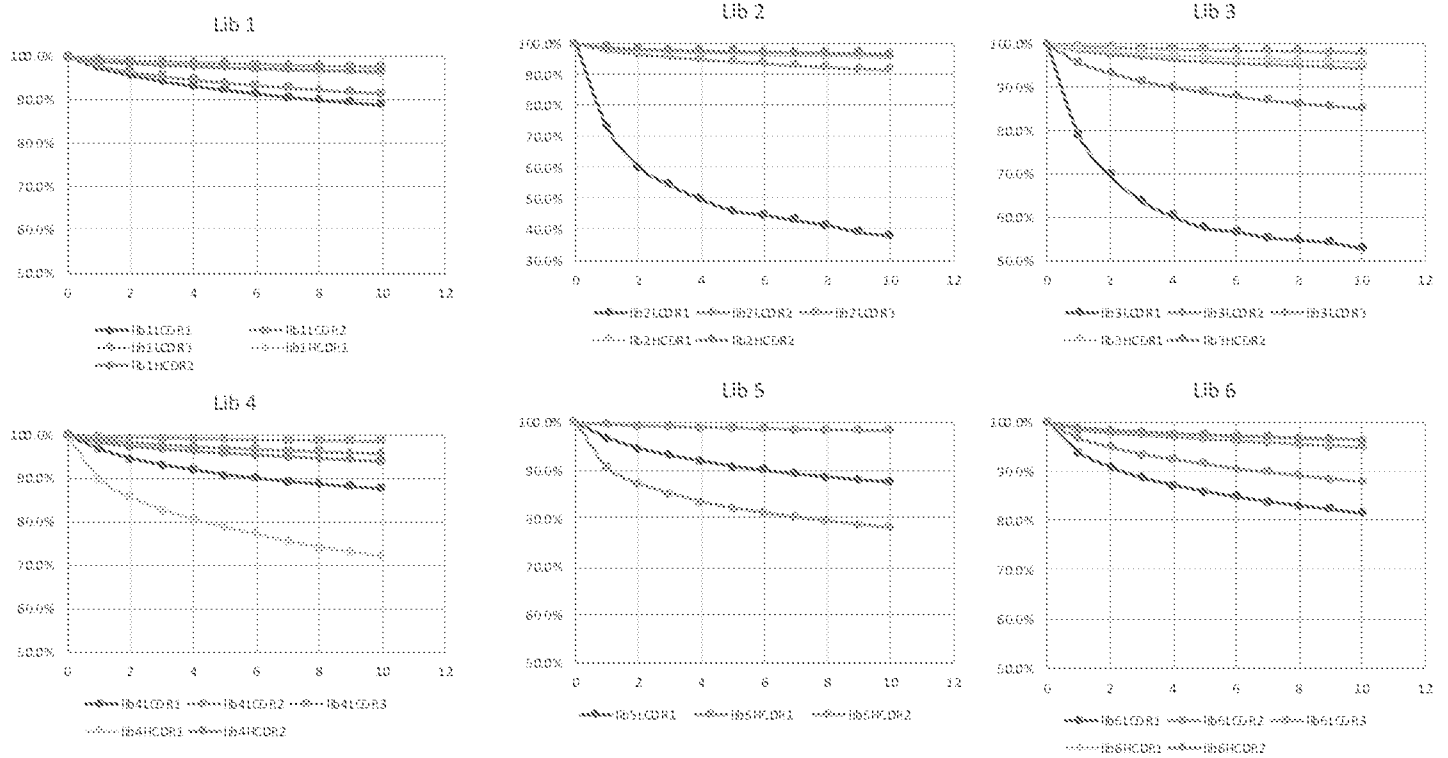
12/58

Figure 12



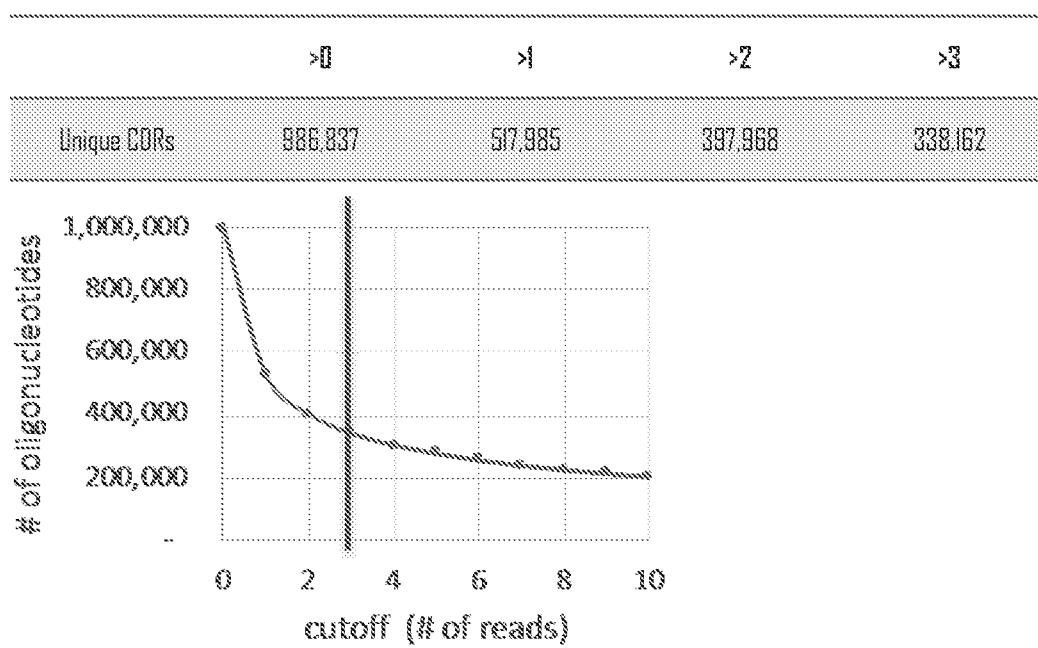
14/58

Figure 14



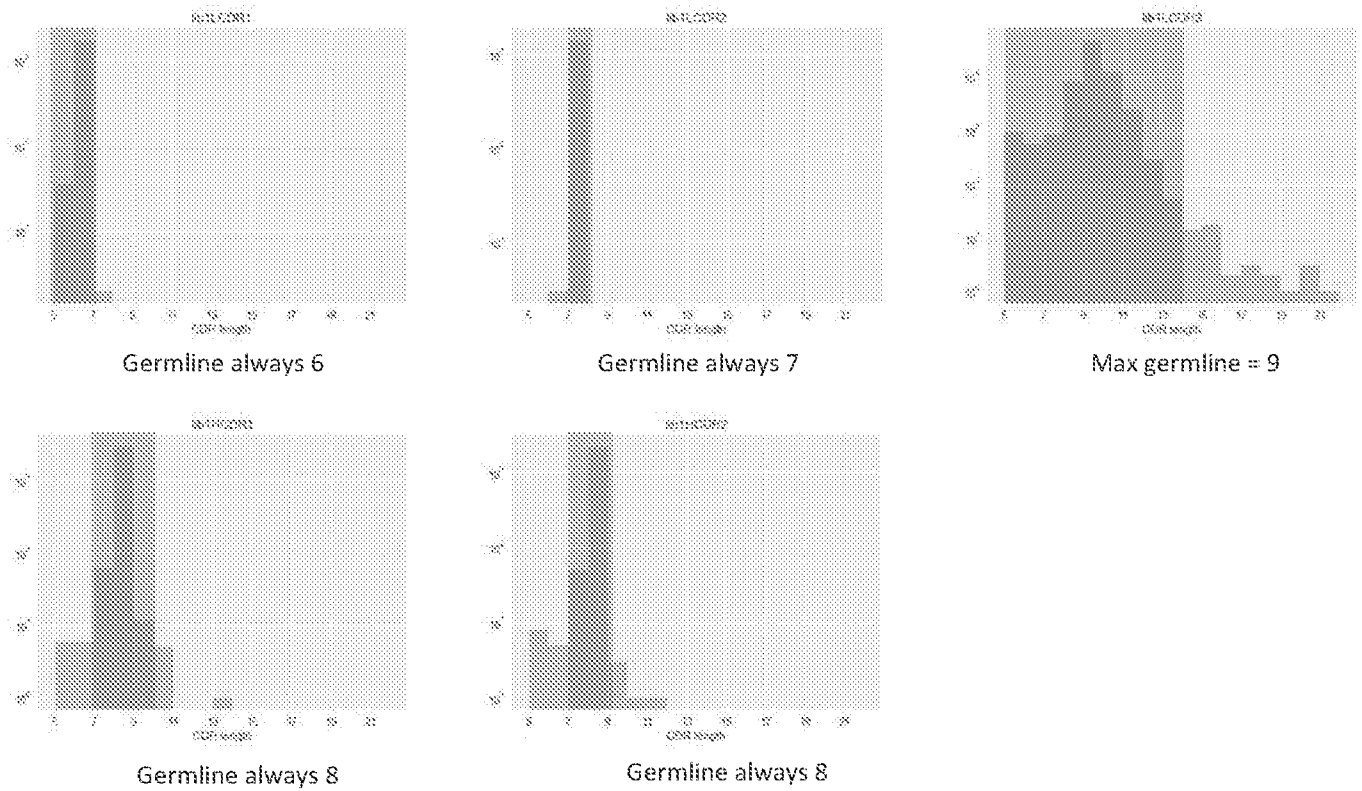
15/58

Figure 15



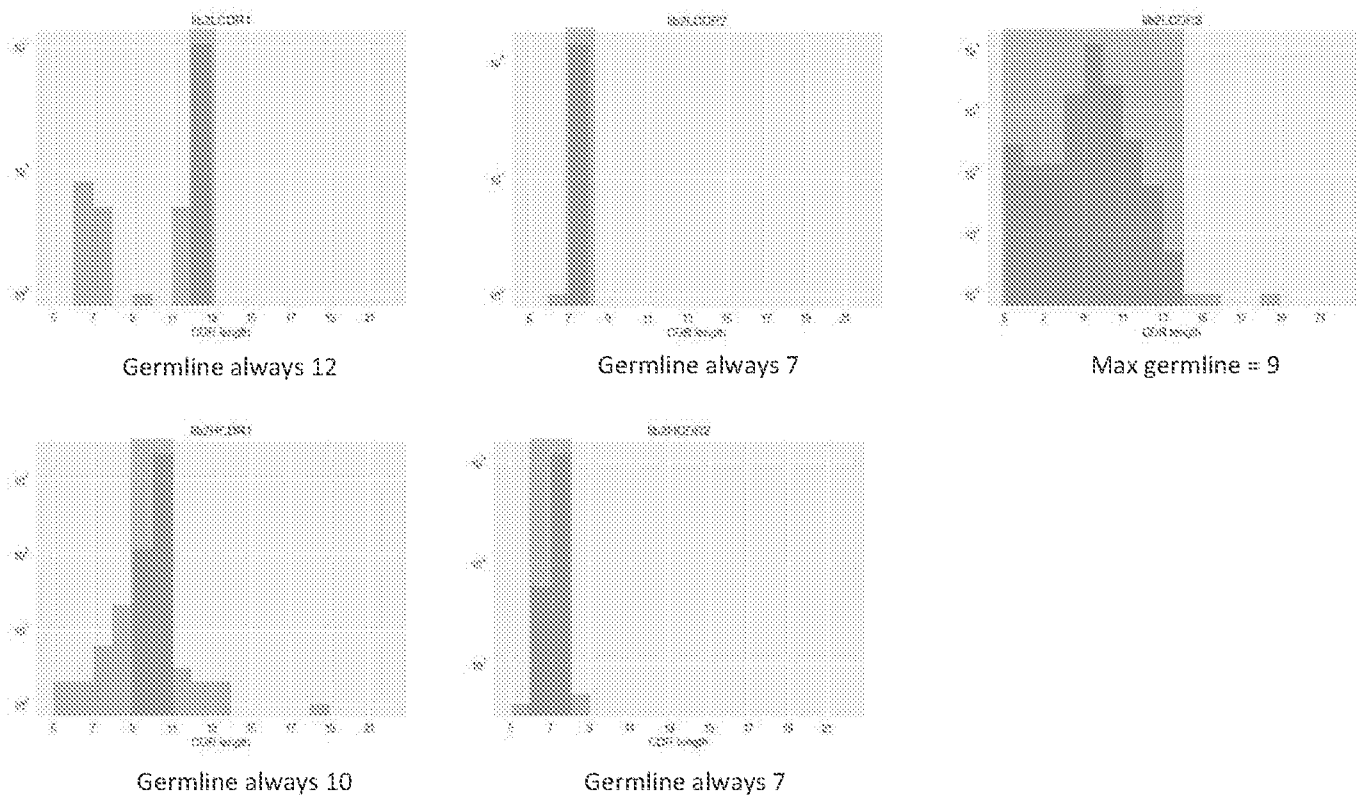
16/58

Figure 16



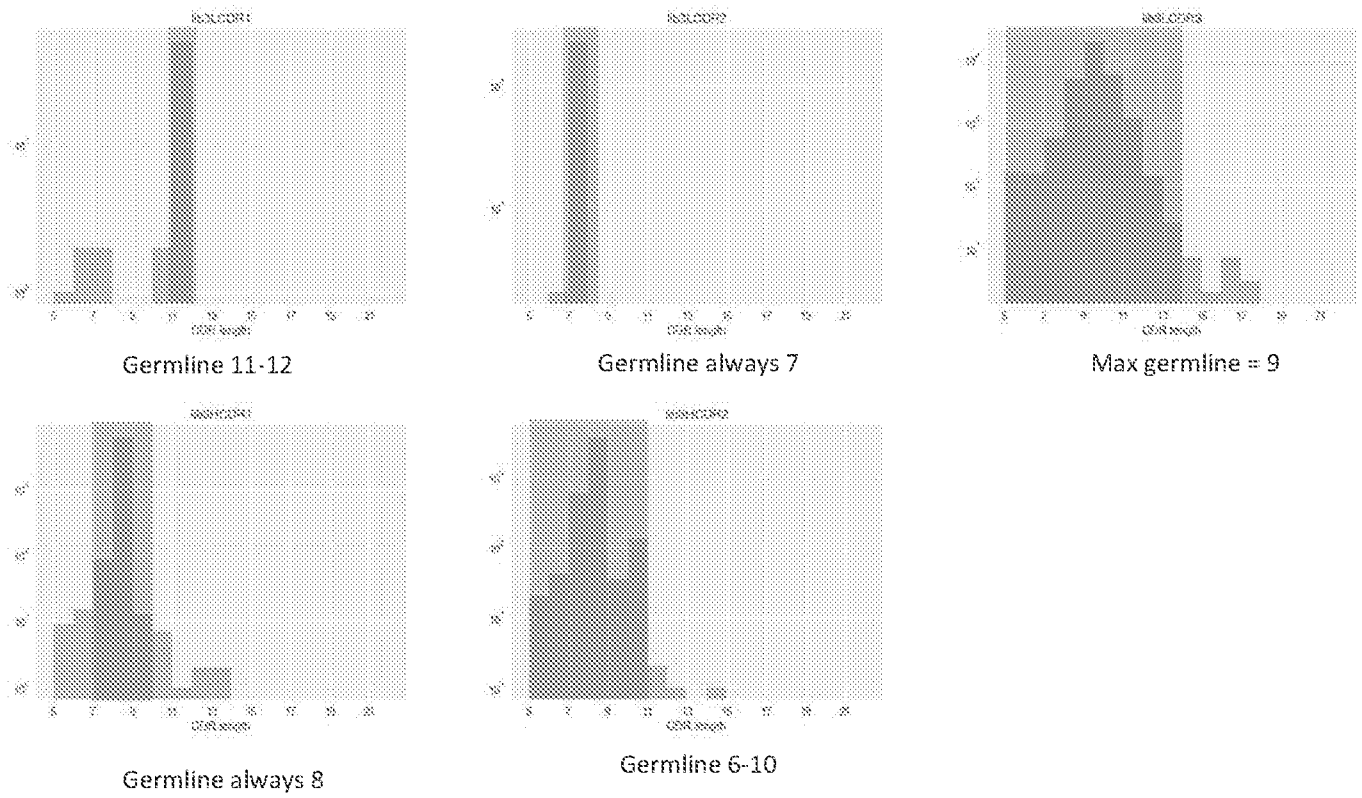
17/58

Figure 17



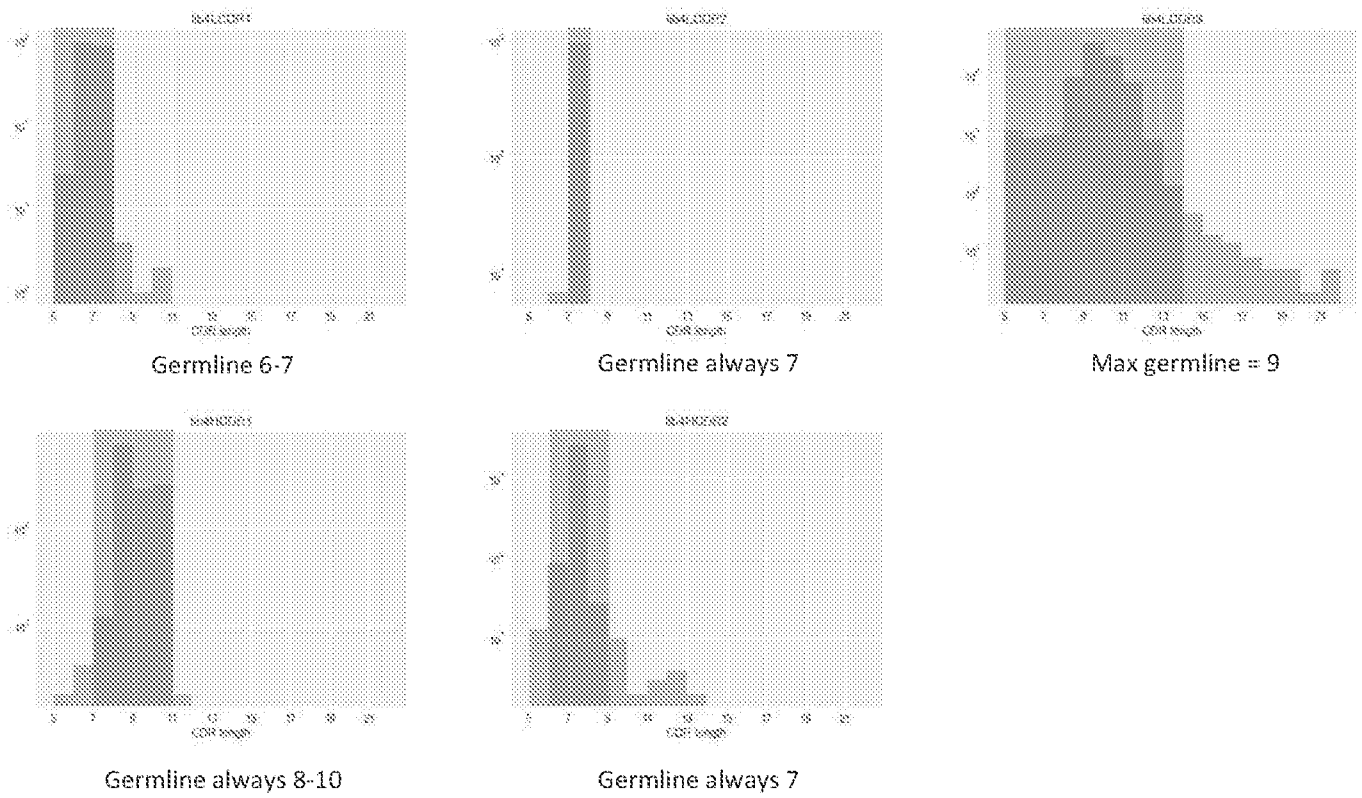
18/58

Figure 18



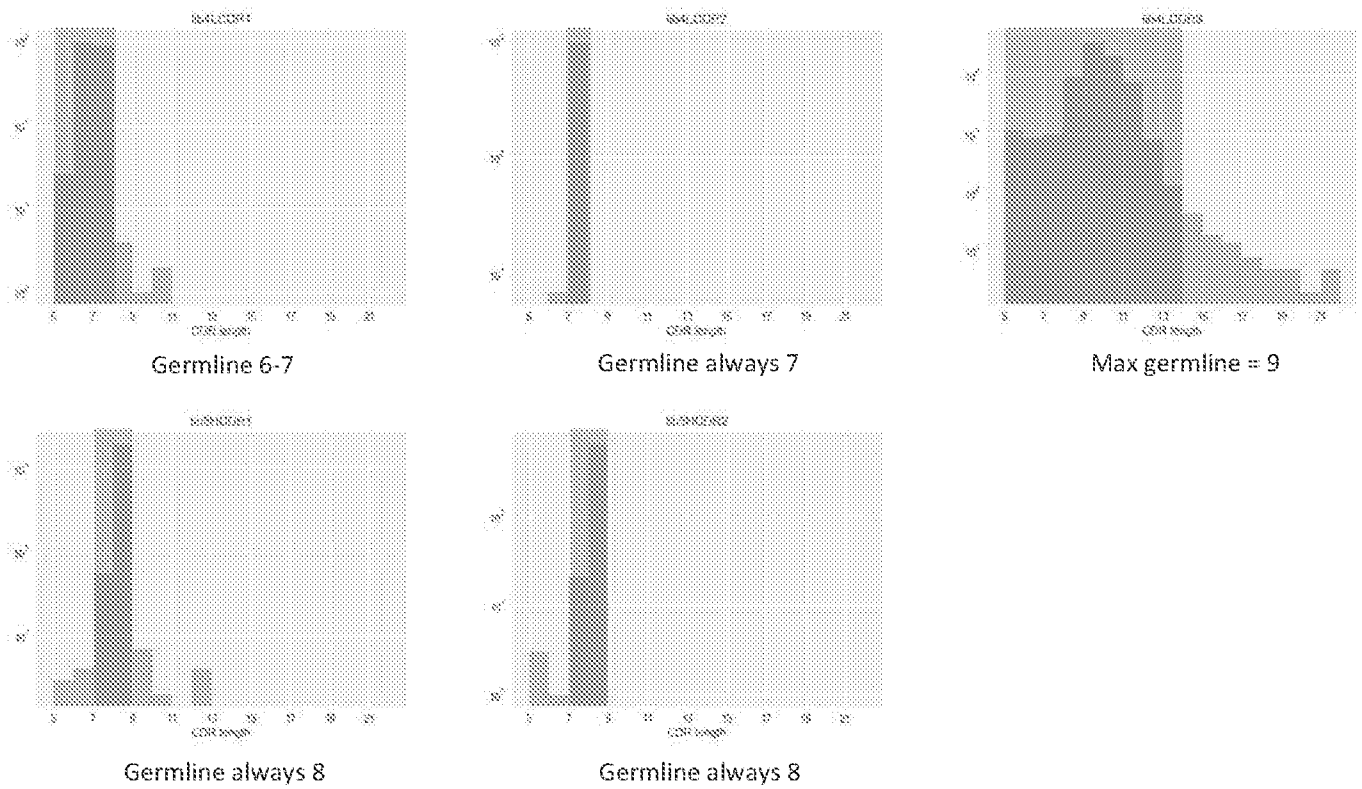
19/58

Figure 19



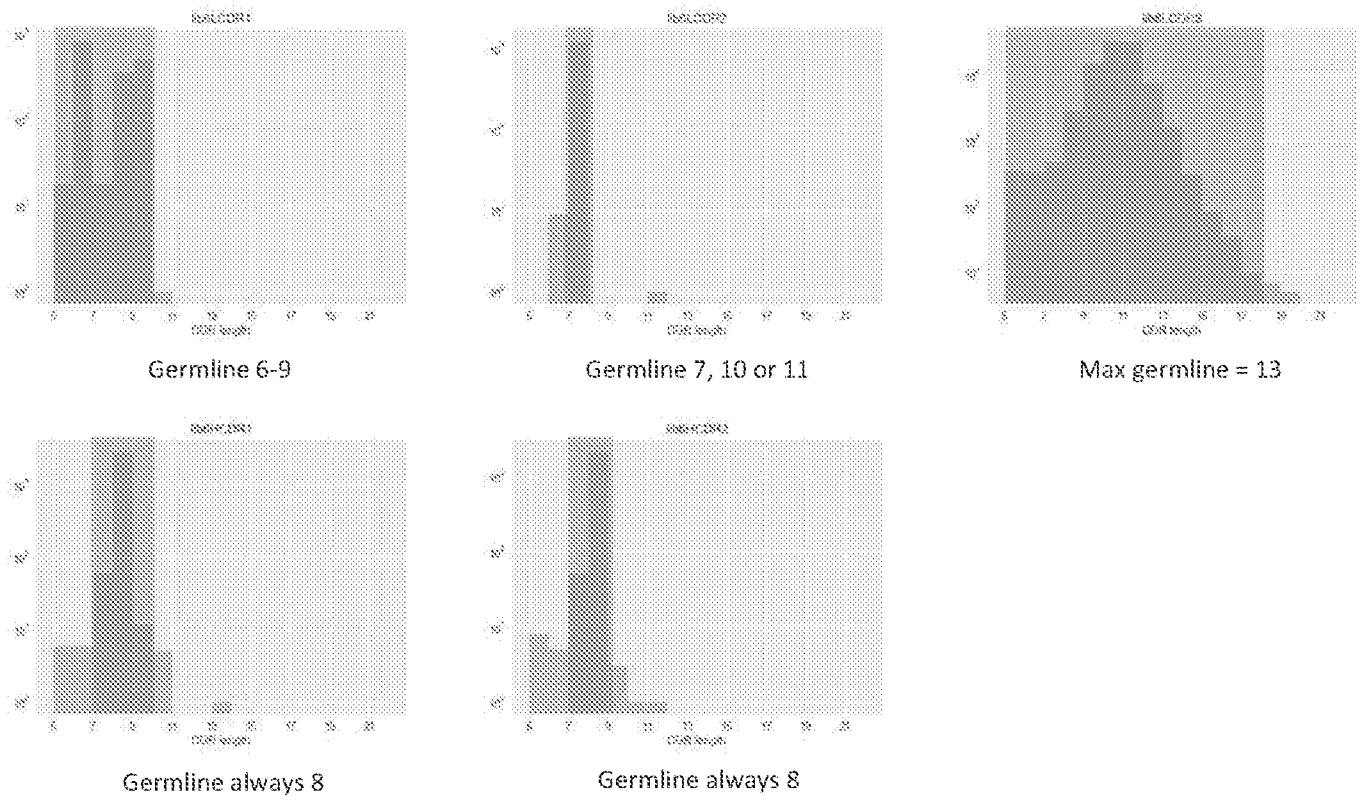
20/58

Figure 20



21/58

Figure 21



22/58

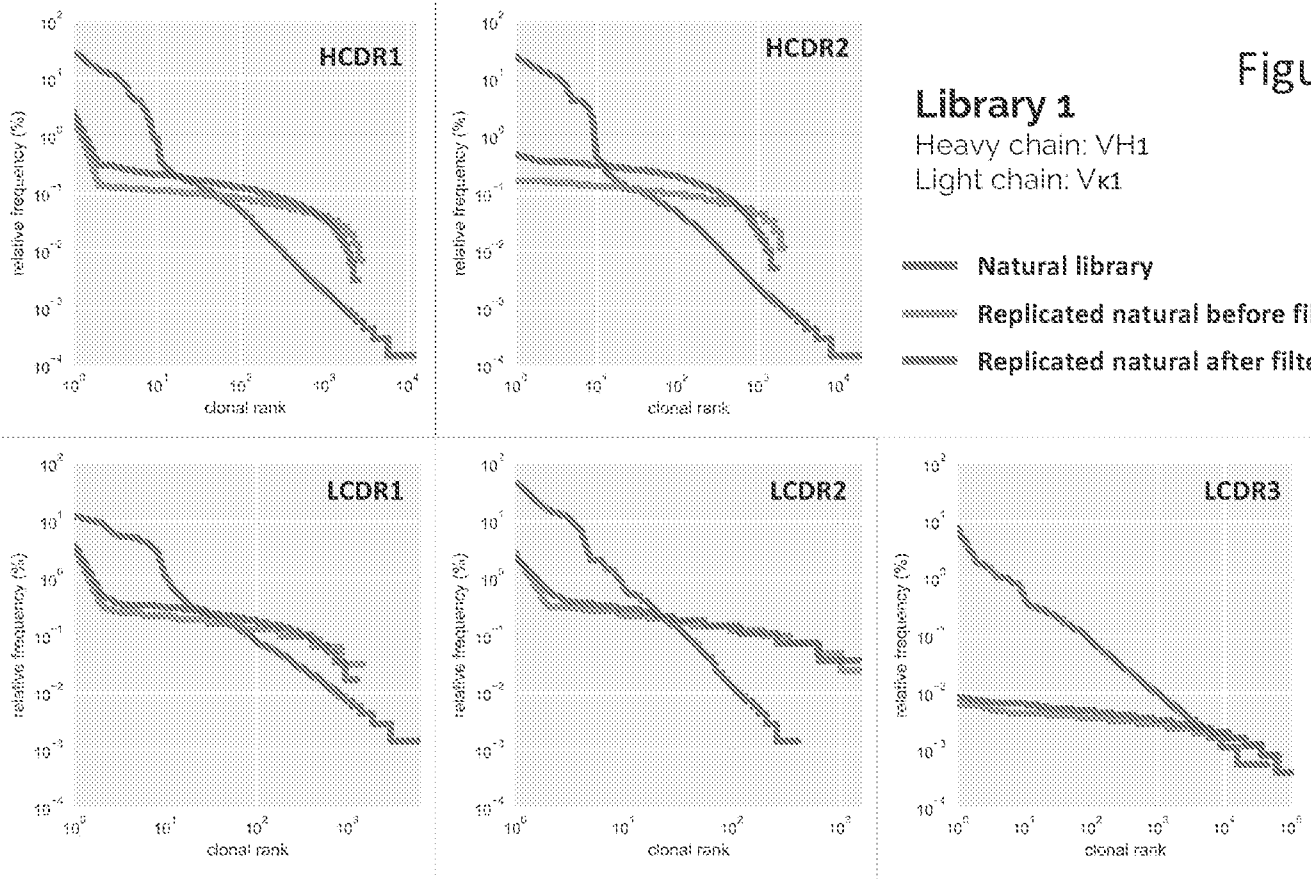
Figure 22A

Library 1

Heavy chain: VH1

Light chain: Vk1

- ▨ Natural library
- ▨ Replicated natural before filtering
- ▨ Replicated natural after filtering



23/58

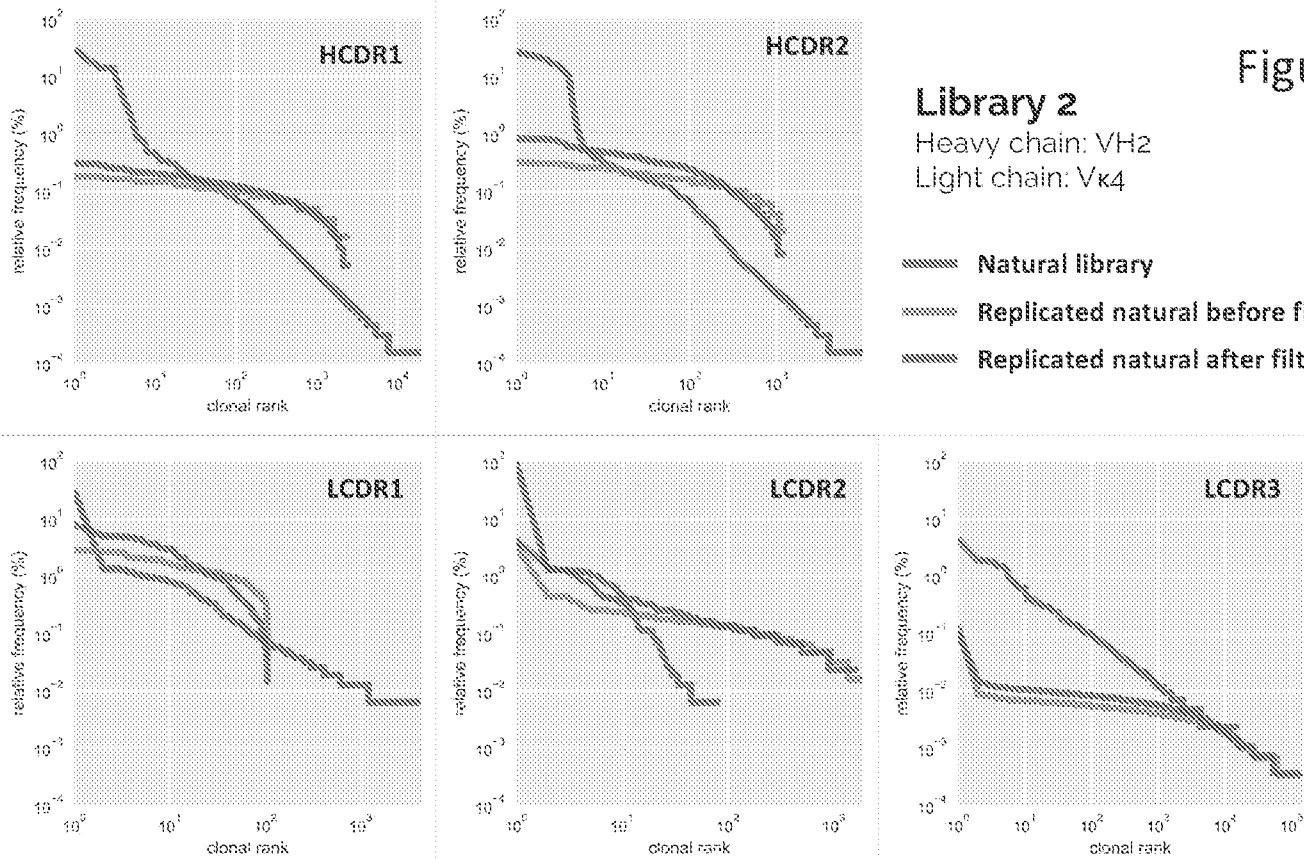
Figure 22B

Library 2

Heavy chain: VH2

Light chain: Vk4

- ▨ Natural library
- ▨ Replicated natural before filtering
- ▨ Replicated natural after filtering



24/58

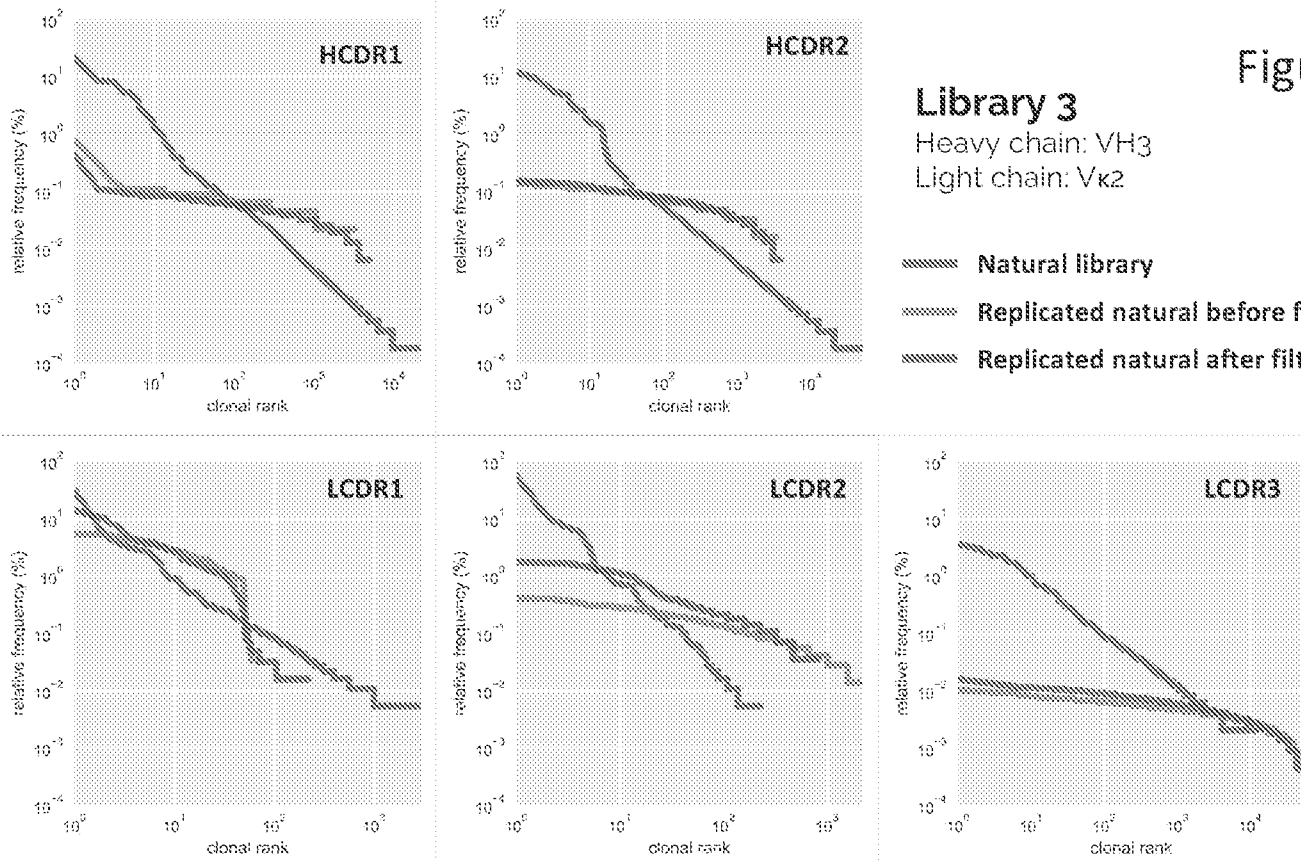
Figure 22C

Library 3

Heavy chain: VH3

Light chain: Vk2

- ▨ Natural library
- ▨ Replicated natural before filtering
- ▨ Replicated natural after filtering



25/58

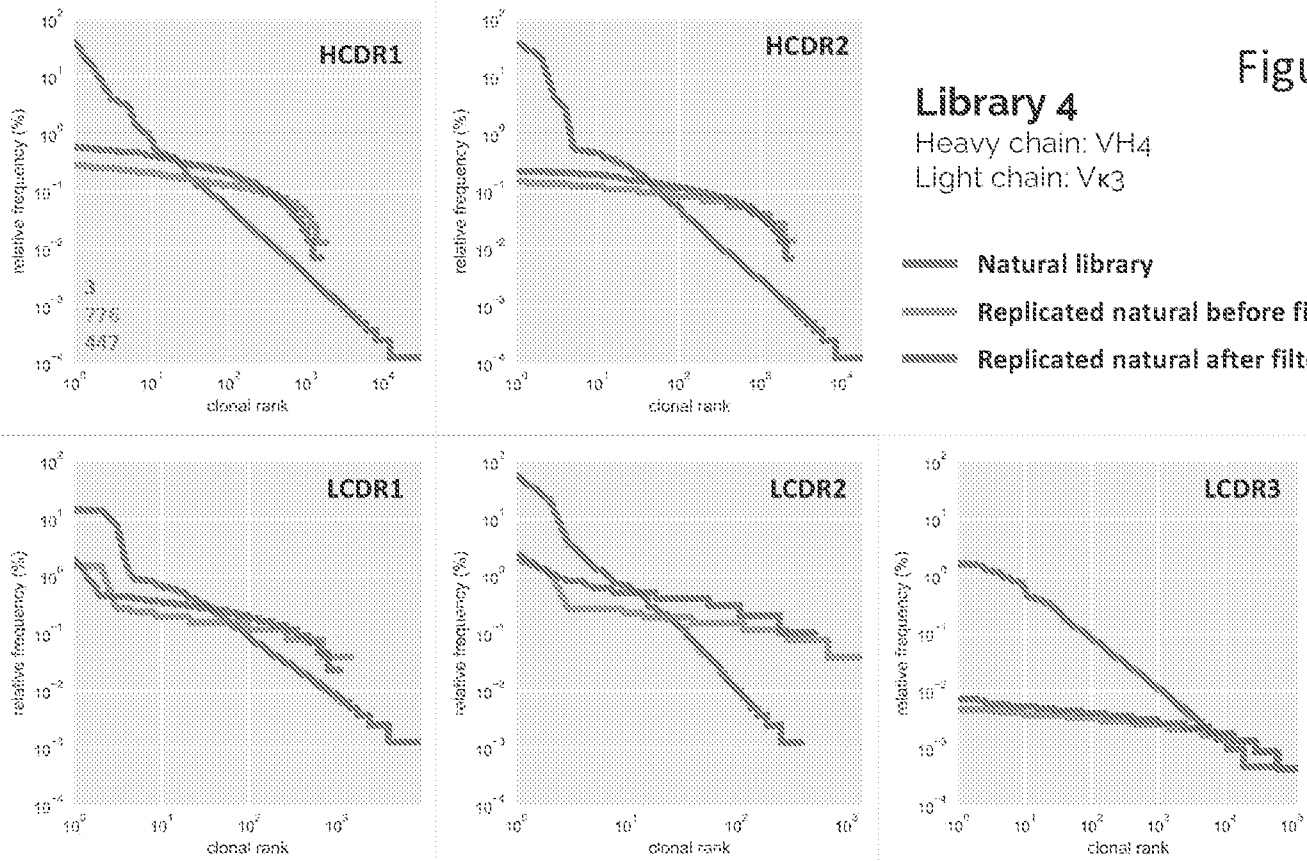
Figure 22D

Library 4

Heavy chain: VH4

Light chain: Vk3

- ▨ Natural library
- ▨ Replicated natural before filtering
- ▨ Replicated natural after filtering



26/58

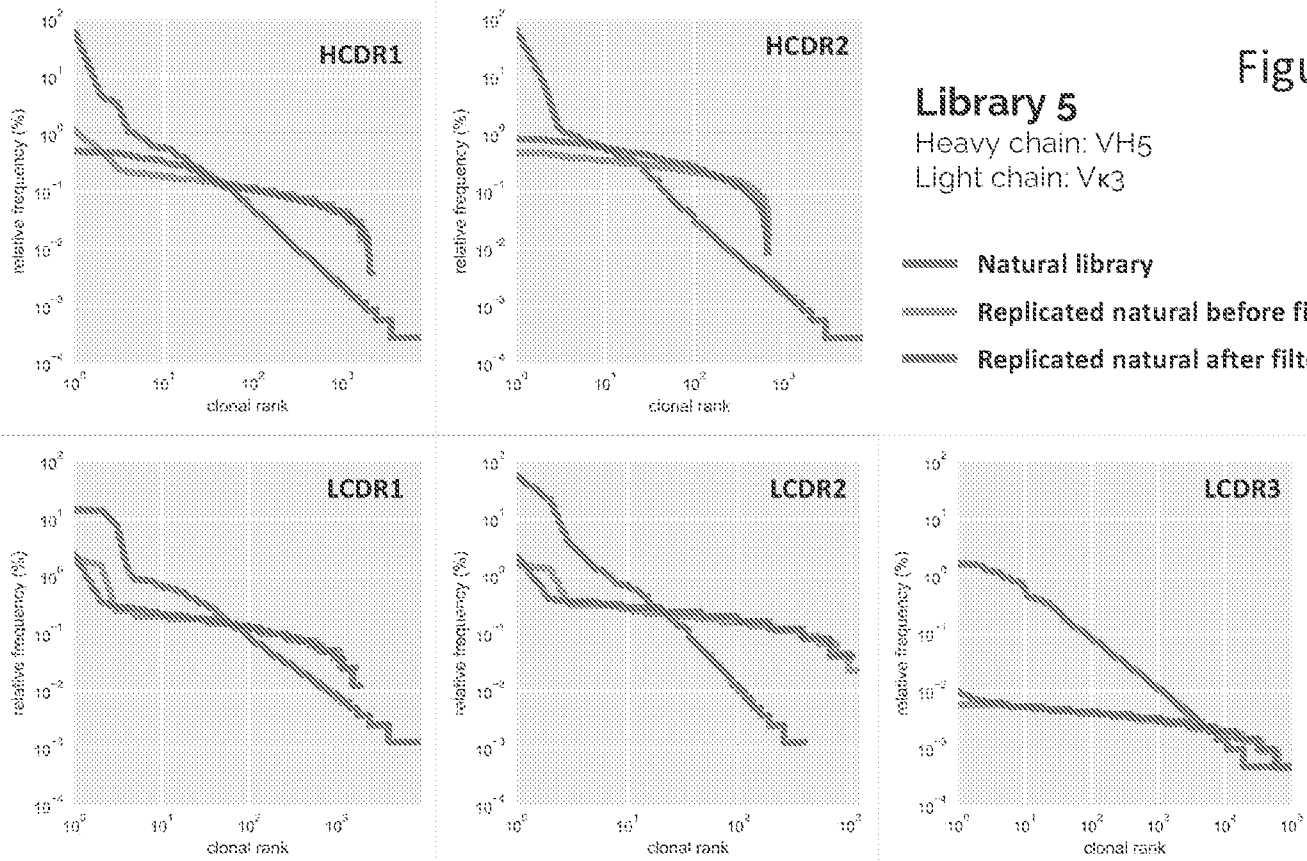
Figure 22E

Library 5

Heavy chain: VH5

Light chain: Vk3

- ▨ Natural library
- ▨ Replicated natural before filtering
- ▨ Replicated natural after filtering



27/58

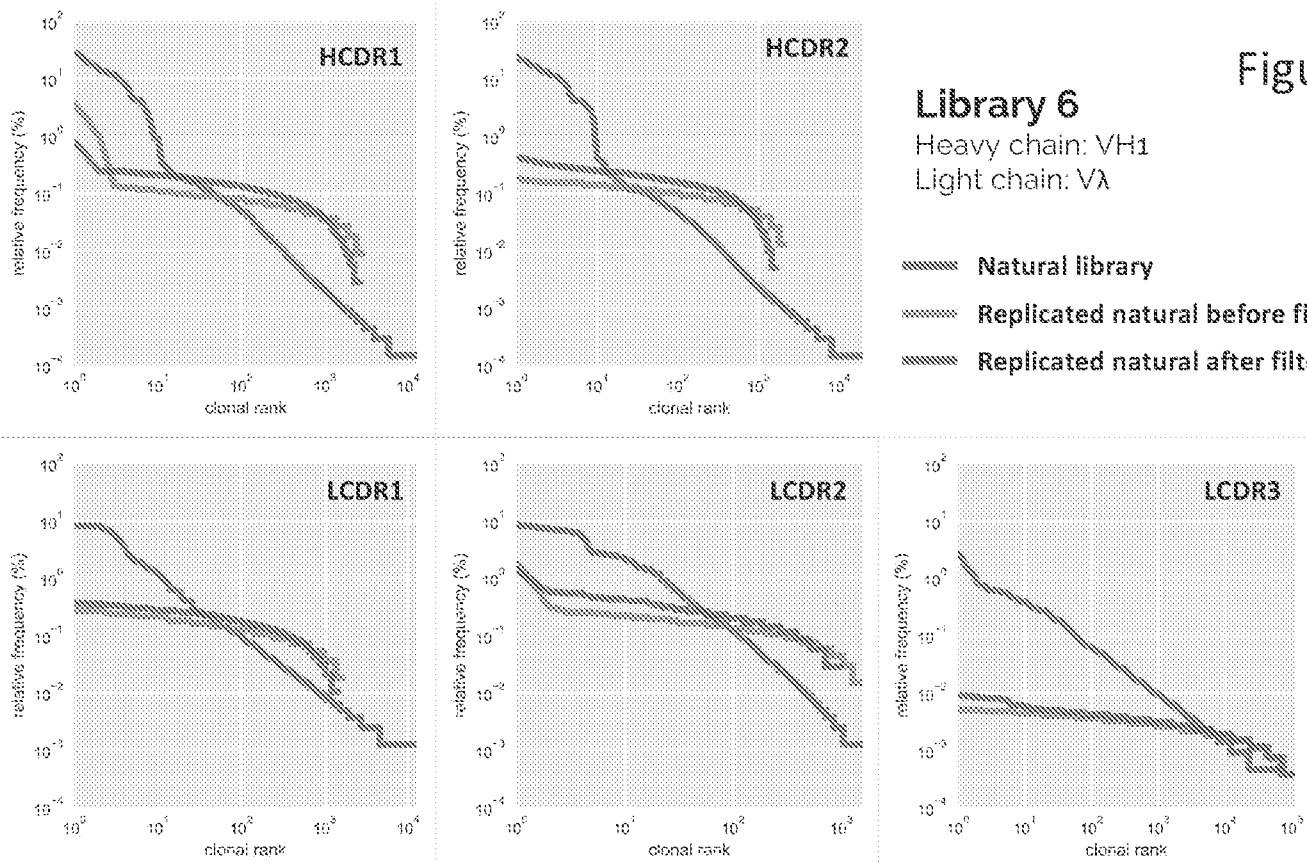
Figure 22F

Library 6

Heavy chain: VH1

Light chain: Vλ

- ▨ Natural library
- ▨ Replicated natural before filtering
- ▨ Replicated natural after filtering



28/58

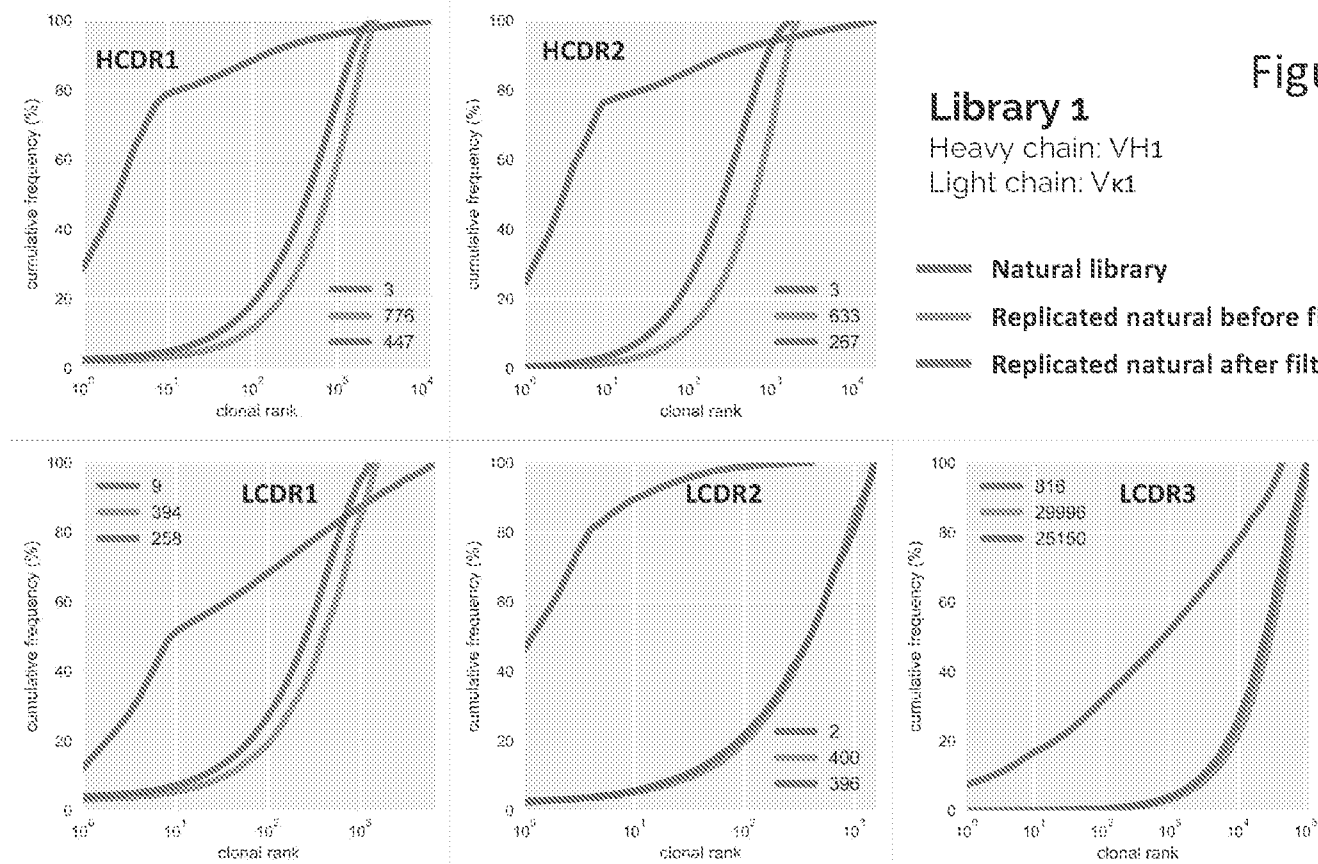
Figure 22G

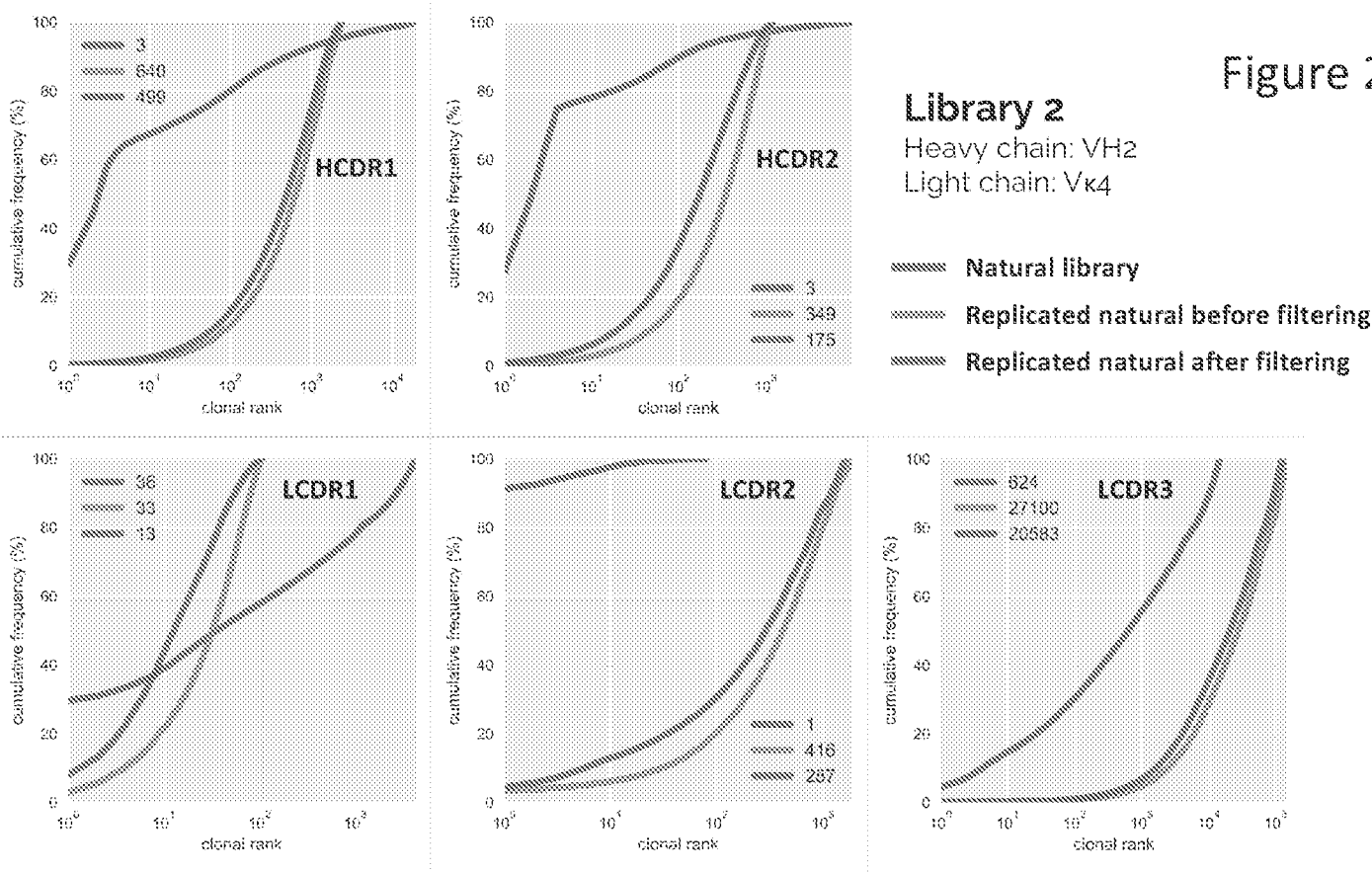
Library 1

Heavy chain: VH1

Light chain: Vk1

- ▬ Natural library
- ▬ Replicated natural before filtering
- ▬ Replicated natural after filtering





30/58

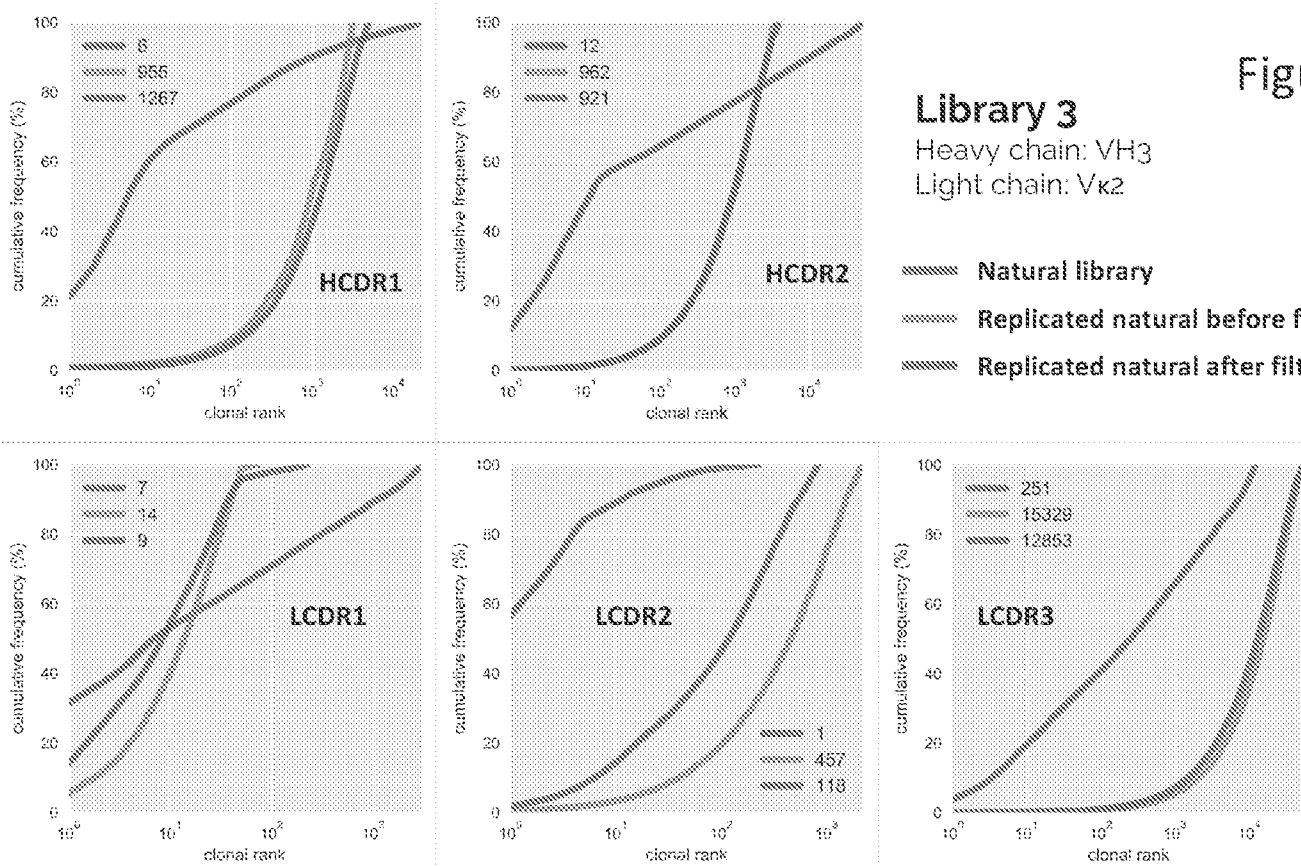
Figure 22I

Library 3

Heavy chain: VH3

Light chain: Vk2

- ▨ Natural library
- ▨ Replicated natural before filtering
- ▨ Replicated natural after filtering



31/58

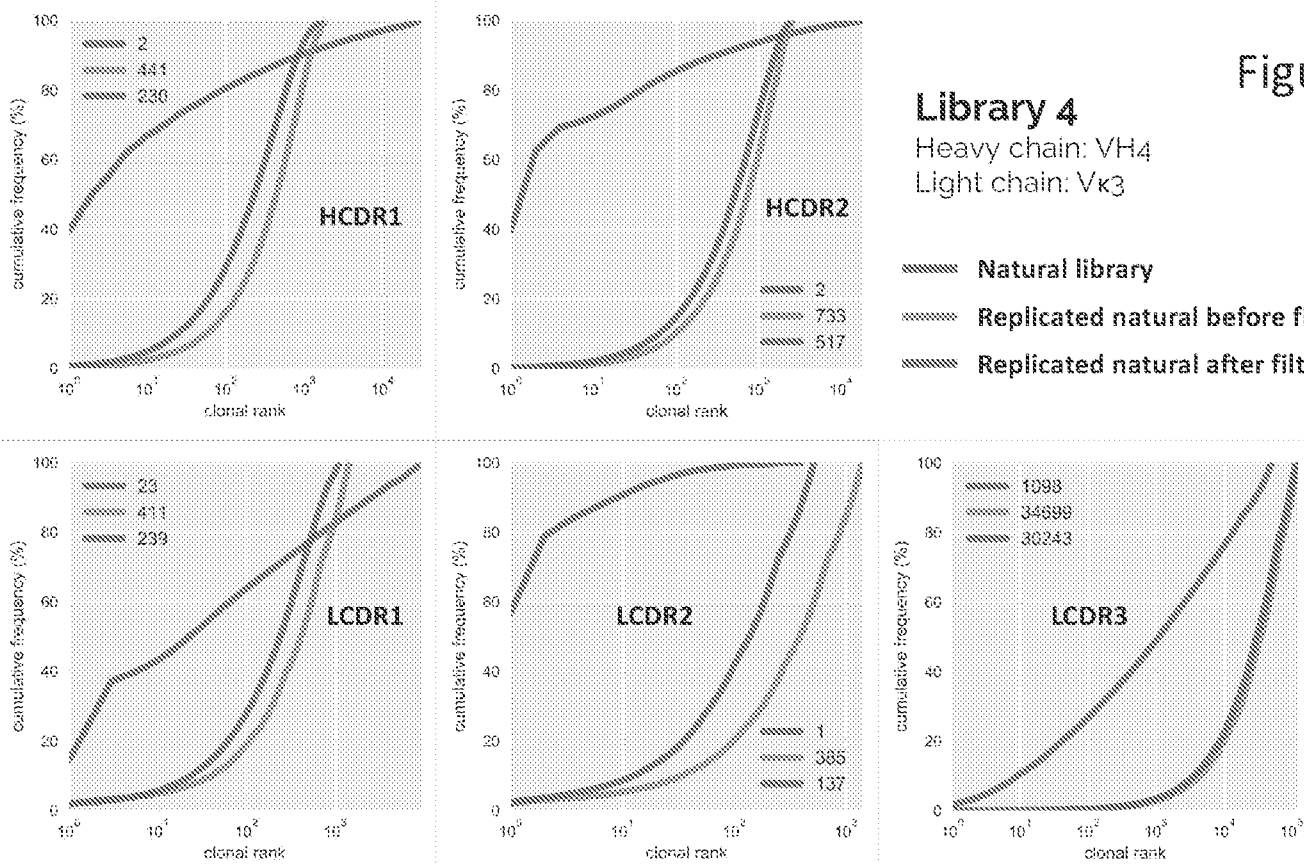
Figure 22J

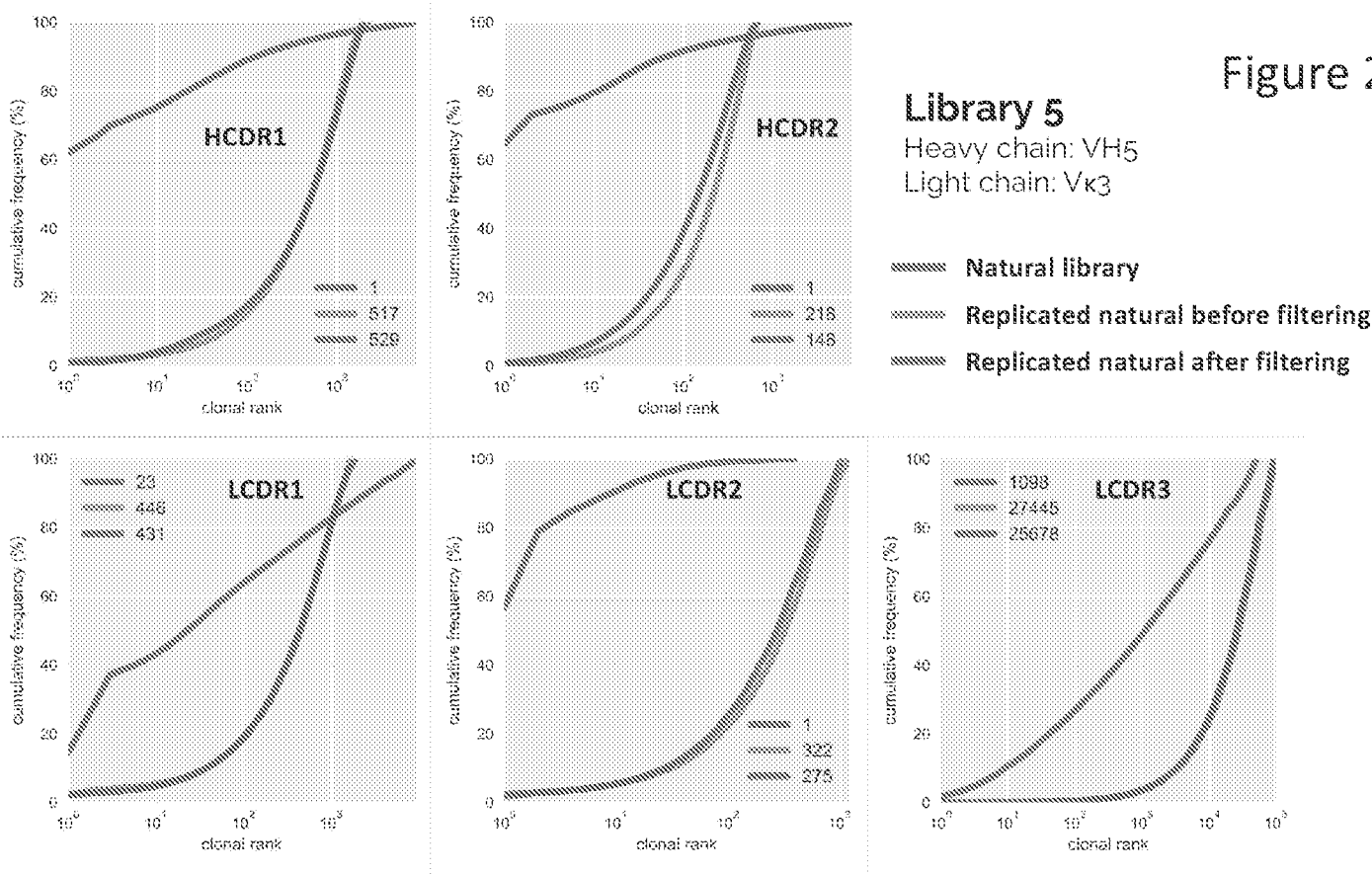
Library 4

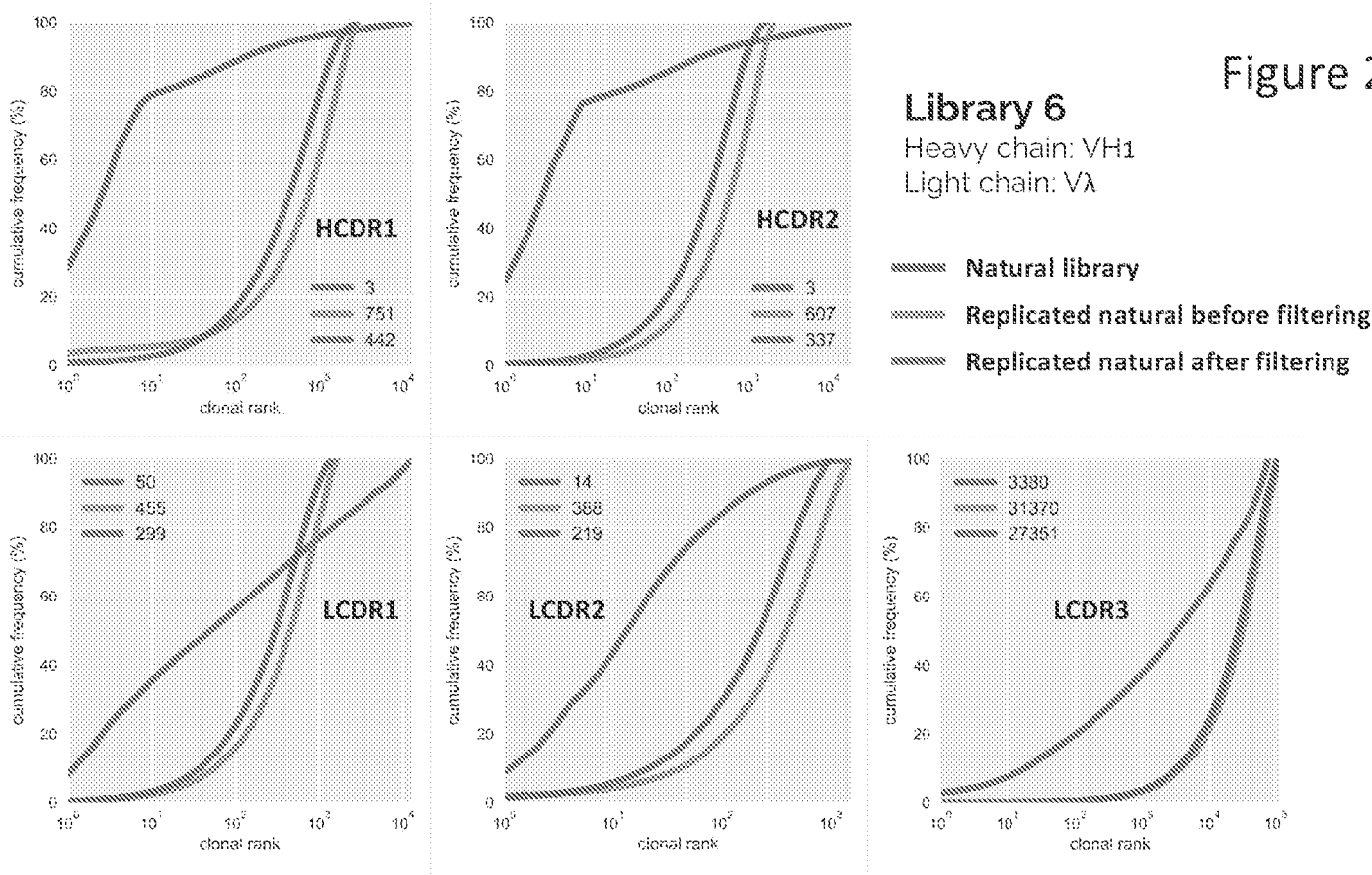
Heavy chain: VH4

Light chain: Vk3

- ▬ Natural library
- ▬ Replicated natural before filtering
- ▬ Replicated natural after filtering

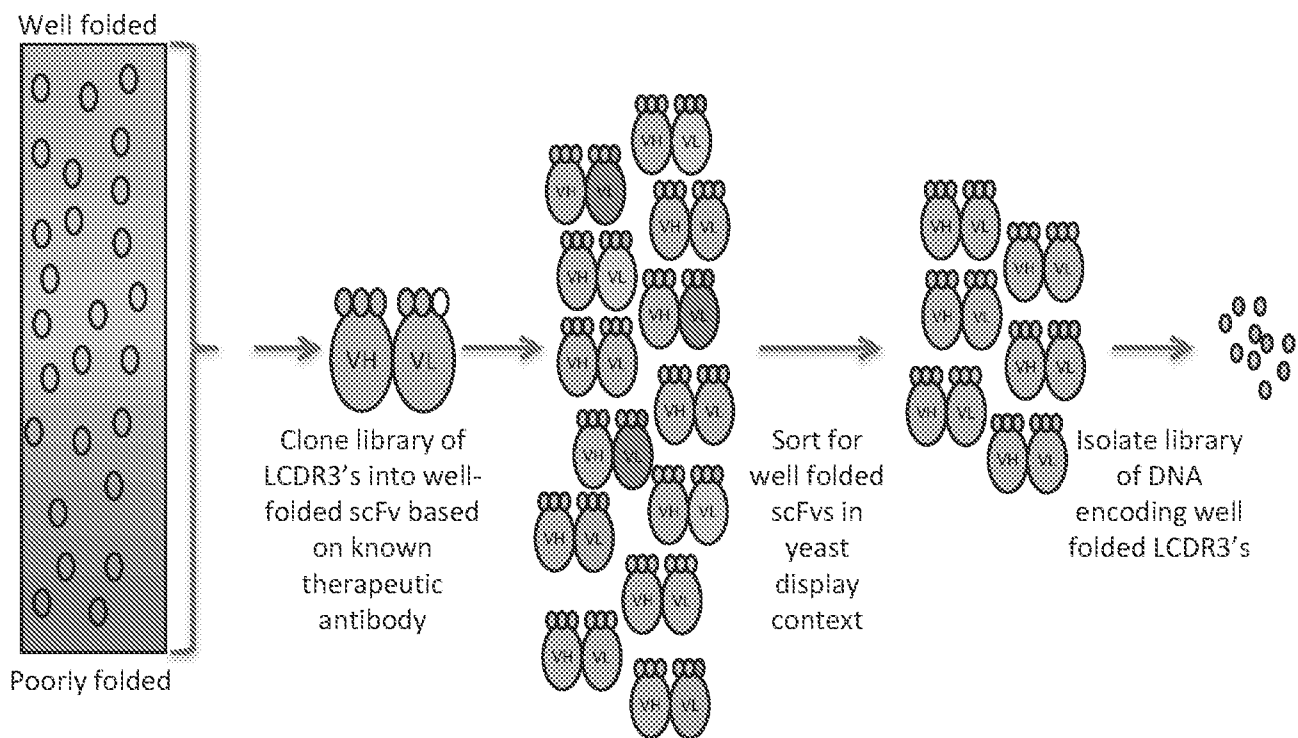






34/58

Figure 23



35/58

Figure 24

GGTGATCGTGTTACTATTACCTGTCGTGCTTCTAGGAACCATGGCCAGTAAGGCGGTCCTCTGGCTGGTACCAGCAGAAACCAGGTAAAGCTC
 CCACTAGCACAAATGATAATGGACAGCACGAAGAATCTGGTACTGGTCATTTCGTCAGAGAGAACCGAAACATGGTCGTCTTTGGTCCATTTCGAG

Bsal SfiI Bsal

Cleave with Bsal and SfiI

Yeast display vector cleaved at LCDR1

GGTGATCGTGTTACTATTACCTGTCGTGCTTCTCTGGCTTGGTACCAGCAGAAACCAGGTAAAGCTC
 CCACTAGCACAAATGATAATGGACAGCACGAAGA GAACCATGGTCGTCTTTGGTCCATTTCGAG

X Homologous recombination X

GGTGATCGTGTTACTATTACCTGTCGTGCTTCT-----LCDR1-----CTGGCTTGGTACCAGCAGAAACCAGGTAAAGCTC
 CCACTAGCACAAATGATAATGGACAGCACGAAGA-----LCDR1-----GACCGAACCATGGTCGTCTTTGGTCCATTTCGAG

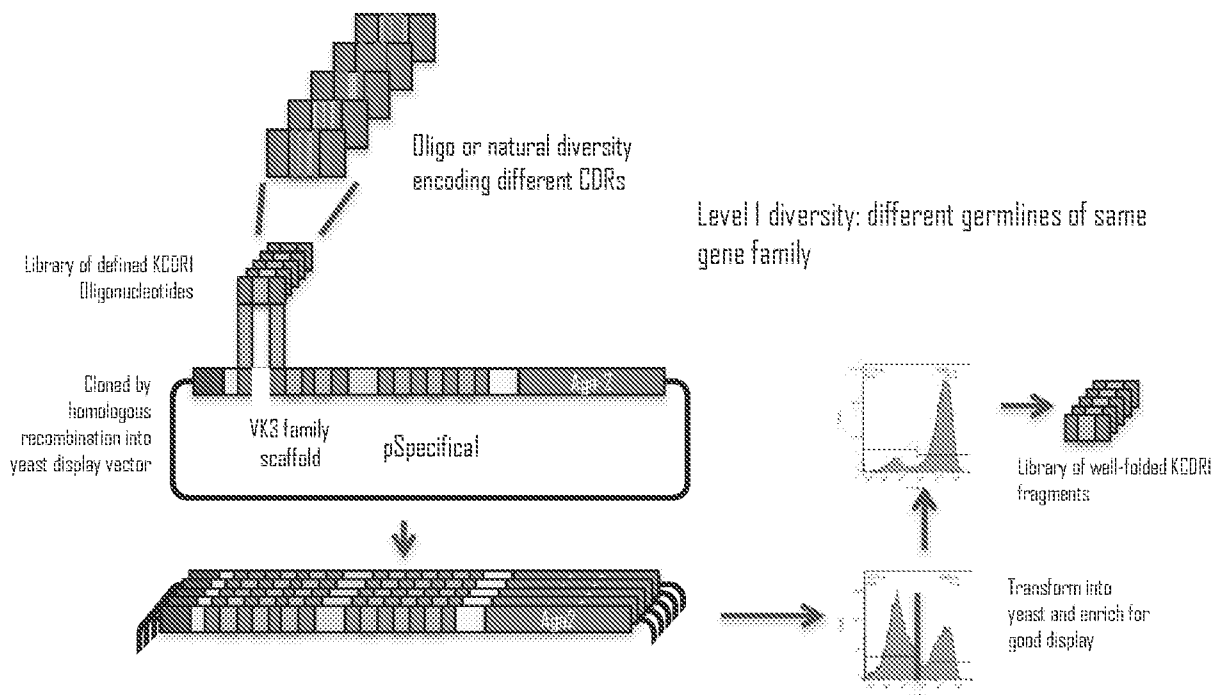
Amplified LCDR1 oligo

Amplified LCDR1 inserted into yeast display vector

GGTGATCGTGTTACTATTACCTGTCGTGCTTCT-----LCDR1-----CTGGCTTGGTACCAGCAGAAACCAGGTAAAGCTC
 CCACTAGCACAAATGATAATGGACAGCACGAAGA-----LCDR1-----GACCGAACCATGGTCGTCTTTGGTCCATTTCGAG

36/58

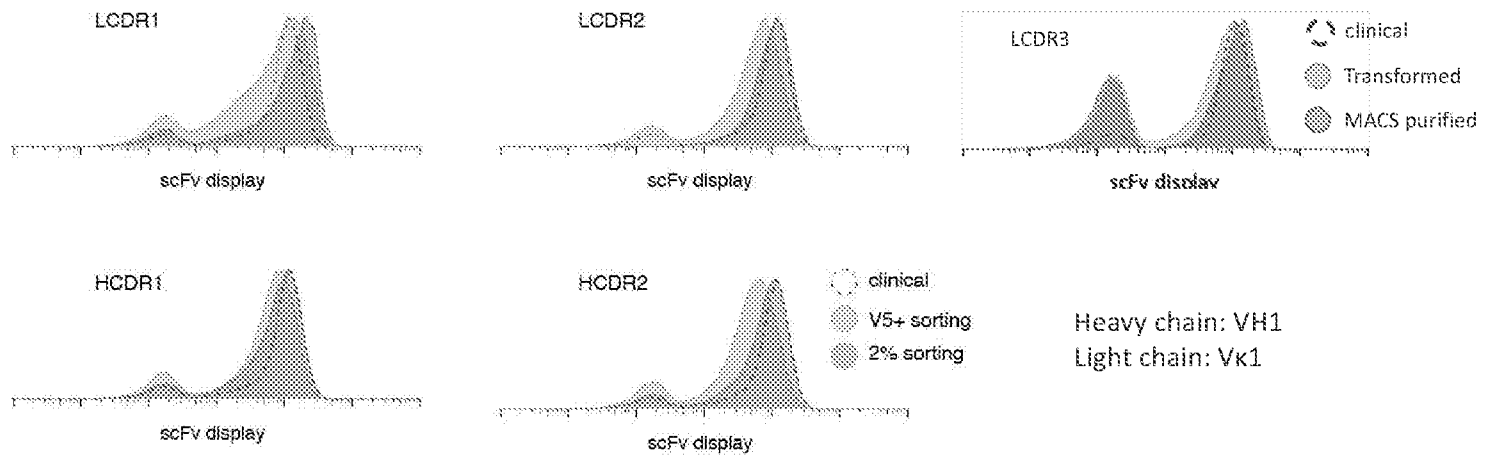
Figure 25



37/58

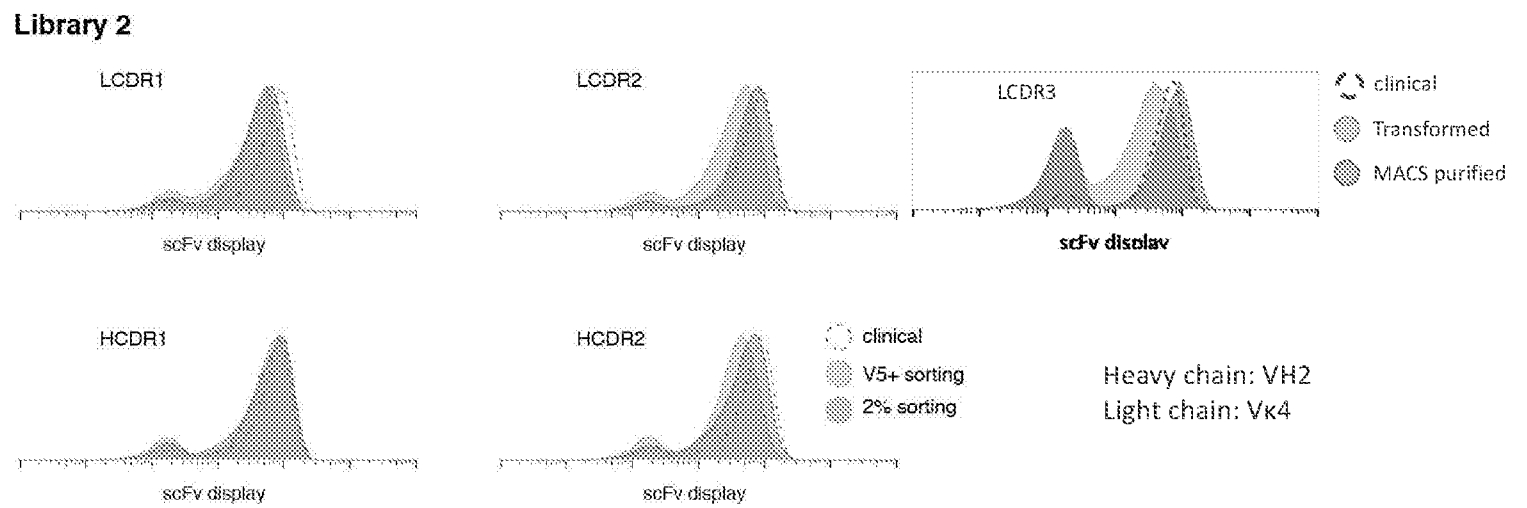
Figure 25B

Library 1



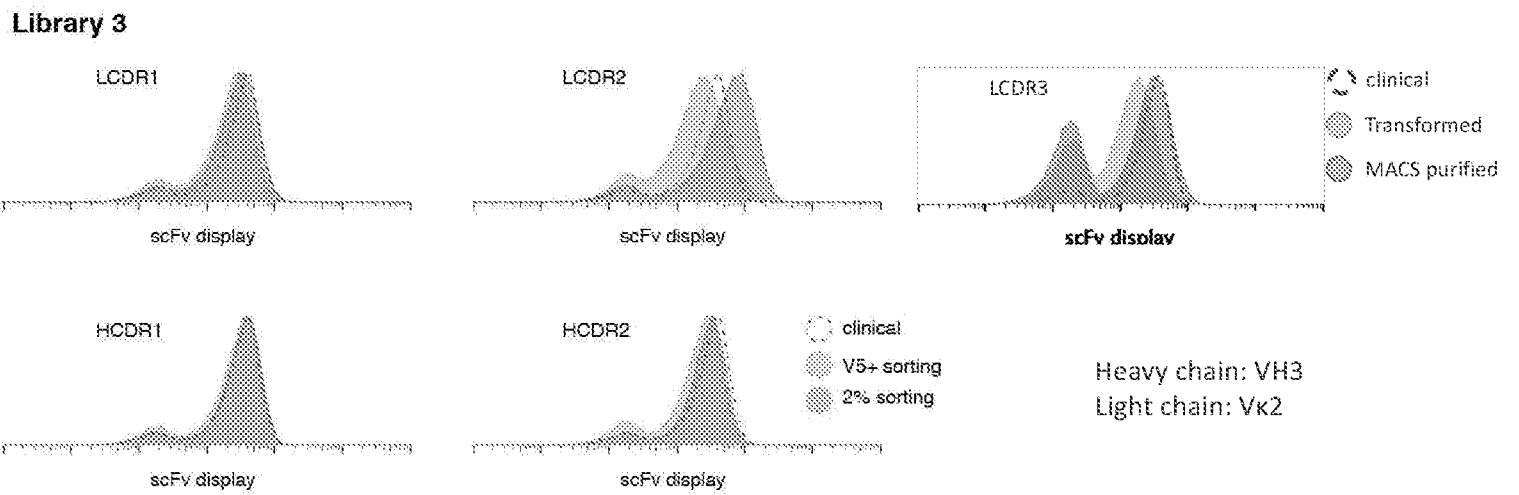
38/58

Figure 25C



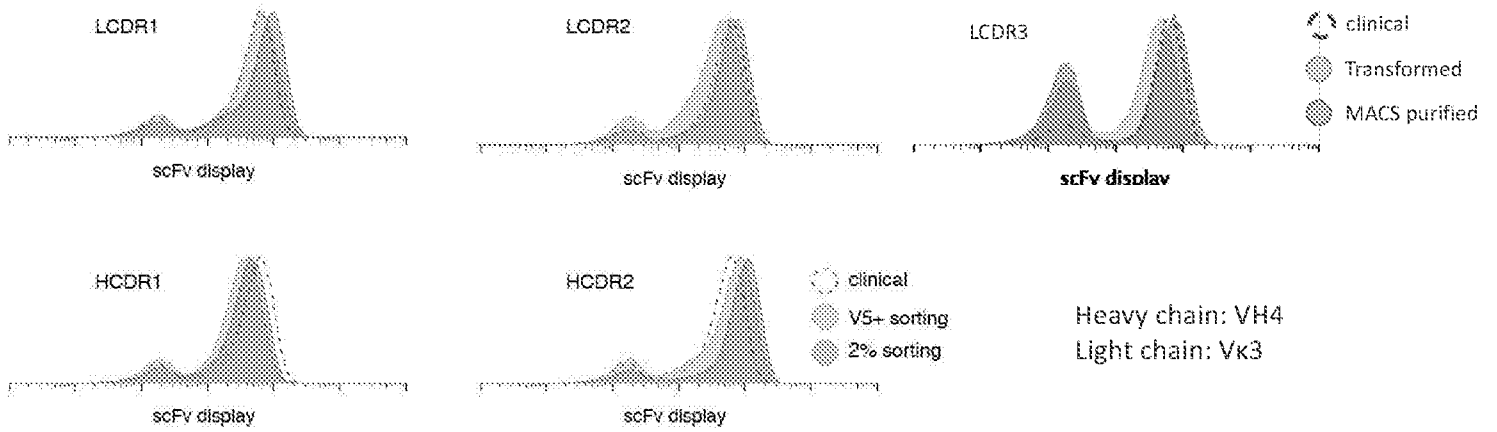
39/58

Figure 25D



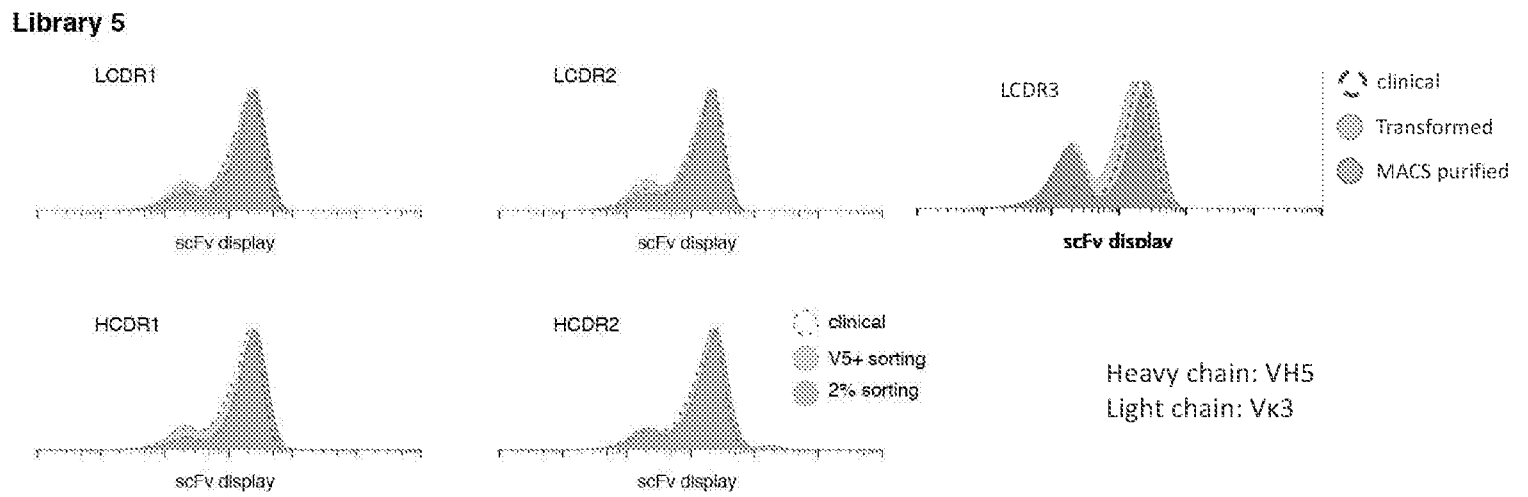
40/58

Figure 25E

Library 4

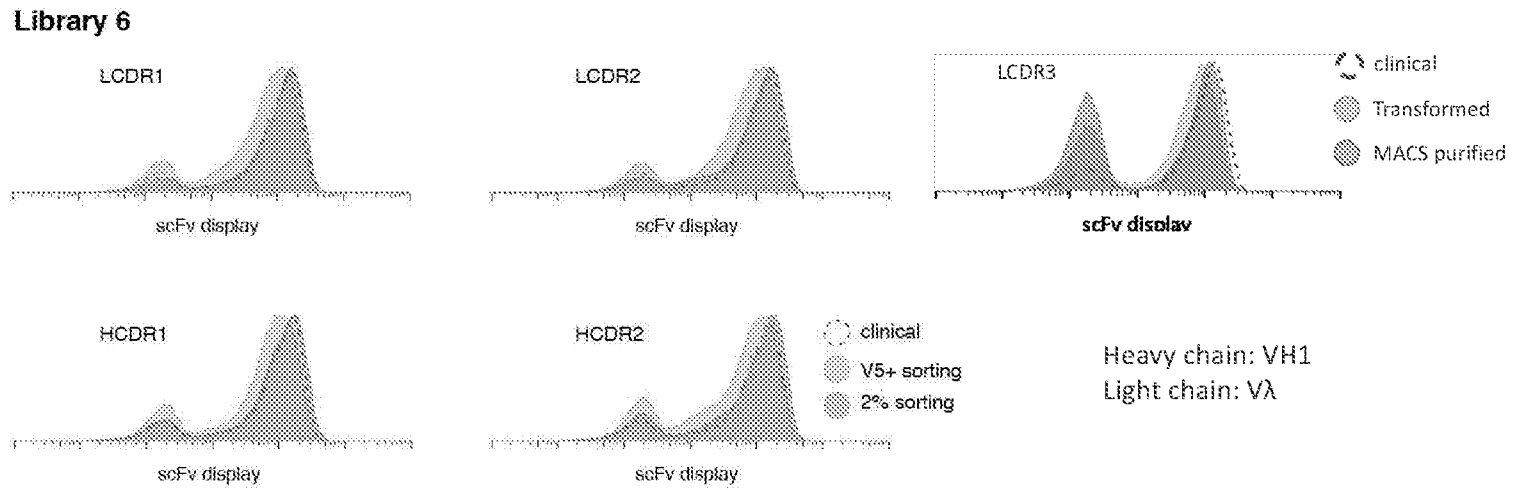
41/58

Figure 25F



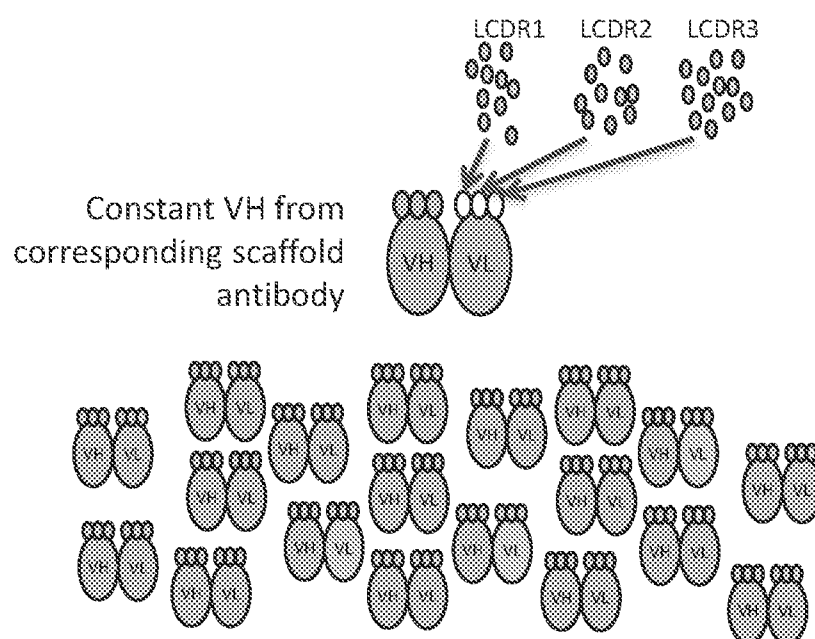
42/58

Figure 25G



43/58

Figure 26



44/58

Figure 27

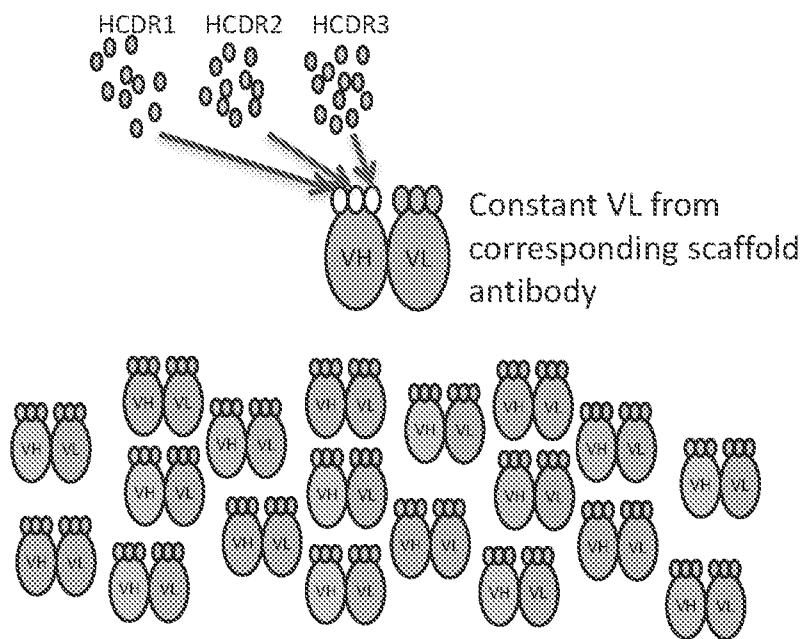
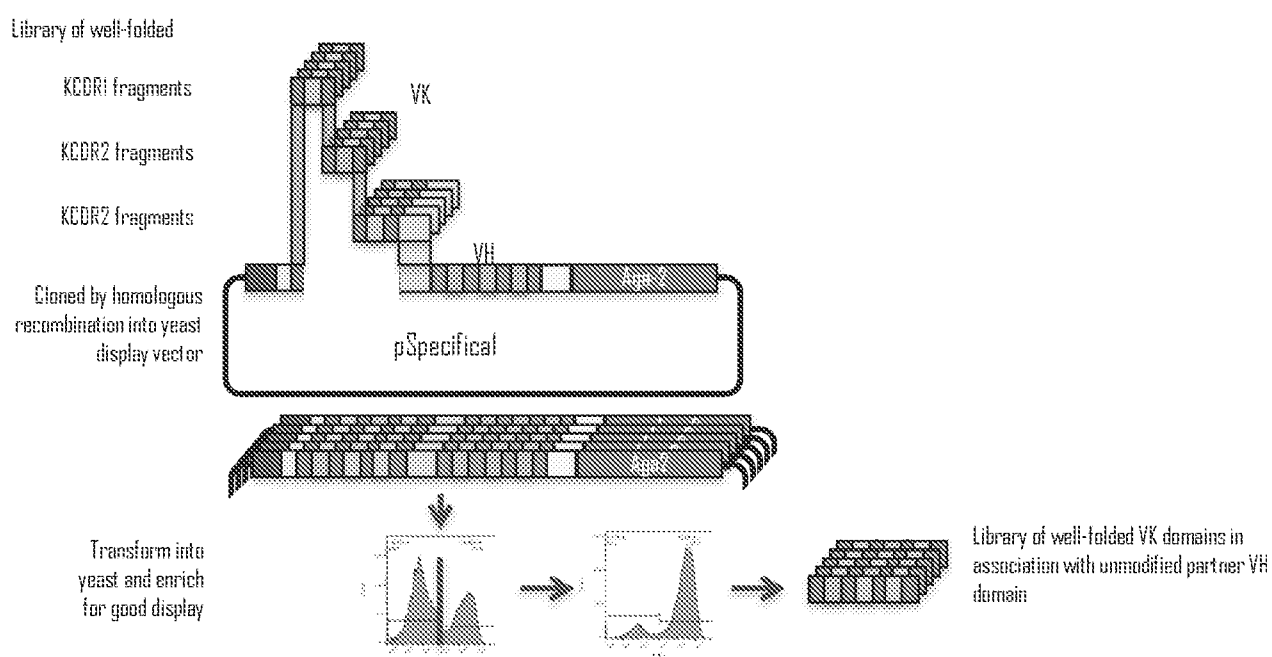


Figure 29: Assembling functional antibodies, e.g. VH libraries

45/58

Figure 28

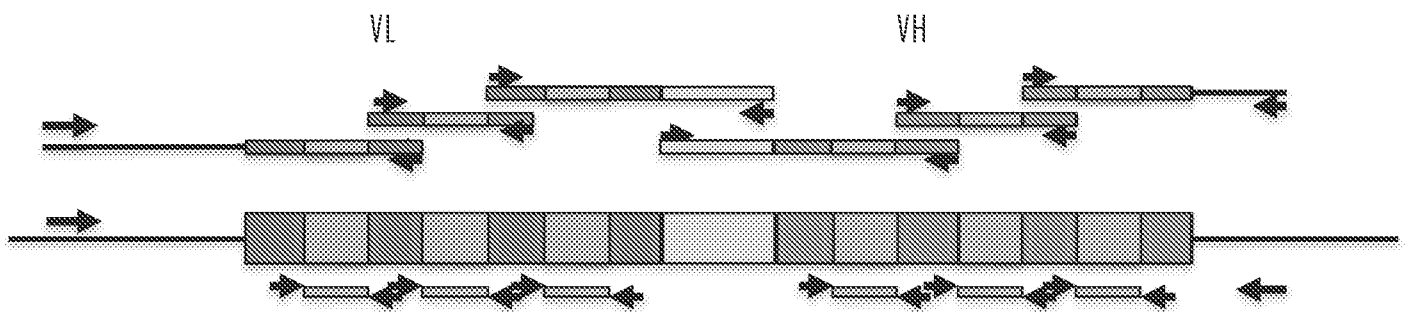


46/58

Figure 29

Assembly amplification 8 primers/library (48 + 4 primers)

- Each product has maximum possible overlap
- Matching Tm
 - Have to keep flanking sequence
 - Have to have homology for recombination

**Construct amplification (2 primers)**

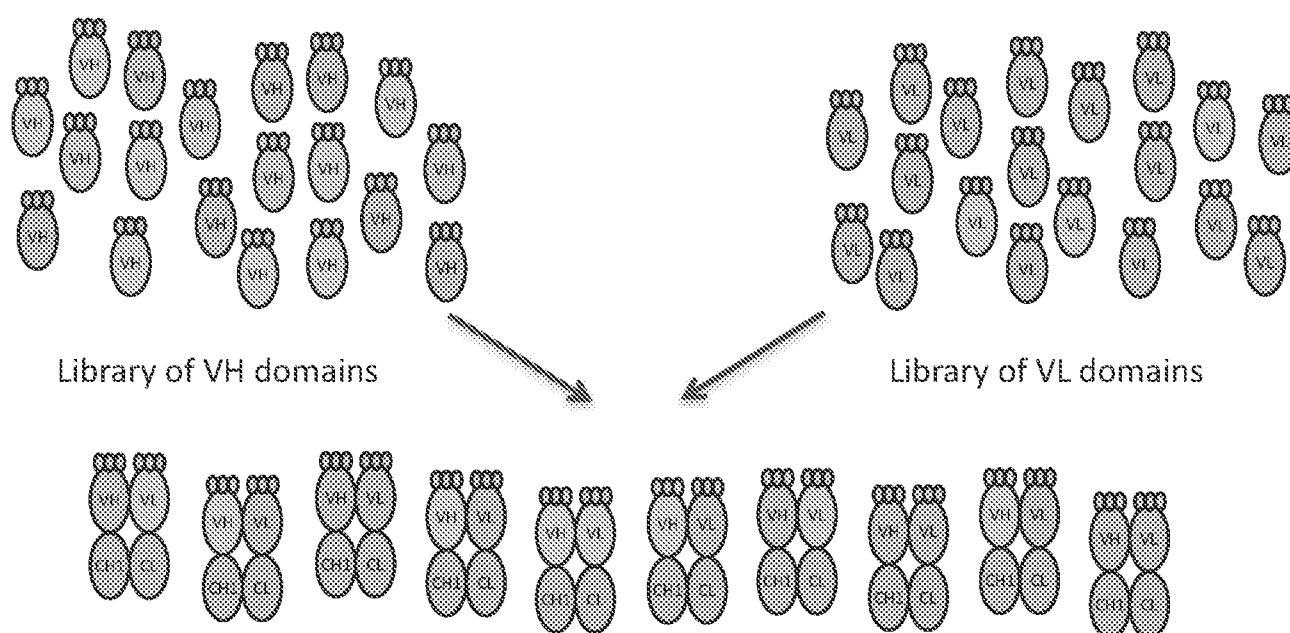
- Matching Tm
- Works for all constructs

CDR amplification 12 primers/library (72 primers)

- 42 nt each
- No matching Tm
 - Have to keep flanking sequence
 - Have to have homology for recombination

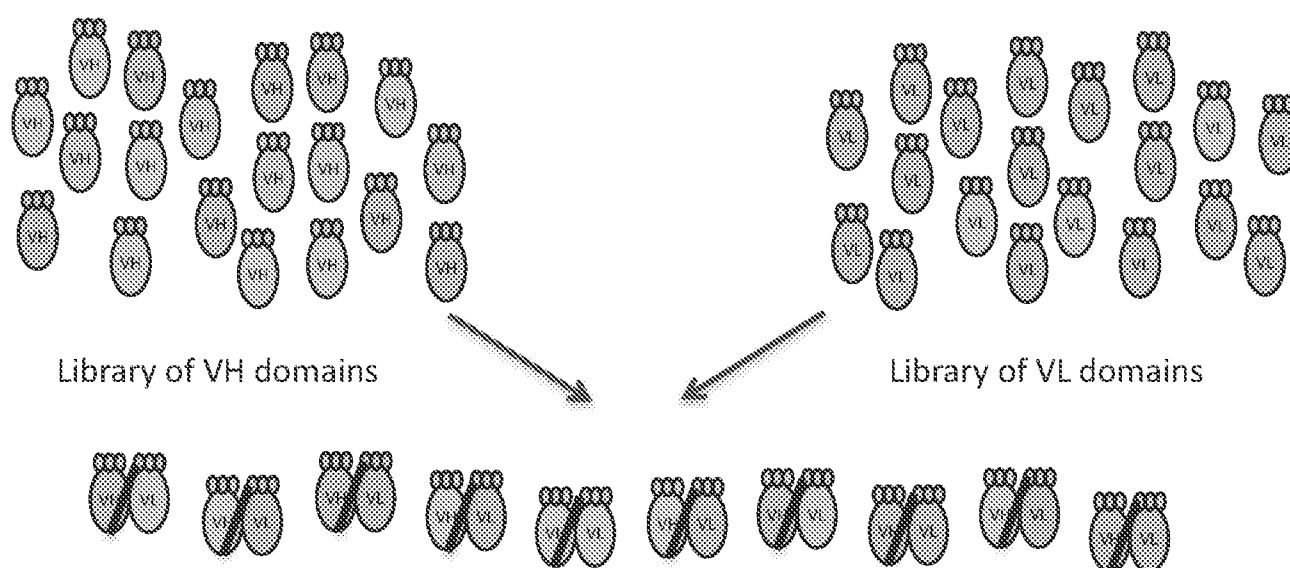
47/58

Figure 30



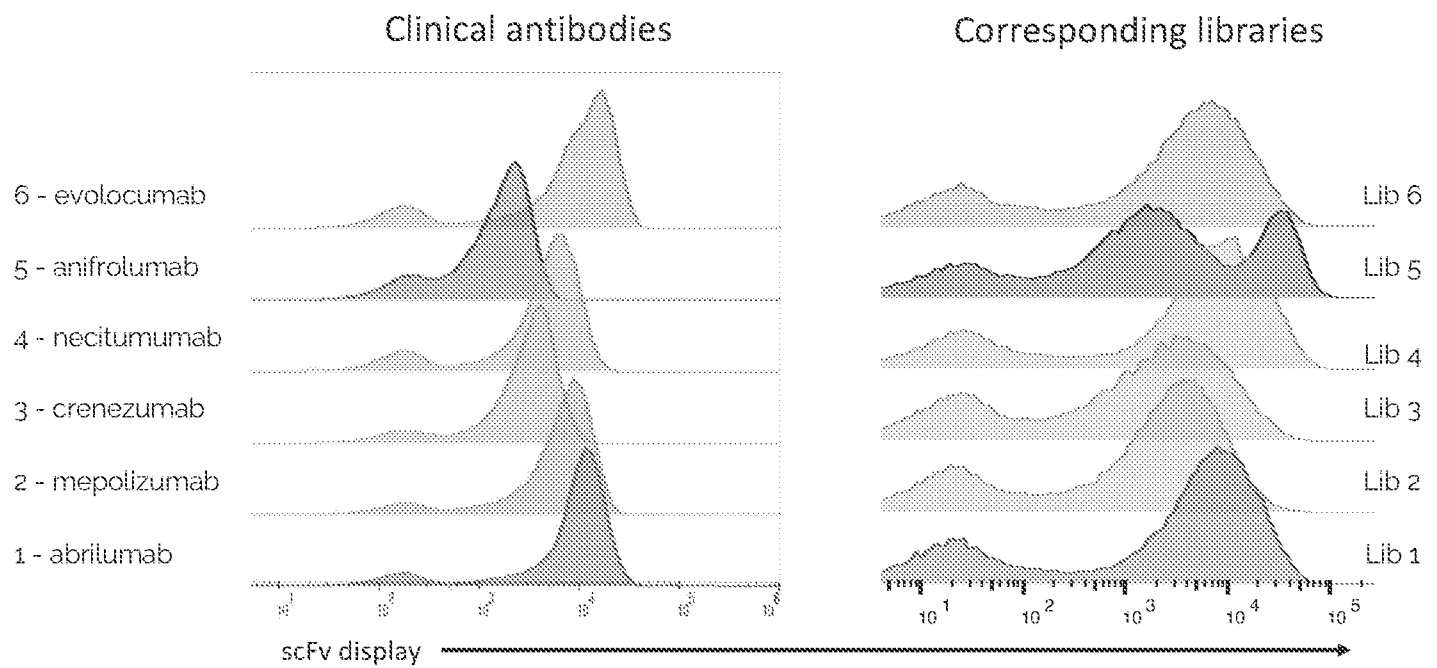
48/58

Figure 31



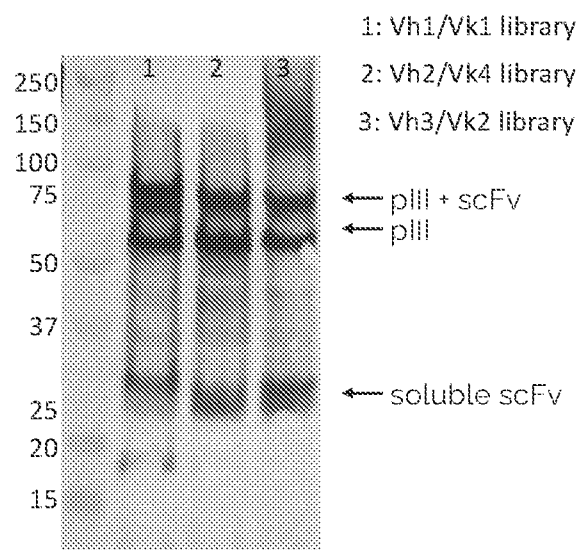
49/58

Figure 32

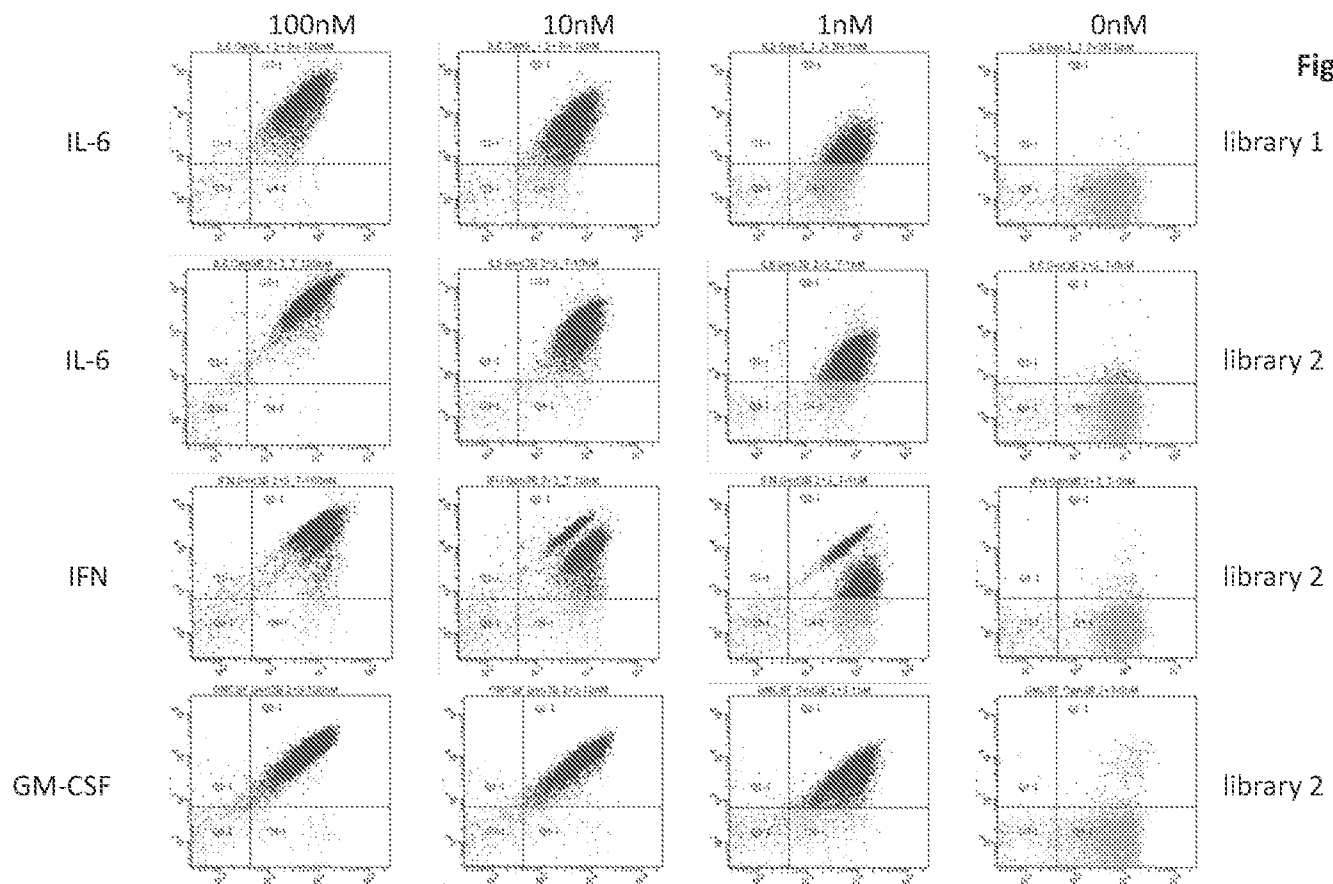


50/58

Figure 33



51/58



52/58

Figure 35

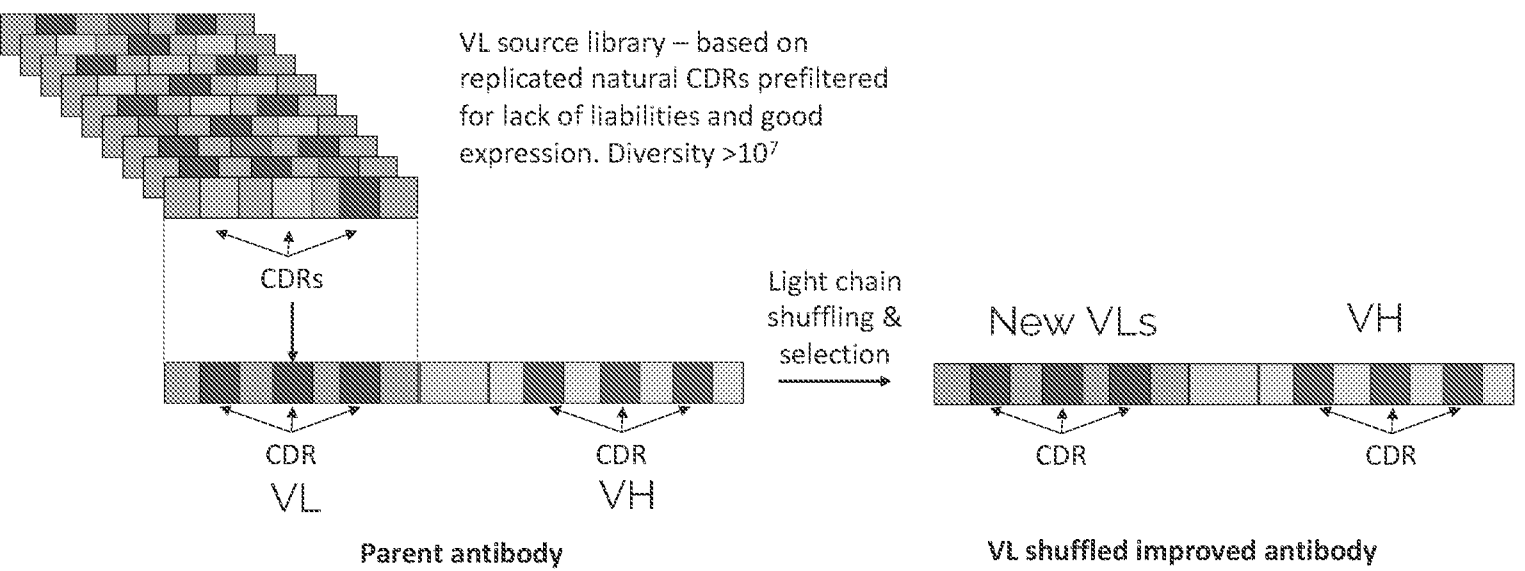
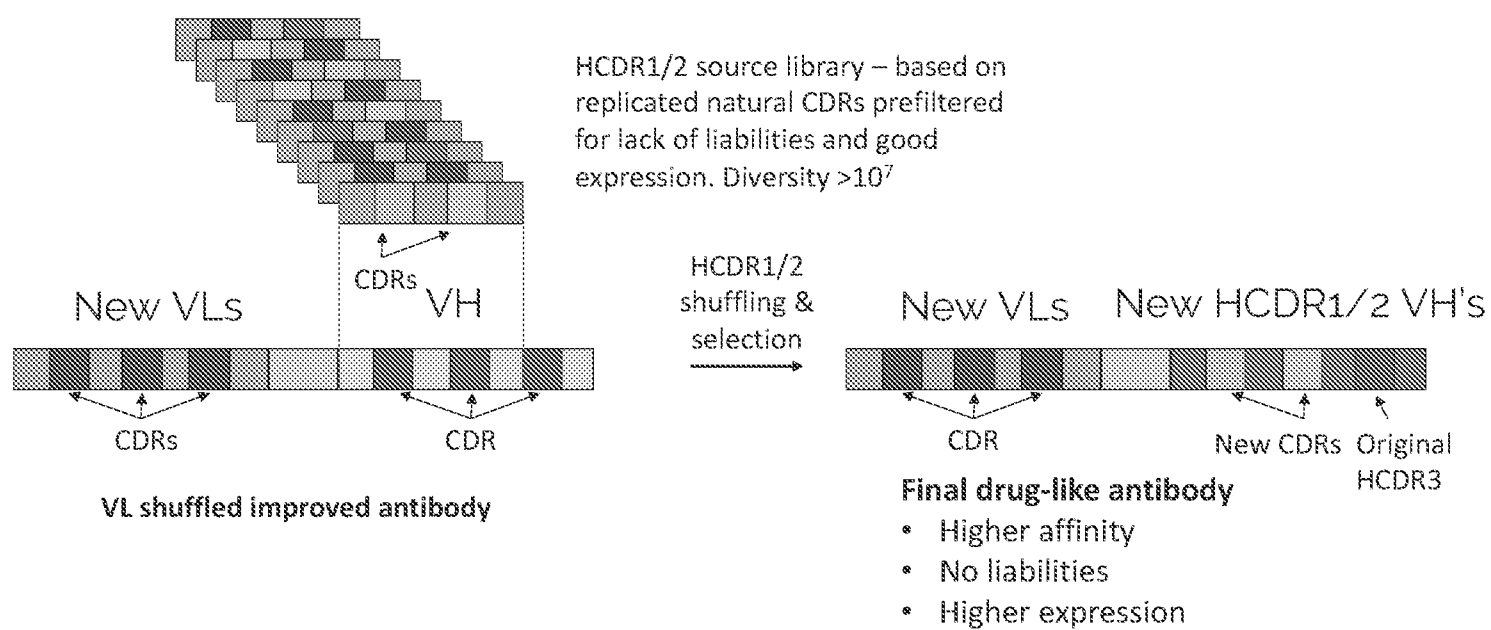


Figure 36



HCDR1/2
shuffling &
selection

Final drug-like antibody

- Higher affinity
- No liabilities
- Higher expression

Figure 37

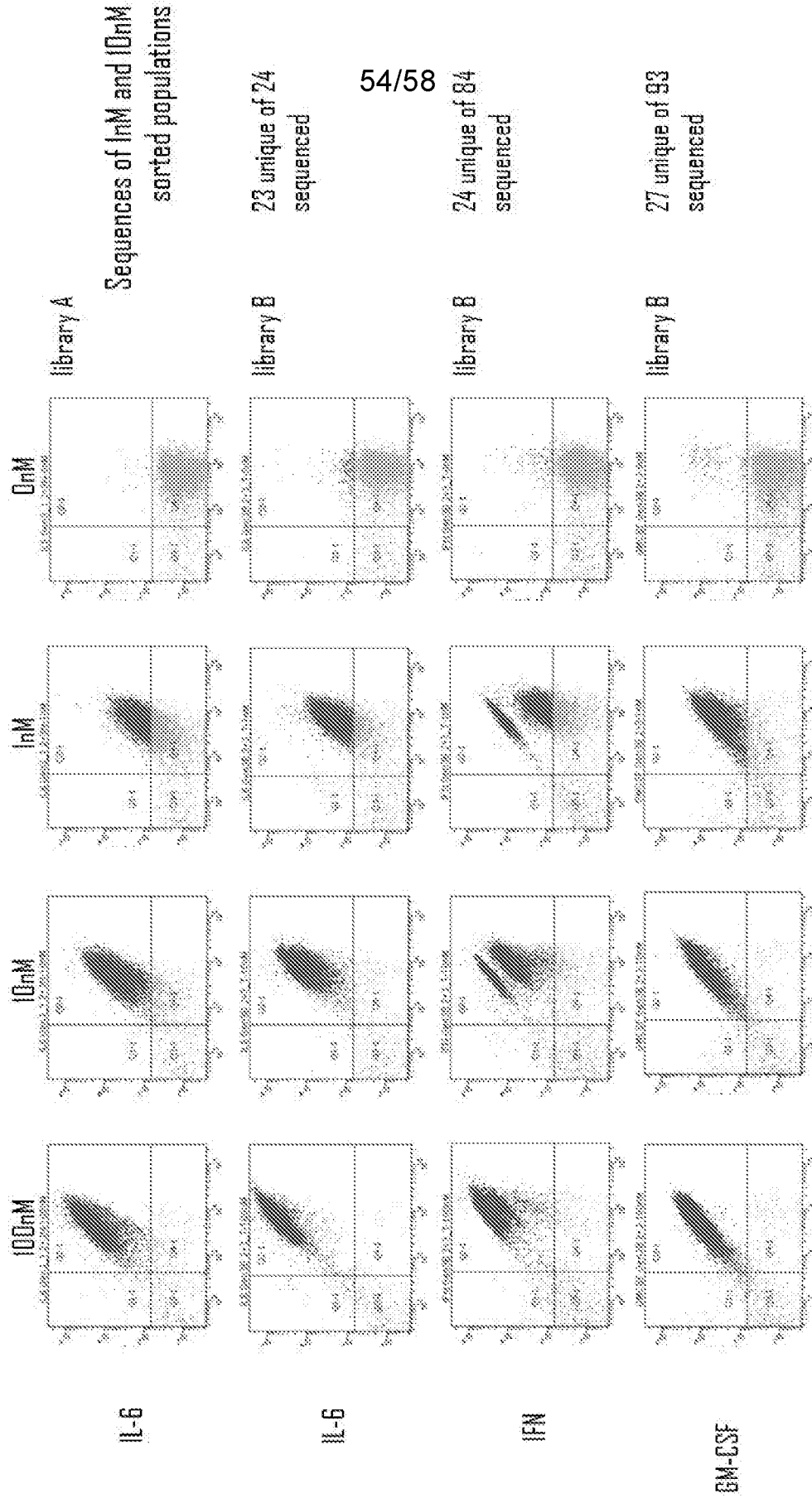


Figure 38A

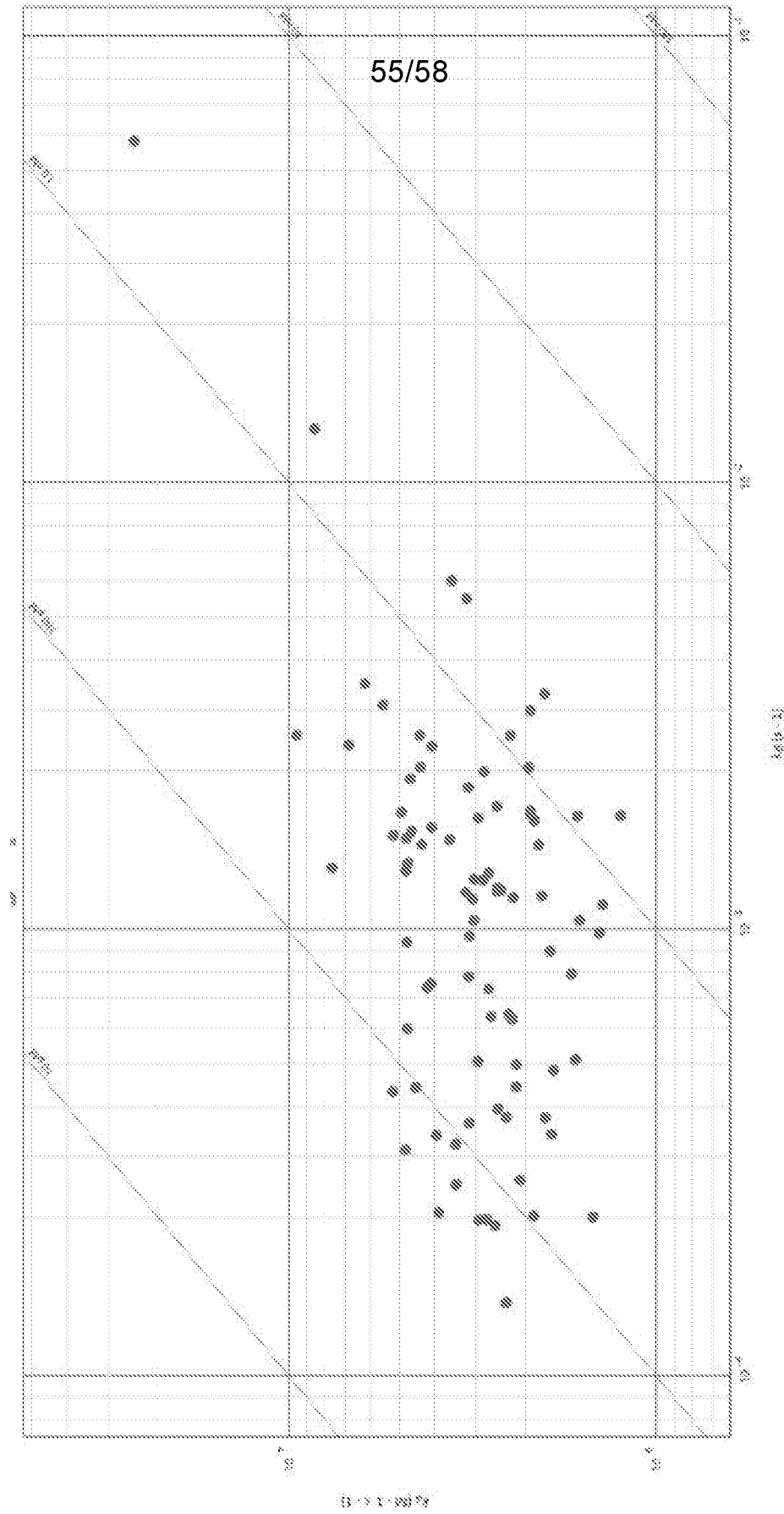


Figure 38B

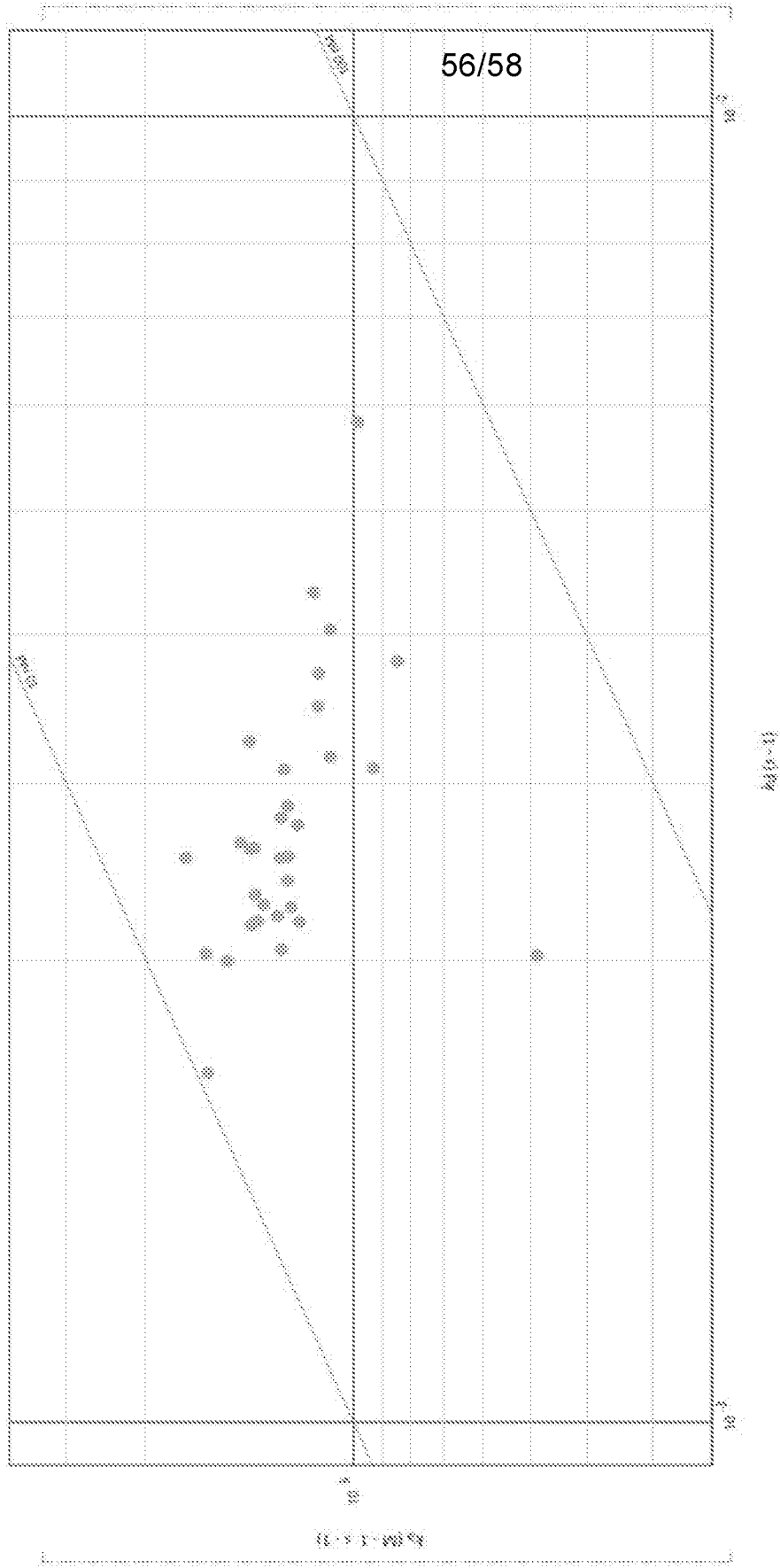


Figure 38C

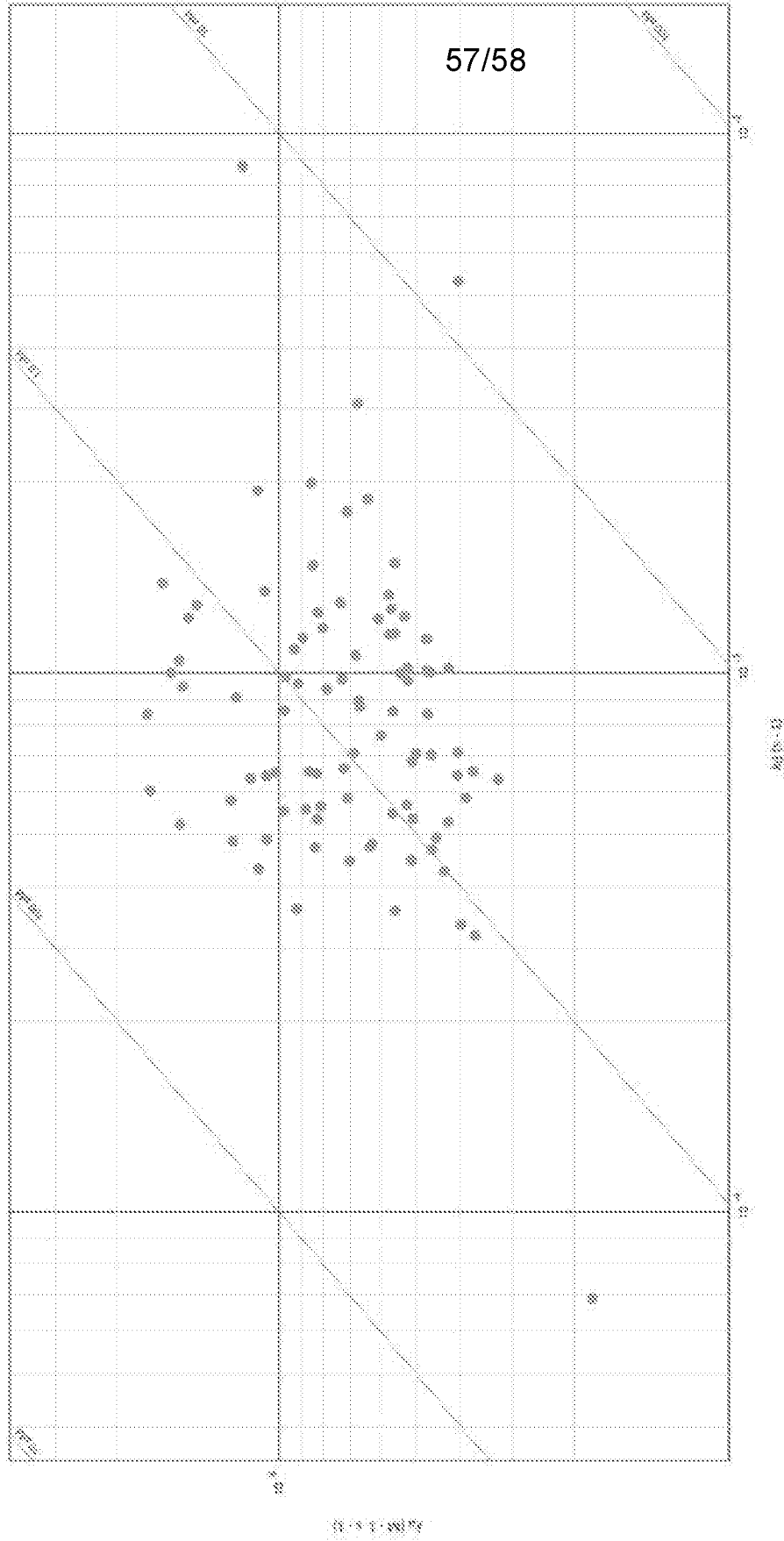


Figure 39

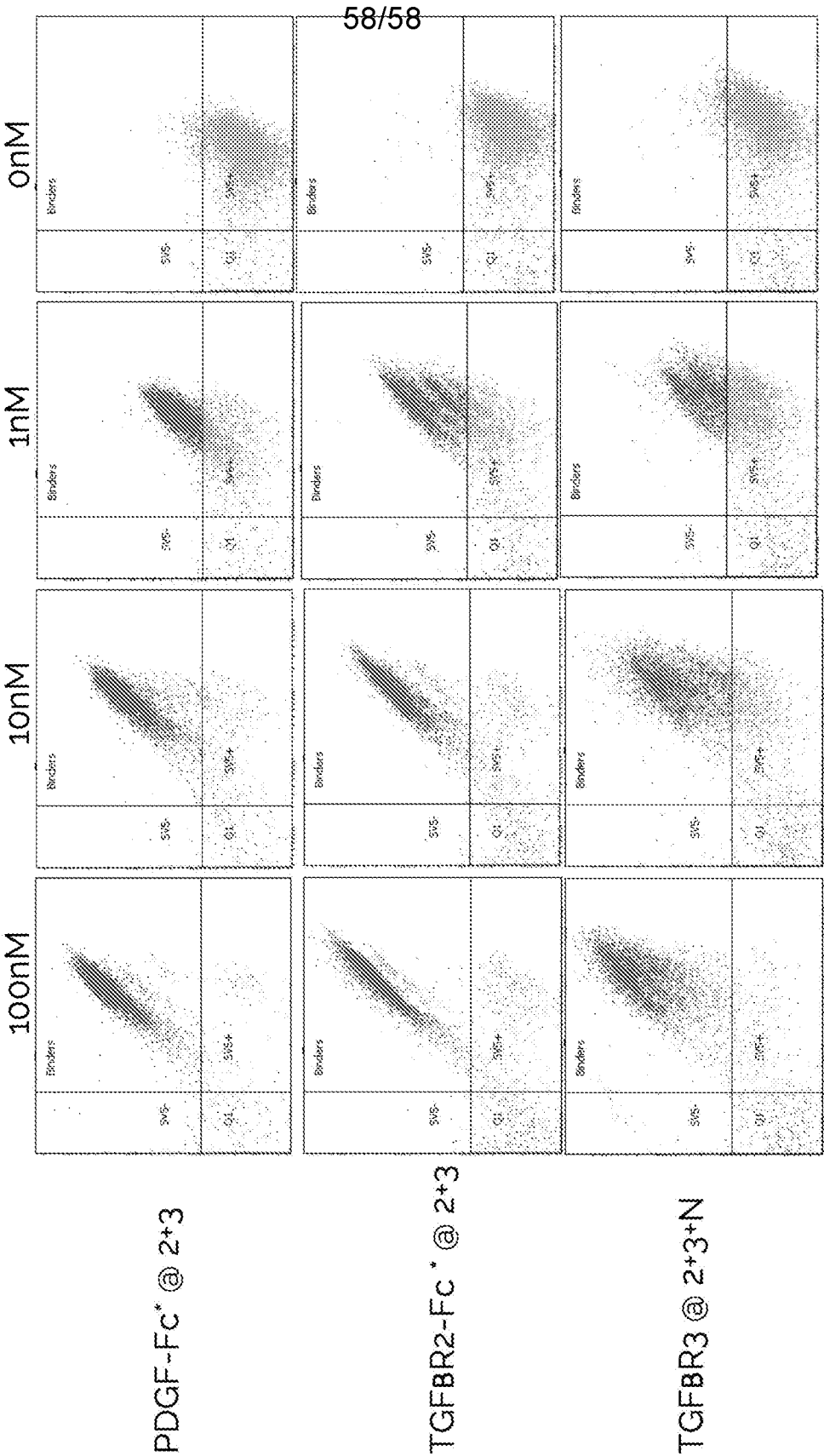


Figure 1

