

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2007 (06.12.2007)

PCT

(10) International Publication Number
WO 2007/137616 A1

(51) International Patent Classification:
C12N 15/10 (2006.01) **C40B 40/02** (2006.01)

(21) International Application Number:
PCT/EP2006/006339

(22) International Filing Date: 30 May 2006 (30.05.2006)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (for all designated States except US): **MILLE-GEN** [FR/FR]; Immeuble BIOSSTEP - Bat A, Rue Pierre et Marie Curie, BP 38183, F-31681 Labège Cedex (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MONDON, Philippe** [FR/FR]; 11, rue de la Bataille, Lieudit Coteau de Fonbazi, F-31450 Donneville (FR). **BOUAYADI, Khalil** [MA/FR]; 15, allée des Glycines, F-31520 Ramonville Saint-Agne (FR). **KHARRAT, Abdelhakim** [FR/FR]; 69 Chemin Al Cers, F-31450 Montgiscard (FR).

(74) Agents: **NARGOLWALLA, Cyra** et al.; Cabinet Plasseraud, 52, rue de la Victoire, F-75440 Paris Cedex 09 (FR).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HIGHLY DIVERSIFIED ANTIBODY LIBRARIES

(57) Abstract: The present invention provides an in vitro method for obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof, comprising the step of performing random mutagenesis of a polynucleotide encoding the variable region of a heavy chain and/or the variable region of a light chain of a light chain, wherein random mutagenesis is performed on a library of polynucleotides comprising a sequence encoding the variable region of a heavy chain and/or the variable region of a light chain; and wherein the random mutagenesis process creates randomly distributed mutations along at least 70% of the sequence encoding the variable region.

WO 2007/137616 A1

Highly diversified antibody libraries.**FIELD OF THE INVENTION**

This invention relates to highly diversified antibody libraries and methods for
5 generating highly diversified antibody libraries.

BACKGROUND OF THE INVENTION

Human antibody fragments can be directly selected from antibody gene repertoires
expressed on the surface of filamentous bacteriophages (Winter et al 1994 Annu Rev
10 Immunol 12:433-455), of yeast cells (Feldaus et al 2003 Nat Biotechnol 21:163-170),
of bacterial cells or of ribosomes (Hanes J. and Pluckthun A. 1997 PNAS 94:4937-
42).

The use of large and diverse human antibodies libraries provide several advantages:
15 Animal immunization can be bypassed. Antibodies may be generated against antigens
which are toxic upon immunization or have a low immunogenicity. Furthermore, in
contrast to murine monoclonal antibodies, the use of human antibodies diminishes the
risk of allergic response.

20 The diversity of the library determines the probability to isolate an antibody with high
affinity for a given antigen.

Antibodies libraries may be obtained from natural sources. Diverse libraries of
immunoglobulin heavy (VH) and light (VL: V kappa and V lambda) chain variable
25 (V) genes were prepared from B-cells of unimmunized donors (Marks et al 1991 J
Mol Biol 222:581-597) or of immunized donors (Barbas et al, 1991 PNAS 88:7978-
7982; Clackson et al 1991 Nature 352:624-628) by polymerase chain reaction (PCR)
amplification. Genes encoding single chain antibody (scFv) were made by randomly
combining heavy and light chain V-genes using PCR, and the combinatorial library
30 was cloned for display on the surface of a phage.

Alternatively libraries may be obtained through the artificial introduction of mutations into the complementarity determining regions (CDR) of the heavy chains or of the light chain domains.

- 5 The CDRH3 loop of the variable heavy genes varies in size and sequence during the rearrangement of the V-D-J segments in the process of forming the unmutated VH repertoire and plays a dominant role in the antibody diversity. CDRH3 is located at the center of the antigen binding site and it is the most variable among the CDRs in natural antibody. Synthetic libraries were constructed by the randomization with
10 degenerate primers of the CDRH3. Several studies showed that medium size libraries (5×10^7 members) with variation in the CRDH3 have provided a successful selection of novel antibody specificities (Barbas et al 1992 PNAS 89:4457-4461; Hoogenboom and Winter 1992 J Mol Biol 227:381-388). Larger libraries $>10^8$ members with CDRH3 sequence lengths of 4-21 residues from 50 VH (Nissim et al 1994 EMBO
15 13:692-698) and 6-15 residues in 49 different VH genes (de Kruif et al 1995 J Mol Biol 248:97-105) allowed the selection of antibody fragments with different specificities. The third CDR strongly contributes to the overall specificity of an antibody. However the drawback of this approach is to shadow the remaining five CDR-loops which also contribute to the specificity and affinity of the antibody.
- 20 Another approach combines all the CDR using a CDR-Implantation Technology (cf. WO0175091). The degree of functional variation is achieved by means of simultaneous and random combination of six biologically derived CDRs (Soderlind et al 2000 Nature18:852-856). The CDRs from a cDNA library prepared from
25 peripheral blood B cell were combined within a selected framework of the DP-47 germline gene (VH3 family) by overlap extension PCR. The genetic diversity produced with this process is different from naturally created in the immune system. This means that this new type of antibody could be potentially immunogen.

Furthermore, the library is based on a single framework and this hinders the ability of the antibodies of the library to bind all types of antigens.

Another approach is disclosed in US6828422 and in Knappik et al 2000 J mol Biol
5 296:57-86. Each of the human VH and VL subfamilies that is frequently used during
an immune response was represented by one consensus framework, resulting in seven
master genes for VH and seven master genes for VL to obtain 49 combinations.
Diversity was created by replacing the VH and VL CDR3 regions of the master genes
by CDR3 library cassettes, generated from mixed trinucleotides and biased towards
10 natural human antibody CDR3 sequences. The sequencing of 257 members of the
unselected libraries indicated that the frequency of correct and thus potentially
functional sequences was 61%. However, the limited variability found in key residues
in the CDR encoded by the few VH and VL used limits its ability to recognize certain
target antigens. Structural incompatibility between these foreign CDRs and the fixed
15 framework may potentially prevent the formation of functional antibody.

Various *in vitro* strategies were used to optimize an antibody selected from a library
screening. These include site specific mutagenesis based on structural information or
combinatorial mutagenesis of CDR. Mutations are usually restricted to the antigen
20 binding surface (CDR). However, although the framework regions of the variable
domains VH and VL are not directly in contact with the antigen, framework residues
can have indirect effects on binding by affecting the CDR conformation. It was
demonstrated that the affinity optimization could be obtained through association of
mutations in the CDR and also in the framework variable fragments (Schier et al
25 1996 J Mol Biol 255:28-43, Boder et al 2000 PNAS 97:10701-10705, Hanes et al
2000 Nature18:1287-1292, Irving et al 2001 J Immunol Methods 248:31-45). In the
humanization process of the antibodies, simple grafting of the CDR sequences often
yields humanized antibodies that bind the antigen much more weakly than the parent
murine antibody. The affinity of an antibody is optimized by replacing key residues

in the framework regions (cf. Baca et al 1997, J Biol Chem 272:10678-10684 and EP1325932). However, optimization of the framework residues is time consuming and tedious.

- 5 There remains a general need in the art for antibody libraries with an increased diversity which will enable the rapid selection and production of antibodies, derivatives thereof or fragments thereof, which bind to a bait molecule with a high affinity.

10 SUMMARY OF THE INVENTION

The present invention provides an *in vitro* method for obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof, comprising the step of performing random mutagenesis of a polynucleotide encoding the variable region of a heavy chain and/or the variable region of a light chain,
15 wherein random mutagenesis is performed on a library of polynucleotides comprising a sequence encoding the variable region of a heavy chain and/or the variable region of a light chain; and wherein the random mutagenesis process creates randomly distributed mutations along at least 70% of the sequence encoding the variable region. By creating randomly distributed mutations along at least 70% of the sequence
20 encoding the variable region, mutations within the CDRs and within the framework are created. By performing this random mutagenesis process on a library of polynucleotides comprising a sequence encoding the variable region, a highly diversified library of polynucleotides with mutations within the CDRs and/or within the framework is generated. This library with increased diversity enables the rapid
25 selection and production of antibodies, derivatives thereof or fragments thereof, which bind to a bait molecule with a high affinity. With this library, the time consuming and tedious optimization of the framework residues is not necessary any more.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides an *in vitro* method for obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof, comprising the step of performing random mutagenesis of a polynucleotide encoding
5 the variable region of a heavy chain and/or the variable region of a light chain, wherein random mutagenesis is performed on a library of polynucleotides comprising a sequence encoding the variable region of a heavy chain and/or the variable region of a light chain; and
wherein the random mutagenesis process creates randomly distributed mutations
10 along at least 70% of the sequence encoding the variable region.

By library of polynucleotides it is meant a large collection (i.e. $> 10^4$ members) of diverse polynucleotides available for screening in order to isolate interesting members.

15 The terms “antibodies, derivatives thereof or fragments thereof” are used in the broadest sense and specifically cover immunoglobulins, such as IgG, IgA, IgM, IgE, IgD, immunoglobulins with polyepitopic specificity, antibody fragments (e.g. the variable region of a heavy chain, the variable region of a light chain, Fab, Fab’,
20 F(ab’)₂, Fv, as well as antibody derivatives (e.g. single chain antibody (scFv)), so long they encompass the variable region of a heavy chain or the variable region of a light chain.

Typically, the random mutagenesis process creates randomly distributed mutations
25 along at least 75%, 80%, 90% or 95% of the sequence encoding the variable region. With this process a library of polynucleotides comprising a sequence with one or more mutations within the CDRs encoding region and/or with one or more mutations within the framework encoding regions is created. In a preferred embodiment, random mutagenesis process creates randomly distributed mutations along the whole
30 sequence encoding the variable region.

Typically, a method according to the invention may comprise the step of performing random mutagenesis of a polynucleotide encoding the variable region of a heavy chain and the step of performing random mutagenesis of a polynucleotide encoding the variable region of a light chain.

5

In a preferred embodiment of the invention, the polynucleotide encoding the variable region is a polynucleotide encoding a single chain antibody.

The single chain antibody technology is well known in the art, this technology has been described in EP0281604, EP0318554 and EP0573551 for example.

10

In a preferred embodiment of the invention, the library of polynucleotides comprising a sequence encoding the variable region of a heavy chain and/or the variable region of a light chain, is obtainable from B-cells isolated from donors selected from the group consisting of healthy donors and donors afflicted with a disease. The B-cells may be obtained from diverse lymphoid sources including peripheral blood lymphocytes (PBLs), bone marrow, spleen or tonsil. Typically the library may be obtained from at least 50 donors. The donors may all be healthy or afflicted with a disease. Alternatively, part of donors may be healthy. Afflicted donors may suffer from bacterial infections, viral infections, autoimmune diseases, endocrine diseases (e.g. diabetes) and/or cancer.

15

It falls within the ability of the skilled person to choose the method for performing random mutagenesis and the suitable working conditions, which will enable the creation of randomly distributed mutations along at least 70% of the sequence encoding the variable region. Typically the skilled person may use error prone PCR. Alternatively, the skilled person may use the method disclosed in WO0238756 wherein random mutagenesis is performed by using one or more mutases selected from the group consisting of DNA polymerases beta, iota, eta and kappa.

20

Typically the sequence encoding the variable region of a heavy chain and/or the variable region of a light chain may be from any origins. Typically they may originate from primates, mice or camels. In a preferred embodiment the sequence encoding the variable region of a heavy chain and/or the variable region of a light chain is from human origin.

Typically the library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof is a library of polynucleotides encoding derivatives or fragments selected from the group consisting of the variable region of a heavy chain, the variable region of a light chain, Fv, a single chain antibody, Fab, Fab' and F(ab')₂. Typically the polynucleotides encoding antibodies, derivatives thereof or fragments thereof are expression vectors.

"Expression vector" refers to polynucleotide sequences containing a desired coding sequence and control sequences in operable linkage, so that hosts transformed with these sequences are capable of producing the encoded proteins.

An embodiment of the present invention provides a library of polynucleotide encoding antibodies, derivatives thereof or fragments thereof obtainable by a method according to the invention.

Various display methods are known in the art, among them, phage display, yeast display, bacterial display and ribosome display are widely used for screening studies (cf. for example Winter et al 1994 Annu Rev Immunol 12:433-455, Feldaus et al 2003 Nat Biotechnol 21:163-170, Hanes J. and Pluckthun A. 1997 PNAS 94:4937-42). Typically, these techniques allow large antibody libraries to be screened against a bait molecule. Alternatively, methods allowing the detection of intracellular interaction have been developed. Within these methods, a bait protein is expressed in the cell. Examples of such methods are the yeast two hybrid method (cf. Fields and Song, 1989 and WO0200729), the cyto Trap method (cf. Broder et al, 1998 and US5776689), the protein fragment complementation (PCA) method (cf. Johnsson and

Varshavsky, 1994 and WO9529195) and the intracellular method described in the application filed by MilleGen with application number EP 06290641.

5 An embodiment of the present invention provides a method for obtaining a library of cells, each comprising a polynucleotide encoding antibodies or fragments thereof, comprising the steps of:

- a) obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof by the method according to the invention; and
- b) transforming cells with the polynucleotides obtained in step a).

10 An embodiment of the present invention provides a library of cells obtainable by this method.

15 Typically, the cells comprising a polynucleotide encoding an antibody, a derivative thereof or a fragment thereof, express on their surface the antibody, the derivative thereof or the fragment thereof. Typically, the cells are bacterial cells or yeast cells.

20 An embodiment of the present invention provides a method for obtaining a library of phages, each comprising a polynucleotide encoding an antibody, a derivative thereof or a fragment thereof and displaying on its surface said antibody, said derivative thereof or said fragment thereof, comprising the step of:

- a) obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof by a method according to the invention; and
- b) generating the phages.

25 An embodiment of the present invention provides a library of phages obtainable by this method.

30 An embodiment of the present invention provides a method for producing an antibody, a derivative thereof or a fragment thereof which binds to a bait molecule comprising the steps of:

- a) obtaining a library of polynucleotide encoding antibodies, derivatives thereof or fragments thereof obtainable by a method according to the invention; and
- b) isolating from the library a polynucleotide which encodes an antibody, a derivative thereof or a fragment thereof which binds to said bait molecule.

5

In the following, the invention will be illustrated by means of the following examples as well as the figures. In all the figures, the CDRs (usually in bold) and the framework regions (FR) were defined according to structural criteria defined by Chothia et al (1989) Nature 342, 877-883.

10

Figure 1: Amino acids sequences of the variable heavy chain gene VH5 (SEQ ID N°1) and VH10 (SEQ ID N°2).

Figure 2: Mutations of the heavy variable genes VH5 and VH10.

Figure 3: Amino acids sequences of three variable lambda genes VL6 (SEQ ID N°8), VL9 (SEQ ID N°9) and VL18 (SEQ ID N°10).

15

Figure 4: Mutations of the lambda variable genes VL6, VL9 and VL18.

Figure 5: Amino acids sequences of two variable kappa chain genes, VK5 (SEQ ID N°12) and VK11 (SEQ ID N°13). The CDRs (in bold) were defined according to structural criteria defined by Chothia *et al* (1989) Nature 342, 877-883.

20

Figure 6: Mutations of the kappa variable genes VK5 and VK11

Figure 7: Representation of the number of mutations along the sequence of single chain antibodies (scFv) of the MB_scFv2e_A_E library (A) and MB_scFv2e_B_N library (B).

25

EXAMPLES

In the following description, all molecular biology experiments are performed according to standard protocol (Sambrook J, Fritsch EF and Maniatis T (eds) Molecular cloning, A laboratory Manual 2nd Ed, Cold Spring Harbor Laboratory Press).

30

Example 1: Random mutagenesis of isolated variable domains of human antibodies**A. Amplification and cloning of the human antibody repertoires of variable heavy and light chain domains.**

Total peripheral blood monocyte (PBMC) RNA was isolated from 100 donors. The
 5 donors were healthy or were afflicted with diverse diseases: Bacterial or viral
 infections, autoimmune diseases, endocrine diseases (e.g. diabetes) and cancer.
 PBMC mRNA was isolated and antibody heavy and light chains (Vlambda and
 Vkappa) cDNA produced from the reverse transcription were PCR amplified with
 several mixtures of primer pair characteristics of the N-terminal and C-terminal
 10 extremities of all different families of the variable genes.

The VH were PCR amplified using 14 different VH forward primers and a mixture of
 4 JH reverse primers. A second PCR with tagged primers were used to introduce
 restriction sites SalI in 5' and EcoRI in 3' of the VH fragments. The VH PCR
 15 products were cloned into SalI and EcoRI restriction sites of pUC18 vector.

The Vlambda cDNA were PCR amplified with 15 different Vlambda forward primers
 and 3 JL reverse primers and the Vkappa cDNA were PCR amplified with a mix of
 13 VK forward primers and 5 JK reverse primers. The Vlambda and Vkappa
 20 fragments were PCR amplified with tagged primers to introduce XbaI in 5' and SalI
 in 3' and cloned into SalI and XbaI restriction site of pUC18.

Each library HS_VH, HS_Vlambda and HS_Vkappa contained $1-3 \times 10^6$ members:

Table 1: Human antibody heavy and light chain libraries cloned

25

Library name	HS_VH	HS_Vlambda	HS_Vkappa
Vector	pUC18	pUC18	pUC18
Size	1×10^6	3×10^6	1.25×10^6

B. Mutagenesis of two variable heavy chain domains from the human antibody repertoire

The DNA encoding the variable heavy domains of two randomly picked clones, VH5 (SEQ ID N°1) and VH10 (SEQ ID N°2), from the human library repertoire HS_VH were mutated with the MutaGenTM process. The MutaGenTM is described in WO0238756. The DNA coding for the VH5 and VH10 clones corresponded to the germline gene IGHV1-3 and IGHV1-24, respectively. For this determination the blast program was used on the IMGT web site: <http://imgt.cines.fr/>. The sequences are described in figure 1. The mutagenesis process was designed to modify the complete variable gene without the JH region. This region was not mutated in prevision of possible future use as template in an overlap PCR to associate the VH and the VL libraries. Two randomly mutated libraries were constructed MB-VH5 and MB-VH10.

a. Mutagenesis of the VH domains

The VH genes were double replicated with human polymerase beta using the 5' primer Mut-PD-S1 5'-CGAGCGTCTACTAGCGCATGCCTGCAGGTCGAC-3' (SEQ ID N°3), the 3' primer as an equimolar mixture of 4 primers [JH-1/2-for-mut: 5'-ATGCGTGAATTCTGAGGAGACGGTGACCAGGGTGCC (SEQ ID N°4), JH-3-for-mut: 5'-ATGCGTGAATTCTGAAGAGACGGTGACCATTGTCCC-3' (SEQ ID N°5), JH-4/5-for-mut: 5'-ATGCGTGAATTCTGAGGAGACGGTGACCAGGGTTCC-3' (SEQ ID N°6), JH-6-for-mut: 5'-ATGCGTGAATTCTGAGGAGACGGTGACCGTGGTCCC-3' (SEQ ID N°7)] and 1 µg of plasmid pUC18-VH (pUC18-VH5 or pUC18-VH10) as template in the replication buffer A, B or E (cf. table 2).

Table 2: Conditions of replication for the random mutagenesis

	dATP	dCTP	dTTP	dGTP	Mn ²⁺	DMSO
Condition A	50 μ M	50 μ M	100 μ M	100 μ M		
Condition B	20 μ M	100 μ M	100 μ M	100 μ M	0.5 mM	
Condition C	20 μ M	100 μ M	100 μ M	100 μ M	0.5 mM	10%
Condition E	20 μ M	100 μ M	100 μ M	100 μ M	0.25mM	
Condition N	100 μ M	100 μ M	100 μ M	100 μ M		

The mixture was denatured for 5 min at 95°C and cooled down to 4°C. An equal
 5 volume of a solution containing 4 units of human polymerase beta in replication
 buffer A, B or E was added to each replication reaction A, B and E. The reactions of
 replication were carried out at 37°C for one hour. The replication products were then
 purified through phenol-chloroform and ethanol precipitation.

10 **b. Selective amplification and cloning of mutated fragments**

The replication products obtained in replication conditions A, B and E were
 selectively amplified through a selective PCR amplification with tail primers. The
 primers were designed with a tail that is non-specific to the template and allowed to
 specifically amplify the DNA fragments synthesized by the mutases.

15 A fraction of each replication product obtained in the replication conditions A, B and
 E was added to a mixture containing the PCR buffer (20 mM Tris HCl pH 8.4, 50
 mM KCl) (Life Technologies), 1.5 mM MgCl₂, 10 pmol of the 5' and 3' primers, 200
 microM of the 4 dNTPs and 1.25 U Platinum Taq DNA polymerase (Life
 Technologies). This mixture was incubated 5 min at 95°C, 5 sec at 55 C, 30 sec at
 20 72°C following by 30 selective cycles of 20 sec at 94 °C and 30 sec at 72°C.

The amplified replication products were cloned into SalI and EcoRI restriction sites
 of pUC18 to obtain MB-VH5 and MB-VH10 libraries.

c. Analyze of randomly mutated libraries of the selected VH domains.

A repertoire of randomly mutated VH domains was obtained. Test screening of individual randomly picked library clones by DNA sequencing revealed an intact insert that contained randomly mutated VH5 or VH10 gene (cf. figure 2). The mutations are distributed randomly along the entire VH gene (i.e. the framework regions and the CDR loops).

The modifications of the mutated sequences were analyzed with the software Mutanalyse 2.5 (MilleGen). The frequency of mutations (cf. table 3) of the MB-VH5 and MB-VH10 libraries were 5.55 and 9.64 mutation per kilo bases, respectively. The average frequency of mutations of 7.58 per kilobases of the MB-VH5 and MB-VH10 libraries means 2.41 base mutations per gene.

By active part of the library, it is meant the quantity of clones with open reading frames and thus potentially coding for functional variable antibody domains.

Table 3: Mutation frequencies of the heavy and light variable genes of the isolated genes

	MB-VH5	MB-VH10	MB-VL6	MB-VL9	MB-VL18	MB-VK5	MB-VK11
Number of replicated bases sequenced per gene	319	318	333	324	324	339	318
Number of sequences studied	28	32	63	61	67	63	64
Active part of the library (%)	78.5	46.8	50.7	41	49.2	36.5	56.2
Base Mutation frequency (on the active part of the library /kb)	5.55	9.64	8.72	8.14	8.89	8.97	5.15
Number of amino acid residues (a.a) replicated per gene	106	106	111	108	108	113	106
Number of sequences with a stop	2	1	2	1	1	0	0
Number of sequences with silence mutations	4	4	9	3	4	6	7
Number of sequences studied	16	10	21	21	28	17	29
Amino acids Mutation frequency (%)	1.27	1.97	1.92	1.68	2.10	2.26	1.28

All different kinds of base mutations are represented (cf. table 4). The different transitions represented 82 % (MB-VH5) and 52 % (MB-VH10) and the transversions 18% and 48% (cf. table 4). The majority of the clones have 1 base mutation, however, some clones could have up to 11 base mutations (cf. table 5). Because of the silence mutation, this high level of base mutations does not result in a high level of amino acid change. The base mutations could also induce a silence mutation or a stop codon. The amino-acid mutation frequencies of the active part of the library were 1.27% and 1.97% for MB-VH5 and MB-VH10, respectively (cf. table 3). Table 6 shows the frequency of amino acid mutations according to the position in the frameworks and the CDR.

Table 4: Mutation diversity

	MB-VH5	MB-VH10	MB-VL6	MB-VL9	MB-VL18	MB-VK5	MB-VK11
% A=>G or T=>C	43.58	28.26	52.22	36.36	41.86	45.71	54.23
% G=>A or C=>T	38.46	23.91	23.33	30.3	24.41	34.28	32.2
% A=>T or T=>A	2.56	0	3.33	7.57	3.48	2.85	1.69
% A=>C or T=>G	7.69	28.26	13.33	13.63	18.6	11.42	5.08
% G=>C or C=>G	2.56	8.69	1.11	7.57	5.81	1.42	3.38
% G=>T or C=>A	5.12	10.86	6.66	4.54	5.81	4.28	3.38

Table 5: Base mutation distribution per sequence

	MB-VH5	MB-VH10	MB-VL6	MB-VL9	MB-VL18	MB-VK5	MB-VK11
1 mutation / seq	11	10	15	13	17	9	23
2 mutations / seq	7	0	6	5	5	5	6
3 mutations / seq	2	1	2	0	3	2	5
4 mutations / seq	2	0	3	1	2	1	1
5 mutations / seq	0	0	0	0	0	2	1
6 mutations / seq	0	1	1	3	3	1	0
7 mutations / seq	0	0	1	2	0	1	0
8 mutations / seq	0	2	1	0	0	0	0
9 mutations / seq	0	0	2	0	0	2	0

10 mutations / seq	0	0	1	1	1	0	0
11 mutations / seq	0	1	0	0	1	0	0
12 mutations / seq	0	0	0	0	1	0	0

Table 6: Frequency of amino acid mutations in the different frameworks (FR) and CDR of the MB-VH5 and MB-VH10

	FR1	CDR1	FR2	CDR2	FR3	CDR3
Number of Mutation	12	4	7	5	11	4
Number of residues/genes	25	7	19	6	41	6-8
Total Number of residues*	650	182	494	156	1066	182
Frequency of aa mutation (%)	1.84	2.19	1.4	3.2	1.03	2.19

* number of residues x number of mutated genes studied.

FR: framework of the variable domain, CDR: Complementary determining region of the variable domain

C. Mutagenesis of three variable lambda (VL) domains from the human antibody repertoire

The DNA encoding the variable lambda domains of three clones, VL6 (SEQ ID N°8), VL9 (SEQ ID N°9) and VL18 (SEQ ID N°10), picked randomly from the human library repertoire HS_Vlambda were randomly mutated with the MutaGenTM process. The sequences of the clones are described in figure 3. The VL6, VL9 and VL18 clones corresponded to the germline gene IGLV2-8, IGLV1-51 and IGLV3-21, respectively. Three highly diversified libraries were constructed MB-VL6, MB-VL9 and MB-VL18.

a. Mutagenesis of the selected Vlambda domains

The VL genes were double replicated with human polymerase beta using the 5' primer Mut-PD-S1 5'-CGAGCGTCTACTAGCGCATGCCTGCAGGTCGAC-3'

(SEQ ID N°3), the 3' primer VL-K1-R 5'-TCAGTCTATCGTCACGTCAACGGCGGCGGATCTTCTAGA-3' (SEQ ID N°11) and 1 µg of plasmid pUC18-VL (pUC18-VL6, pUC18-VL9 or pUC18-VL18) as template in the replication buffer A, B or E (cf. table 2). Each mixture was denatured for 5 min at 95°C and cool down to 4°C. An equal volume of a solution containing 4 units of polymerase beta in replication buffer A, B or E was added. The reactions of replication were carried out at 37°C for one hour. The replication products were then purified.

b. Selective amplification and cloning of mutated fragments

The replication products were selectively amplified through a selective PCR amplification with tail primers. The primers were designed with a tail that is not specific to the template and allow specific amplification of DNA fragments synthesized by the mutases.

A fraction of each replication product obtained in the replication conditions A, B and E was added to a mixture containing the PCR buffer (20 mM Tris HCl pH 8.4, 50 mM KCl) (Life Technologies), 1.5 mM MgCl₂, 10 pmol of the 5' and 3' primers, 200 microM of the 4 dNTPs and 1.25 U PlatinumTaq DNA polymerase (Life Technologies). This mixture was incubated 5 min at 95°C, 5 sec at 55°C, 30 sec at 72°C following by 30 selective cycles of 20 sec at 94°C and 30 sec at 72°C.

The amplified replication products were cloned into XbaI and SalI restriction sites of pUC18 to obtain: MB-VL6, MB-VL9 and MB-VL18 libraries.

c. Analysis of randomly mutated libraries of the selected Vlambda domains

A repertoire of randomly mutated VL domains was obtained: MB-VL6, MB-VL9 and MB-VL18. Test screening of individual randomly picked library clones by DNA sequencing revealed an intact insert that contained randomly mutated VL6, VL9 or VL18 gene (cf. figure 4). The mutations are distributed randomly along the entire VL

gene (i.e. the framework regions and the CDR loops). The modifications of the mutated sequences were analyzed with the software Mutanalyse 2.5 (MilleGen).

The frequencies of mutations (cf. table 3) of the MB-VL6, MB-VL9 and MB-VL18 libraries were 8.72, 8.14 and 8.89 mutation per kilo bases, respectively. The average
5 frequency of mutations of 8.58 per kilobases of the three MB-VL libraries means 2.80 base mutations per gene.

All different kinds of base mutations are represented (cf. table 4). The average of the different transitions represented 69 % and the transversions 31% (cf. table 4). The
10 majority of the clones have 1 base mutation, however, some clones could have up to 12 base mutations (table 5). Because of the silence mutation, this high level of base mutations does not result in a high level of amino acid change. The base mutations could also induce a silence mutation or a stop codon. The amino-acid mutation frequencies of the active part of the library were 1.92 % 1.68 % and 2.10 % for MB-
15 VL6, MB-VL9 and MB-VL18, respectively (cf. table 3). Table 7 shows a homogenous distribution of the amino acid mutations according to the position in the frameworks and the CDR.

Table 7: Frequency of amino acid mutations in the different frameworks (FR) and
20 CDR of the MB-VL6, MB-VL9 and MB-VL18

	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4
Number of Mutations	14	15	12	5	48	15	14
Number of residues/genes	22	11-14	15	7	32	11	8-10
Total number of residues*	1386	795	945	441	2016	693	588
Frequency of amino acid mutation (%)	1.01	1.89	1.26	1.13	2.38	2.16	2.38

* number of residues x number of mutated genes studied.

FR: framework of the variable domain, CDR: Complementary determining region of the variable domain

5 **D. Mutagenesis of the variable kappa (VK) domains from the human antibody repertoire**

The DNA coding the variable kappa domains of two clones, VK5 (SEQ ID N°12) and VK11 (SEQ ID N°13), picked randomly from the human library repertoire HS-Vkappa were randomly mutated with the MutaGenTM process. The sequences of the
10 clones are described in figure 5. The VK5 and VK11 clones corresponded with the germline gene IGKV4-1 and IGKV6-57, respectively. Two randomly mutated libraries were constructed MB-VK5 and MB-VK11.

a. Mutagenesis of the selected VK domains

15 The VK genes were double replicated with human polymerase beta using the 5' primer Mut-PD-S1 (SEQ ID N°3), the 3' primer VL-K1-R (SEQ ID N°11) and 1 µg of plasmid pUC18-VK (pUC18-VK5 or pUC18-VK11) as template in three different replication buffers A, B or E (cf. table 2). Each mixture was denatured for 5 min. at 95°C and cooled down to 4°C. An equal volume of a solution containing 4 units of
20 polymerase beta in replication buffer A, B or E was added. The reactions of replication were carried out at 37°C for one hour. The replication products were then purified.

b. Selective amplification and cloning of mutated fragments

25 The replication products were selectively amplified through a selective PCR amplification with tail primers. The primers were designed with a tail that is not specific to the template and allow specific amplification of DNA fragments synthesized by the mutases.

30 A fraction of each replication product obtained in conditions A, B and E was added to a mixture containing the PCR buffer (20 mM Tris HCl pH 8.4, 50 mM KCl) (Life

Technologies), 1.5 mM MgCl₂, 10 pmol of the 5' and 3' primers, 200 μM of the 4 dNTPs and 1.25 U Platinum Taq DNA polymerase (Life Technologies). This mixture was incubated 5 min at 95°C, 5 sec at 55°C, 30 sec at 72°C followed by 30 selective cycles of 20 sec at 94°C and 30 sec at 72°C.

5

The amplified replication products were cloned into XbaI and SalI restriction sites of pUC18 to obtain MB-VK5 and MB-VK11 libraries.

c. Analysis of randomly mutated libraries of the selected VK domains

10 A repertoire of randomly mutated VK domains was obtained: MB-VK5 and MB-VK11. Test screening of individual randomly picked library clones by DNA sequencing revealed an intact insert that contained randomly mutated VK5 or VK11 gene (cf. figure 6). The mutations are distributed randomly along the entire VK gene, in the frameworks and/or in the CDR loops. The modifications of the mutated
15 sequences were analyzed with the software Mutanalyse 2.5 (MilleGen).

The frequencies of mutations (cf. table 3) of the MB-VK5 and MB-VK11 libraries were 8.97 and 5.15 mutations per kilo bases, respectively. The average frequency of mutations of 7.06 per kilobases of the MB-VK libraries means 2.31 base mutations per gene.

20

All different kinds of base mutations are represented (cf. table 4). The average of the different transitions represented 83% and the transversions 27% (cf. table 4). The majority of the clones have 1 base mutation, however, some clones could have up to 9 base mutations (cf. table 5). Because of the silence mutation, this high level of base
25 mutations did not result in a high level of amino acid change. The base mutations could also induce a silence mutation or a stop codon. The amino-acid mutation frequencies of the active part of the library were 2.26 % and 1.28 % for MB-VK5 and MB-VK11, respectively (cf. table 1). Table 8 shows an homogenous distribution of the amino acid mutations according to the position in the frameworks and the CDR.

Table 8: Frequency of amino acid mutations in the different frameworks (FR) and CDR of the MB-VK5 and MB-VK11

	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4
Nber of Mutations	20	16	16	5	22	3	5
Nber of residues/genes	23	12-17	15	7	32	9	10
Nber of residues*	1058	637	690	322	1472	414	460
Frequency of aa mutation %	1.89	2.51	2.31	1.55	1.49	0.72	1.09

5

* number of residues x number of mutated genes studied.

FR: framework of the variable domain, CDR: Complementary determining region of the variable domain

10 **Example 2: Random mutagenesis applied to variable domain libraries.**

A. Random mutagenesis of the HS-VH library

The HS-VH library (c.f. table 1) obtained as described before was double replicated with human polymerase beta using the 5' primer Mut-PD-S1 (SEQ ID N°3), the 3' primer as an equimolar mixture of 4 primers [JH-1/2-for-mut (SEQ ID N°4), JH-3-for-mut (SEQ ID N°5), JH-4/5-for-mut (SEQ ID N°6), JH-6-for-mut (SEQ ID N°7)]
 15 and 1 µg of plasmid pUC18-VH (HS-VH DNA library) as template in the replication buffer A, B or E (cf. table 2). Each mixture was denatured for 5 min at 95°C and cooled down to 4°C. An equal volume of a solution containing 4 units of polymerase beta in replication buffer A, B or E was added. The reaction of replication was carried out at 37°C for one hour.

The replication products obtained in condition A, B and E were selectively amplified through a selective PCR amplification with tail primers. The primers were designed with a tail that is not specific to the template and allow specific amplification of DNA
 25 fragments synthesized by the polymerase beta. The replication products selectively PCR amplified were cloned into SalI and EcoRI restriction sites of pUC18 to obtain the library MB_HS_VH_AEB01 (cf. table 9).

B. Random mutagenesis of the HS-Vlambda library.

The HS-VL library (cf. table 1) was double replicated with polymerase beta using the 5' primer Mut-PD-S1 (SEQ ID N°3), the 3' primer VL-K1-R (SEQ ID N°11) and 1 µg of plasmid pUC18-VL (MB-VL DNA library) as template in the replication buffer A, B or E (cf. table 2). Each mixture was denatured for 5 min at 95°C and cooled down to 4°C. An equal volume of a solution containing 4 units of polymerase beta in replication buffer A, B or E was added. The reaction of replication was carried out at 37°C for one hour. The replication products obtained in condition A, B and E were selectively amplified through a selective PCR amplification with tail primers. The primers were designed with a tail that is not specific to the template and allow specific amplification of DNA fragments synthesized by the polymerase beta. The replication products selectively PCR amplified were cloned into SalI and EcoRI restriction site of pUC18 to obtain the library MB_HS_VL_AEB01 (cf. table 9).

C. Random mutagenesis of the HS-Vkappa library.

The HS-Vkappa library (cf. table 1) was double replicated with polymerase beta using the 5' primer Mut-PD-S1 (SEQ ID N°3), the 3' primer VL-K1-R (SEQ ID N°11), 1 µg of plasmid pUC18-VK (HS_VK DNA library) as template in the replication buffer A, B or E (cf. table 2). The mixture was denatured for 5 min at 95°C and cooled down to 4°C. An equal volume of a solution containing 4 units of polymerase beta in replication buffer A, B or E was added. Each reaction of replication was carried out at 37°C for one hour. The replication products were selectively amplified through a selective PCR amplification with tail primers. The primers were designed with a tail that is non-specific to the template and allowed to specifically amplify the DNA fragments synthesized by the polymerase beta. The replication products selectively PCR amplify were cloned into SalI and EcoRI restriction site of pUC18 to obtain the library MB_HS_VK_AEB01 (cf. table 9).

D. Analysis of randomly mutated libraries

The random mutagenesis of the human repertoires of HS_VH, HS_VL and HS_VK domains has provided three libraries: MB_HS_VH_AEB01, MB_HS_VK_AEB01 and MB_HS_VL_AEB01, with library sizes of 1.17×10^7 , 9.4×10^6 and 4.8×10^6 clones, respectively (cf. table 9).

A sample of 400 randomly picked clones from the different libraries were DNA sequenced and analyzed. A representative sample of reference sequences of each subgroup VH, Vlambda and Vkappa from <http://imgt.cines.fr/textes/vquest/refseqhb.html#VQUEST> were locally downloaded. The sequences from the libraries were compared to the reference sequences with the bioinformatic tools developed by MilleGen.

A sample of 199 clones of the heavy chain library MB_HS_VH_AEB01 showed a distribution of the subtype different to that was reported in the natural representation (cf. Table 10). The VH1 and VH3 are the most represented as in the natural representation of the subtype, however, VH1 is largely more represented with 56.78%.

The 201 sequenced clones of the light chain libraries (MB_HS_VL_AEB01 and MB_HS_VK_AEB01) showed a high representation of the VL1 (66%) and VK1 (70.36%) subtypes.

The difference of the subtype distribution in the MutalBank libraries compared to the reported frequencies showed the originality of the constructed libraries. This difference is probably due to the origins of donors which are donors afflicted with diverse diseases.

Table 9: Size of the different libraries

Library name	MB_HS_VH_AEB01	MB_HS_VL_AEB01	MB_HS_VK_AEB01
Vector	pUC18	pUC18	pUC18
Size	1.17x10 ⁷	9.4x10 ⁶	4.8x10 ⁶

5 **Table 10:** Representation of the germline families

<i>Name of the library</i>	<i>Germline family</i>	<i>Nber of clone sequenced</i>	<i>Representation</i>	<i>Natural representation⁽¹⁾</i>
MB_HS_VH_AEB01		199		
	IGHV1	113	56.78%	19.57%
	IGHV2	13	6.53%	6.52%
	IGHV3	37	18.59%	45.65%
	IGHV4	5	2.51%	21.74%
	IGHV5	26	13.06%	2.17%
	IGHV6	4	2.01%	2.17%
	IGHV7	1	0.50%	2.17%
MB_HS_VL_AEB01		50		
	IGLV1	33	66,0%	17.24%
	IGLV2	3	6,0%	17.24%
	IGLV3	2	4,0%	27.59%
	IGLV4	3	6,0%	10.34%
	IGLV5	4	8,0%	6.90%
	IGLV6	0	0,0%	3.45%
	IGLV7	3	6,0%	6.90%
	IGLV8	0	0,0%	3.45%
	IGLV9	1	2,0%	3.45%
	IGLV10	1	2,0%	3.45%
	IGLV11	0	0,0%	0.00%
MB_HS_VK_AEB01		151		
	IGKV1	107	70.86%	50.00%
	IGKV2	6	3.97%	26.47%
	IGKV3	13	8.60%	17.65%
	IGKV4	19	12.58%	2.94%
	IGKV5	4	2.64%	2.94%
	IGKV6	2	1.32%	0.00%
	IGKV7	0	0%	0.00%

⁽¹⁾ IMGT <http://imgt.cnusc.fr:8104>

Example 3: Random mutagenesis applied to scFv libraries.

The retrotranscript of the variable heavy and light genes of the mouse hybridoma VEBA76.50 were PCR amplified and cloned. The DNA coding for the VH and VL domains corresponded with the germline gene IGHV5S14 and IGKV1-117 (IMGT web site: <http://imgt.cines.fr/>). A single chain antibody Fv fragment (scFv) having the structure VH-VK was cloned into the vector pCR4-topoTA to obtain the clone MG_scFv2e.

The MG_scFv2e gene was double replicated with human polymerases beta or eta using the 5' primer VH-scFv2e-S3 5'-TGACGAGTACTAGCTGCTACACCAGGATCCGAAGTGAAGTT-3' (SEQ ID N°14) and the 3' primer VK-scFv2e-R3 5' -ACAGCTACGTGATACGACTCACAGAATTCCCGTTTGATTTC-3' (SEQ ID N°15) and 1 µg MG_scFv2e plasmid DNA as template in the replication buffer E or N (cf. table 2). The mixture was denatured for 5 min at 95°C and cool down to 4°C. An equal volume of a solution containing 4 units of human polymerase beta (mutase A) or human polymerase eta (mutase B) was added in replication buffer E or N. The reaction of replication was carried out at 37°C for two hours. The replication products selectively PCR amplified were cloned into BamHI and EcoRI restriction sites of the vector p03.

Two libraries were constructed: MB_scFv2e_A_E obtained from the replication with mutase A in the condition E and MB_scFv2e_B_N from the replication with mutase B in the condition N. The library sizes were 6×10^5 and 5×10^5 clones, respectively. The alignment of 31 and 30 sequences obtained from different clones of MB_scFv2e_A_E and MB_scFv2e_B_N libraries shows an homogenous distribution of the base mutations along the DNA sequences (cf. figure 7). Table 11 shows the modifications of the mutated sequences analyzed with Mutanalyse 2.5 (MilleGen). According to the mutation conditions used, the libraries obtained have a different

- mutation profile. The frequency of base mutations is double in the MB_scFv2e_B_N library (7.51 per kb) compared to MB_scFv2e_A_E (3.14 per kb). Furthermore, the mutase B in condition N seems to induce a higher modification of the amino acids. The amino acid mutation frequency was 4 times higher in the MB_scFv2e_B_N library than in MB_scFv2e_A_E library (cf. table 11). The two different mutagenesis conditions lead to a library with a low mutagenesis frequency (0.57 mutation per 100 amino acids) and to library with a high mutagenesis frequency (1.52 mutation per 100 amino acids).
- 10 The different transitions represented 80% and the transversions 20% with the two libraries (cf. table 12).
- The number of mutations per sequence of the MB_scFv2e_B_N library is widely represented with 59% of the sequences which have between 6 and 12 base modifications (cf. table 13). In contrast, the MB_scFv2e_A_E library has 74 % of the sequences with 1-3 base mutations.
- 15

Table 11: Mutation frequencies of MB_scFv2e_A_E and MB_scFv2e_B_N libraries

	MB_scFv2e_A_E	MB_scFv2e_B_N
Nber of replicated bases sequenced per gene	786	759
Nber of sequence studied	31	30
Active part of the library (%)	48	47
Base Mutation frequency of the active part of the library (/kb)	3.14	7.51
Nber of aa replicated per gene	262	253
Nber of sequence with a stop	0	1
Nber of sequence with silence mutations	4	2
Nber of mutated sequence with a total open reading frame	15	14
Amino acids Mutation frequency (%)	0.57	1.52
Library size (total number of clones)	6x10 ⁵	5x10 ⁵

Table 12: Mutation diversity of MB_scFv2e_A_E and MB_scFv2e_B_N libraries

	MB_scFv2e_A_E	MB_scFv2e_B_N
% de A=>G or T=>C	57.89	62.88
% de G=>A or C=>T	23.68	17.52
% de A=>T or T=>A	0	11.34
% de A=>C or T=>G	10.52	6.18
% de G=>C or C=>G	5.26	2.06
% de G=>T or C=>A	2.63	0

5

Table 13: Base mutation distribution per sequence of MB_scFv2e_A_E and MB_scFv2e_B_N libraries

	MB_scFv2e_A_E	MB_scFv2e_B_N
1 mutation / seq	11	5
2 mutations / seq	1	0
3 mutations / seq	2	1
4 mutations / seq	2	0
5 mutations / seq	1	1
6 mutations / seq	0	3
7 mutations / seq	1	1
8 mutations / seq	1	2
9 mutations / seq	0	1
10 mutations / seq	0	0
11 mutations / seq	0	2
12 mutations / seq	0	1

10 **References:**

Throughout this application, various references describe the state of the art to which this invention pertains. The disclosures of these references are hereby incorporated by reference into the present disclosure.

CLAIMS

1. An *in vitro* method for obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof, comprising the step of performing random
5 mutagenesis of a polynucleotide encoding the variable region of a heavy chain and/or the variable region of a light chain, wherein random mutagenesis is performed on a library of polynucleotides comprising a sequence encoding the variable region of a heavy chain and/or the variable region of a light chain; and wherein the random mutagenesis process creates randomly distributed mutations
10 along at least 70% of the sequence encoding the variable region.
2. The method according to claim 1, wherein the random mutagenesis process creates randomly distributed mutations along at least 90% of the sequence encoding the variable region.
- 15 3. The method according to claim 1, wherein the random mutagenesis process creates randomly distributed mutations along the whole sequence encoding the variable region.
- 20 4. The method according to any one of claims 1 to 3, comprising the step of performing random mutagenesis of a polynucleotide encoding the variable region of a heavy chain and the step of performing random mutagenesis of a polynucleotide encoding the variable region of a light chain.
- 25 5. The method according to any one of claims 1 to 4, wherein the polynucleotide encoding the variable region is a polynucleotide encoding a single chain antibody.
6. The method according to any one of claims 1 to 5, wherein the library of polynucleotides comprising a sequence encoding the variable region of a heavy chain
30 and/or the variable region of a light chain, is obtainable from B cells isolated from

donors selected from the group consisting of healthy donors and donors afflicted with a disease.

7. The method according to any one of claims 1 to 6, wherein random mutagenesis is performed by using one or more mutases selected from the group consisting of DNA polymerases beta, iota, eta and kappa.

8. The method according to any one of claims 1 to 7, wherein the sequence encoding the variable region of a heavy chain and/or the variable region of a light chain is from human origin.

9. The method according to any one of claims 1 to 8, wherein the library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof is a library of polynucleotides encoding derivatives or fragments selected from the group consisting of the variable region of a heavy chain, the variable region of a light chain, Fv, a single chain antibody, Fab, Fab' and F(ab')₂.

10. The method according to any one of claims 1 to 9, wherein the polynucleotides encoding antibodies, derivatives thereof or fragments thereof are expression vectors.

11. A method for obtaining a library of cells, each comprising a polynucleotide encoding antibodies or fragments thereof, comprising the steps of:

- a) obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof by the method according to any one of claims 1 to 10; and
- b) transforming cells with the polynucleotides obtained in step a).

12. The method according to claim 11 wherein the cells comprising a polynucleotide encoding an antibody, a derivative thereof or a fragment thereof, express on their surface the antibody, the derivative thereof or the fragment thereof.

13. The method according to claim 11 or 12, wherein the cells are bacterial cells or yeast cells.

14. A method for obtaining a library of phages, each comprising a polynucleotide encoding an antibody, a derivative thereof or a fragment thereof and displaying on its surface said antibody, said derivative thereof or said fragment thereof, comprising the step of:

a) obtaining a library of polynucleotides encoding antibodies, derivatives thereof or fragments thereof by the method according to any one of claims 1 to 10; and

b) generating the phages.

15. A library of polynucleotide encoding antibodies, derivatives thereof or fragments thereof obtainable by the method according to any one of claims 1 to 10.

16. A library of cells obtainable by the method according to any one of claims 11 to 13.

17. A library of phages obtainable by the method according to claim 14.

18. A method for producing an antibody, a derivative thereof or a fragment thereof which binds to a bait molecule comprising the steps of:

a) obtaining a library of polynucleotide encoding antibodies, derivatives thereof or fragments thereof obtainable by the method according to any one of claims 1 to 10; and

b) isolating from the library a polynucleotide which encodes an antibody, a derivative thereof or a fragment thereof which binds to said bait molecule.

VH5

QVQLVQSGTEVKKPGASVKVSCKVSGY~~TFTT~~SAMHWVRQAP~~EQGPEW~~MGWINAGN~~GDTKYSQK~~
FQGRVTITRDTSASTAYMELSSLGSEDTAVYFCARGWNY~~DDW~~GQGLVTVS

VH10

EVQLVQSGAEVKKPGASVKVSCKVSGY~~TLTE~~LSMHWVRQAPGKGLEWMGGF~~DPED~~GETIYAQK
FQGRVTMTEDTSTD~~TAYMELSSL~~RSEDTAVYYCATVPWMYY~~FDY~~WEPWSPSPQ

Figure 1

	FR1	CDR1	FR2	CDR2	FR3	CDR3 FR4
VH5.TX1_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RAVH5a02_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RAVH5a04_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RBVH5a07_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RBVH5a05_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RBVH5a07_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RBVH5a07_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RBVH5a05_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RBVH5a07_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
RBVH5a08_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
REVH5a04_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
REVH5a04_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
REVH5a01_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	
REVH5a02_1	QVQLVQSGTEVKKPGASVKVSKVSDYTFITTS	AMHWVROAPEQGPPEWMGWI	NAGNGDTKYSQKFQGRVTITRDT	SASTAYMELSSLGSEDTAVYFCAR	GNWYDQWG	

	FR1	CDR1	FR2	CDR2	FR3	CDR3
VH10.TX1_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RAVH10_C5_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RAVH10_C6_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RAVH10_F6_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RAVH10_H8_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RBVH10_A2_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RBVH10_B4_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RBVH10_C2_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RBVH10_D2_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
RBVH10_F4_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
REVH10C10_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	
REVH10D11_1	EVQLVQSGAEVKKPGASVKVSKVSDYTLTSL	SMHWVROAPGKGLEWMGGF	DPEDGETIY	AKKFQGRVTMTEDTSTDTAYMELSSLRSEDTAVYYCAT	IPWMYYFD	

Figure 2

VL6

QPVLTQPPSASGSPGQSVTISCTGTSSDVGDYNHVSQYQHPGKVPKLIIFEVNKRPSGVPDRFSGS
KSGNTASLTVSGLQPEDEADYYCSSFGGRNILFVFGTGTKLTVL

VL9

QSVLTQPPSVSAAPGQKVTISCSGGSSNIGNNYVSWYQQLPGTAPKLLIYEDNKRPSGIPDRFSGSK
SGTSATLGITGLQTGDEADYYCGTWDSLSAGVFGGGTKLTVL

VL18

QSVLTQPPSVSVAPGQTARITCGGDNVGSKSVHWYQKPGQAPVLVYEDSDRPSGIPERYSGSNS
GNTATLTISRVEAGDEADYYCQVWDGNSDHQVFGTGTKVTVL

Figure 3

[illegible][illegible]

Figure 4

VK5

DIVMTQSPESLGVSLGERATINCKSSQSVLYSYNNKNYVAWYQQKPGQPPKLLIYWASTRESGVPD
RFGSGESGTDFTLTISSLQAEDVAVYYCHQYYSVPTFGPGTKVDIK

VK11

EIVLTQSPGTLSPGERATLSCRASQSVNINYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGS
GTDFTLTISRLEPEDFAVYYCQQYGSSPRTFGQGTKVDIK

Figure 5

WKS_TXT_01VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS5g12_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS5g2_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS901_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS301_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS5g2_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS12_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS10_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS10_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS11_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS03_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS5g05_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS5g07_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS12_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS12_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS05_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
RAVKS05_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK
REVKS08_1VMTQSPESLGSVLSGERATINCKSSQSVLSYNNKNKYVAVYQQKPQOPPKLLIYWASTRESGVPRDFS0SES0TDFTLT1SSLOAEDVAVYCHQYYSVPFTFGPGTKVDIK

[illegible]

Figure 6

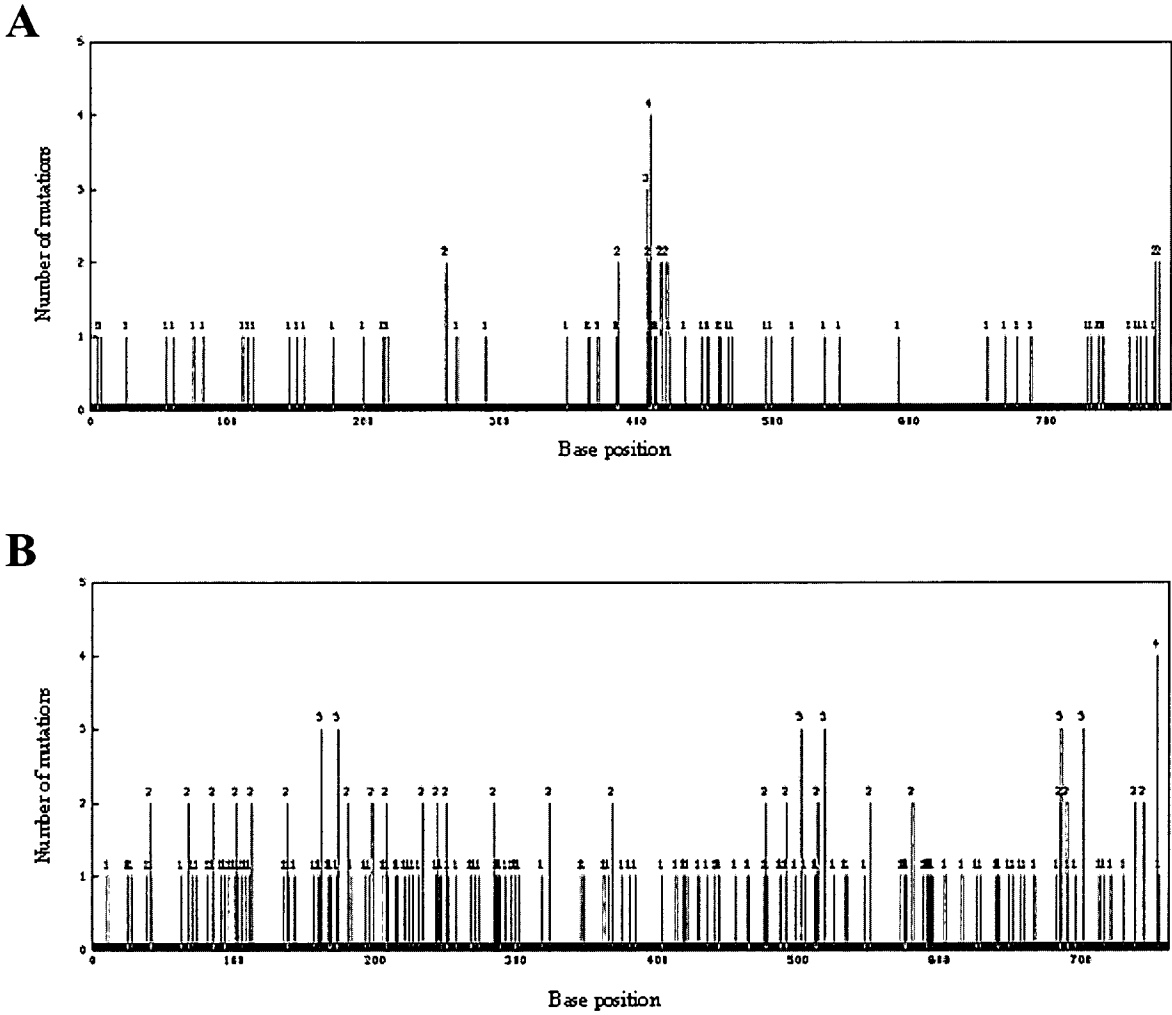


Figure 7

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2006/006339

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/10 C40B40/02

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N C40B

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

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

EPO-Internal, BIOSIS, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GRAM H ET AL: "In vitro selection and affinity maturation of antibodies from a naive combinatorial immunoglobulin library" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, DC, US, vol. 89, no. 8, 15 April 1992 (1992-04-15), pages 3576-3580, XP000384398 ISSN: 0027-8424 page 3576, column 2, line 23 - page 3579, column 1, line 8; figure 2	1-5,9-18
Y	page 3578, column 1, line 16 - line 40 ----- -/-	6-8



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

27 March 2007

Date of mailing of the international search report

29. 03. 2007

Name and mailing address of the ISA/

European Patent Office, P.B. 5510 Patentlaan 2
NL - 2280 HV Rijswijk
Tel: (+31-70) 340-2040, Tx, 81 851 epo nl,
Fax: (+31-70) 340-3016

Authorized officer:

Pérez-Mato, Isabel

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2006/006339

O(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MIYAZAKI C ET AL: "Changes in the specificity of antibodies by site-specific mutagenesis followed by random mutagenesis." PROTEIN ENGINEERING. MAY 1999, vol. 12, no. 5, May 1999 (1999-05), pages 407-415, XP002413822 ISSN: 0269-2139 page 410, column 2, line 3 - page 411, column 1, line 28	1-3,9-18
X	----- CASSON L P ET AL: "Evaluation of loss and change of specificity resulting from random mutagenesis of an antibody VH region." JOURNAL OF IMMUNOLOGY (BALTIMORE, MD. : 1950) 15 DEC 1995, vol. 155, no. 12, 15 December 1995 (1995-12-15), pages 5647-5654, XP002413823 ISSN: 0022-1767 the whole document	1-3,9-18
X	----- SAVIRANTA P ET AL: "Engineering the steroid-specificity of an anti-17beta-estradiol Fab by random mutagenesis and competitive phage panning" PROTEIN ENGINEERING, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 11, February 1998 (1998-02), pages 143-152, XP000941075 ISSN: 0269-2139 page 147, column 1, line 24 - page 149, column 2, line 22; table 1	1-3,9-18
X	----- DAUGHERTY P S ET AL: "Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. 29 FEB 2000, vol. 97, no. 5, 29 February 2000 (2000-02-29), pages 2029-2034, XP002413824 ISSN: 0027-8424 page 2030, column 1, line 22 - page 2032, column 2, line 3; figure 2 -----	1,4,5,9-18

-/--

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2006/006339

G(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HAWKINS R E ET AL: "The Contribution of Contact and Non-contact Residues of Antibody in the Affinity of Binding to Antigen" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 234, 1993, pages 958-964, XP002920236 ISSN: 0022-2836 page 960, column 1, line 3 - line 26; table 2	1,4,5, 9-18
Y	----- KOEFOED K ET AL: "Molecular characterization of the circulating anti-HIV-1 gp120-specific B cell repertoire using antibody phage display libraries generated from pre-selected HIV-1 gp120 binding PBLs" JOURNAL OF IMMUNOLOGICAL METHODS, ELSEVIER SCIENCE PUBLISHERS B.V.,AMSTERDAM, NL, vol. 297, no. 1-2, February 2005 (2005-02), pages 187-201, XP004793118 ISSN: 0022-1759 page 192, column 1, line 1 - column 2, line 2	6,8
Y	----- SALOMONSSON STINA ET AL: "Cloning and characterization of two human Ro52-specific monoclonal autoantibodies directed towards a domain associated with congenital heart block." JOURNAL OF AUTOIMMUNITY, MAR 2004, vol. 22, no. 2, March 2004 (2004-03), pages 167-177, XP002413825 ISSN: 0896-8411 page 173, column 2, line 12 - page 175, column 1, line 3	6,8
Y	----- YIP Y L ET AL: "Evaluation of different lymphoid tissue sources for the construction of human immunoglobulin gene libraries" IMMUNOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS BV, NL, vol. 3, no. 3, October 1997 (1997-10), pages 195-203, XP004097003 ISSN: 1380-2933 page 196, column 1, line 10 - page 202, column 2, line 4	6,8
Y	----- US 2004/110294 A1 (BOUAYADI KHALIL [FR] ET AL) 10 June 2004 (2004-06-10) cited in the application page 1, paragraph 14 - page 2, paragraph 19 -----	7

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2006/006339

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004110294 A1	10-06-2004	AU 2377202 A	21-05-2002
		EP 1332210 A1	06-08-2003
		FR 2816319 A1	10-05-2002
		WO 0238756 A1	16-05-2002
