Nucleic Acids Research, 2005, Vol. 33, No. 9 ***e81***
doi:10.1093/nar/gni080

Oligonucleotide-assisted cleavage and ligation: a novel directional DNA cloning technology to capture cDNAs. Application in the construction of a human immune antibody phage-display library

Sonia Schoonbroodt, Nicolas Frans, Mark DeSouza[1], Rachel Eren[2], Smadar Priel[2], Naama Brosh[2], Judith Ben-Porath[2], Arie Zauberman[2], Ehud Ilan[2], Shlomo Dagan[2], Edward H. Cohen[1], Hennie R. Hoogenboom, Robert Charles Ladner[1] and René M. Hoet*

Dyax s.a., Boulevard du Rectorat 27B, Sart Tilman, B-4000 Liege 1, Belgium, [1]Dyax Corp., Corporate Headquarters, 300 Technology Square, Cambridge, MA 02139, USA and [2]XTL Biopharmaceuticals Ltd, Kiryat Weizmann, Rehovot, 76103, Israel

Received December 7, 2004; Revised March 29, 2005; Accepted April 28, 2005

ABSTRACT

The use of oligonucleotide-assisted cleavage and ligation (ONCL), a novel approach to the capture of gene repertoires, in the construction of a phage-display immune antibody library is described. ONCL begins with rapid amplification of cDNA ends to amplify all members equally. A single, specific cut near 5′ and/or 3′ end of each gene fragment (in single stranded form) is facilitated by hybridization with an appropriate oligonucleotide adapter. Directional cloning of targeted DNA is accomplished by ligation of a partially duplex DNA molecule (containing suitable restriction sites) and amplification with primers in constant regions. To demonstrate utility and reliability of ONCL, a human antibody repertoire was cloned from IgG mRNA extracted from human B-lymphocytes engrafted in Trimera mice. These mice were transplanted with peripheral blood lymphocytes from *Candida albicans* infected individuals and subsequently immunized with *C.albicans* glyceraldehyde-3-phosphate dehydrogenase (GAPDH). DNA sequencing showed that ONCL resulted in efficient capture of gene repertoires. Indeed, full representation of all V_H families/segments was observed showing that ONCL did not introduce cloning biases for or against any V_H family. We validated the efficiency of ONCL by creating a functional Fab phage-display library with a size of 3.3×10^{10} and by selecting five unique Fabs against GAPDH antigen.

INTRODUCTION

Phage-display has been used to generate and select libraries of antibody fragments in scFv or Fab formats (1–7). Non-immune or immune antibody repertoires are routinely cloned from human B-lymphocyte mRNAs using RT–PCR of V_H and V_L genes (8–14). Although the use of PCR technology (15,16) has accelerated the cloning, sequencing, and characterization of genes, the cloning and sequencing of Ig variable (V) genes by PCR has been challenging, primarily due to the great diversity of Ig V region genes. Indeed, the use of PCR priming regions on each side of the targeted Ig V region compromises the equal amplification of all immunoglobulin mRNAs because there are no constant 5′ end mRNA sequences. Over the last decade, several methods have been established to capture the widest repertoire of immunoglobulin genes. One PCR-based cloning approach of immunoglobulin cDNAs used degenerate consensus sequences as 5′ end primers (17). This method risks introducing biases in the capture of the different Ig family repertoires, largely due to the varying efficiency of mRNA amplification among these sequences. As a result, the first 6–8 amino acids of the variable heavy and light chain sequence can differ from the original mRNA sequence which can affect the distribution, the expression and the functionality (antigen binding) of the antibody (18,19). Anchor PCR, a second and more general method, tails the 3′ end of the first strand of cDNA with oligo(dG) to provide a template for reverse priming with oligo(dC) (20,21). In this case, the length of the tail formed with terminal deoxynucleotidyl transferase may vary from a few bases to >200 bases, complicating identification of the appropriate size of the amplified cDNA. Since all methods described above use PCR technology to

---

*To whom correspondence should be addressed. Tel: +32 4 364 2405; Fax: +32 4 364 2499; Email: rhoet@dyax.com

capture antibody diversity, they potentially carry biases introduced by PCR oligonucleotides due to the lack of sequence uniformity in the antibody V-region.

Alternative antibody cloning strategies that do not rely on the priming of oligonucleotides within the antibody V-region, such as PCR adaptors (22), CapFinder (23) and rapid amplification of cDNA ends (RACE) technologies (24,25), have been described. However, those techniques are subject to high background interference and do not provide antibody gene expression in the correct reading frame. Indeed, Schramm *et al.* (26) have described residual primer contamination from the initial reaction resulting in the amplification of non-desired cDNA fragments. To address such contamination problems, Roeder (27) implemented the RACE approach with a solid-phase cDNA synthesis.

In this paper, we describe a method to efficiently clone and capture families of genes, in particular antibody genes. This technique (i) involves amplification using oligonucleotides that prime outside the variable part of the genes (antibody V-region), (ii) has low background, and (iii) allows expression of the captured genes in a correct reading frame.

The utility of this new method is illustrated by the construction of an immune phage-display antibody library directed against *Candida albicans* cell-wall associated glyceraldehyde-3-phosphate dehydrogenase (GAPDH) protein. GAPDH was selected as the target protein because it is expressed on the surface of *Candida* cells. Therefore, *C.albicans*-specific anti-GAPDH antibodies are potential therapeutic leads for the treatment of *Candida* infections. We have used the Trimera mouse technology to obtain an immunized human antibody repertoire (28–30). Trimera mice engrafted with functioning human peripheral blood lymphocytes (PBLs) and immunized with a specific antigen generate antigen-specific human antibodies. In this pilot study, we used Trimera mice engrafted with human PBLs of *C.albicans*-infected patients. The mice were subsequently immunized with a recombinant baculovirus-produced *C.albicans* GAPDH in order to obtain a human anti-*C.albicans* GAPDH immune antibody repertoire. The oligonucleotide-assisted cleavage and ligation (ONCL) method was then used to clone the human V genes, and a Fab phage-display library with ~33 billions Fab clones was built. Because all of the PCR priming sites are outside the variable regions, the method does not introduce variations in efficiency of amplification for different Ig mRNAs. In addition, as the cleavage step involves no amplification, the generation of biases is unlikely. The efficiency of the technology was further demonstrated by Ig gene sequencing followed by selection of anti-GAPDH specific antibodies.

MATERIALS AND METHODS

Expression, purification and biotinylation of *C.albicans* GAPDH

The full-length cDNA of *C.albicans* GAPDH was provided by Dr D. Gozalbo (University of Valencia, Spain) (31) and expressed in the Baculovirus gold system (Pharmingen). GAPDH was purified via nickel agarose chromatography (Qiagen, Hilden, Germany). The biotinylation ratio for GAPDH was five molecules of NHS-SS-biotin (Pierce) per molecule of antigen, in accord with the supplier's recommendations. Biotinylation was checked using non-reducing PAGE and confirmed by western blot using streptavidin horseradish peroxidase for detection.

Peripheral blood lymphocytes

Human PBLs were obtained by leukopheresis from four female donors (age 18–55) who had experienced at least one episode of vaginitis caused by *Candida* species. Immune response to *C.albicans* GAPDH in each donor was evidenced by the presence of anti-GAPDH IgG-specific antibodies in the serum, as well as significant T-cell proliferative response. The PBLs were separated on Ficoll IsoPrep (Robbins Scientific Corporation), washed twice, counted and re-suspended in phosphate-buffered saline (PBS). For *in vitro* activation, one-quarter of the cells were incubated for 16 h in RPMI medium supplemented with 10% fetal calf serum (FCS) (Biological Industries), glutamine (200 mM), sodium pyruvate (100 mM), 1% Eagle's non essential amino acids, 0.1% HEPES buffer and β-mercaptoethanol in the presence of 10 μg/ml GAPDH and Pokeweed mitogen (Gibco) diluted 1:100.

Mice

BALB/c mice, 6–10 weeks old, were obtained from the Weizmann Institute Animal Breeding Center (Rehovot, Israel). All mice were fed sterile food and given acidified water containing ciprofloxacin (20 μg/ml) (Bayer, Leverkusen, Germany). Mice were injected daily with 1 mg Fortum (Glaxo Operations UK, Greenford, England) intraperitoneally (i.p.) for 5 days post-bone marrow transplantation.

The Trimera system

BALB/c mice were exposed to split-dose total body irradiation (4 Gy followed 3 days later by 10–11 Gy), from a gamma beam 150 Å ^{60}Co source produced by Atomic Energy of Canada (Kanata, Ontario, Canada) with focal skin distance of 75 cm and a dose rate of 0.7 Gy/min. Following irradiation, each recipient mouse was immediately injected intravenously (i.v.) with $4–6 \times 10^6$ of SCID/NOD bone marrow cells and i.p. with 100×10^6 human PBLs (28). On the day of PBL transplantation, Trimera mice were immunized once i.p. with 150 μg per mouse of GAPDH absorbed in 0.15% aluminum hydroxide gel (Merck, Darmstadt, Germany). One day after transplantation, mice were boosted with 15×10^6 *in vitro* activated human PBLs and 75 μg per mouse GAPDH. On the following day, mice were boosted again with 150 μg/mouse GAPDH. Animals were bled from the retro-orbital vein using heparin-coated glass capillaries. Plasma was kept for human immunoglobulin determination. Animals were killed by cervical dislocation; their spleens were removed, cut into pieces, then pressed through stainless steel sieves to make a cell suspension in PBS.

Fluorescence-activated cell sorter (FACS) analysis of engrafted human cells

Spleen cells were incubated for 30 min on ice with mouse anti-human CD45-fluorescein isothiocyanate (pan-leukocyte antigen, Caltag Laboratories, South San Francisco, CA). After washing, fluorescent analysis was performed on a FACScan analyzer (Becton Dickinson, Mountain View, CA).

Determination of human immunoglobulin and antigen specific antibodies

Trimera mice sera were tested for total human immunoglobulin and for antigen-specific antibodies. Total human immunoglobulin was quantified by sandwich enzyme linked immunosorbent assay (ELISA) using purified goat $F(ab)_2$ anti-human IgG (Chemicon International, Temecula, CA) as the capture agent (1 μg/well) and purified goat anti-human IgG peroxidase conjugates (Chemicon International) as a detection reagent. Purified human IgG was used as the standard (Sigma, Rehovot, Israel). Concentration of anti-GAPDH antibodies in human or mice sera was determined by an antigen-specific ELISA. Microtiter plates (Nunc, Roskilde, Denmark) were coated with 10 μg/ml of GAPDH in PBS for 18 h at 4°C. The plates were blocked with PBS, 0.1% Tween-20 and 20% FCS, for 1 h at room temperature and reacted with human sera or Trimera mice sera diluted 1:20 and 1:200, respectively, for 2 h at room temperature. Bound anti-GAPDH antibodies were detected with a 1:10 000 dilution of goat anti-mouse/human peroxidase conjugated polyclonal antibodies (Chemicon International). Following 1 h incubation and subsequent washings, substrate solution (TMB, Chemicon International) was added. Results were read at a wavelength of 450 nm using an ELISA reader.

RNA extraction and mRNA isolation

Spleens from 12 immunized Trimera mice were either frozen in liquid nitrogen immediately after removal from the animal and placed at −70°C for long-term storage or immersed in 1.6 ml of RNAlater (Qiagen) and kept at 4°C for no longer than 1 month. Spleens were homogenized in 4 ml of RNAzol B, using a Kinematica Polytron PT 2100 homogenizer. Chloroform was added to each sample up to 20% of the final volume and the samples were left on ice for 5 min. Samples were then centrifuged for 15 min at 4°C and an equal volume of isopropanol was added to the upper phase. The RNA was left to precipitate on ice for 15 min, a step that was followed by 15 min of centrifugation at 21 900 *g*. The RNA pellet was washed with 75% ethanol, dried and re-suspended in 100 μl of DEPC-treated water. Optical density was measured using GeneQuant-pro (Amersham Pharmacia), and the RNA integrity was visualized on a 1.2% agarose TAE gel. One-tenth volume of 3 M sodium acetate and 2.5 vol of ethanol were added to the samples and stored at −70°C. Recovery of the RNA was carried out by 10–15 min centrifugation at 4°C followed by 70% ethanol wash of the pellet and re-suspension in DEPC-treated water to the desired volume. The RNA amount recovered from each spleen ranged between 80 and 220 μg. For mRNA isolation, 4 μg of total RNA from each of the 12 spleen samples was submitted to the PolyATtract® mRNA Isolation Systems (Promega) according to the manufacturer's instructions.

Capture and amplification of variable region genes

Cleavage of the pyrophosphate bond of the 5′ terminal methylated guanine nucleotide mRNA 'cap' was achieved using Tobacco Acid Pyrophosphatase enzyme (TAP) (Epicentre Technologies): in practice, 400 ng mRNA was treated for 60 min at 37°C with 4 U of TAP in TAP reaction buffer [50 mM sodium acetate (pH 6.0), 1 mM EDTA, 0.1% β-mercaptoethanol and 0.01% Triton X-100 supplemented with 160 U of RNAsin (40 U/μl) (Promega)]. Decapped mRNA was then purified over an RNAeasy column (Qiagen). Next, 100 ng of decapped mRNA was heat-denatured for 5 min at 65°C in the presence of 240 ng (50× molar excess) of an RNA adapter (5′-GCUGAUGGCG AUGAAUGAAC ACUGCGUUUG CUGGCUUUGA UGAAA-3′), then cooled for 2 min on ice. Ligation was performed for 1 h at 37°C using 10 U of T4 RNA ligase in ligation buffer (Promega), supplemented with 10 U RNAsin (40 U/ml) (Promega) and was followed by purification over a RNAeasy column (Qiagen). Subsequently, 50 ng of the ligated RNA was converted into cDNA: 100 pmol oligo$(dT)_{15}$ primer (Roche), buffer and DTT were added to the mixture according to the supplier's instructions (Life Technologies, Inc.), as well as 400 μM dNTPs (Amersham), 40 U of RNAsin (40 U/μl; Promega) and 200 U of Moloney murine leukemia virus reverse transcriptase (Life Technologies, Inc.) in a total volume of 25 μl. After 1 h at 42°C, enzyme was heat-inactivated at 70°C for 15 min, and cDNA was immediately used for PCR amplification. Primary PCR amplification was carried out on 3 μl of the cDNA reaction volume using a 5′ primer complementary to the RNA adapter tailed sequence (Outer Inv primer) combined with a 3′ primer either complementary to the human IgG-derived heavy chain (CH1IgGXTLfor) or to the κ- or λ-light chain constant regions (HuCκforAsc, Cλ2,7forAsc). PCR was performed in a volume of 25 μl using the Advantage 2 PCR enzyme system (Clontech) and 10 pmol of each primer for 35 cycles (1 min at 95°C, 1 min at 60°C and 2 min at 68°C). The PCR products were re-amplified with a combination of a 5′ end biotinylated nested primer (Inner Inv primer) and a 3′ end NheI-tagged nested CH1 reverse primer for the heavy chains, and κ- or λ-light chain constant region primers identical to the first PCR. Similar to primary PCR, the second PCR was performed in a volume of 50 μl using the Advantage 2 PCR enzyme system (Clontech) and 10 pmol of each primer for 25 cycles (1 min at 95°C, 1 min at 60°C and 2 min at 68°C); 100 reactions were performed. All biotinylated products were purified from an agarose gel with the GFX extraction kit (Amersham).

ONCL cloning method

(i) *DNA immobilization.* An aliquot of 6 μg of double-stranded DNA (dsDNA) biotinylated on the 5′ end of the top strand was bound to 900 μl of streptavidin-coupled magnetic beads (Seradyn) in 1 ml B&W buffer [5 mM Tris–HCl (pH 7.5), 0.5 mM EDTA, 1 M NaCl and MQH_2O] for 30 min under rotation at room temperature. Unbound DNA was washed away two times with 1 ml of 1× B&W buffer.

(ii) *Conversion of DNA to single-stranded (ss) form.* DNA immobilized on the beads was incubated with 900 μl 0.1 M NaOH for 3 min, washed once with 1 ml of 0.01 M NaOH followed by one additional washing step using 1 ml Tris–NaCl buffer [10 mM Tris–HCl (pH 7.5), 100 mM NaCl].

(iii) (a). *ONCL cleavage of the light chains.* During a controlled temperature decrease (from 90 to 50°C at 1°C per 45 s), single-stranded DNA (ssDNA) was annealed to the set of adapters shown in Table 1 (40× molar excess for κ chains and 80× molar excess for λ chains) in the appropriate restriction enzyme buffer (New England Biolabs). After annealing,

Table 1. Oligonucleotide primers, adapters and bridge–extenders used for library construction

Oligonucleotide name	Oligonucleotide sequence	Restriction enzyme
Outer Inv primer	GCTGATGGCGATGAATGAAC	
Inner Inv primer	GAACACTGCGTTTGCTGGC	
Adapter for Vλ family 1	GAGGGTGGCT *GAGTC* AGCAC	HinfI
Adapter for Vλ family 2	GAGGCAGGCT *GAGTC* AGGGC	HinfI
Adapter for Vλ family 3 (1)	GCAGGGTCCT *GAGTC* AGCTC	HinfI
Adapter for Vλ family 3 (2)	GAGGGTGGCT *GAGTC* AGCTC	HinfI
Bridge for Vλ family 1	GGCTGAGTCAAGACGCTCT *GTGCAC* TTCGCTGTCGAGG	ApaLI
Extender for Vλ family 1	CCTCGACAGCGAA *GTGCAC* AGAGCGTCTTG	ApaLI
Bridge for Vλ family 2	GGCTGAGTCAAAGCGCTCT *GTGCAC* TTCGCTGTCGAGG	ApaLI
Extender for Vλ family 2	CCTCGACAGCGAA *GTGCAC* AGAGCGCTTTG	ApaLI
Bridge for Vλ family 3 (1)	GGCTGAGTCAATTCGCTCT *GTGCAC* TTCGCTGTCGAGG	ApaLI
Extender for Vλ family 3 (1)	CCTCGACAGCGAA *GTGCAC* AGAGCGAATTG	ApaLI
Bridge for Vλ family 3 (2)	GGCTGAGTCAATTCGTACT *GTGCAC* TTCGCTGTCGAGG	ApaLI
Extender for Vλ family 3 (2)	CCTCGACAGCGAA *GTGCAC* AGTACGAATTG	ApaLI
ONPlePCR	CCTCGACAGCGAA *GTGCAC* AG	ApaLI
Cλ2forAsc	ACCGCCTCCACCG *GGCGCGCC* TTATTATGAACATTCTGTAGGGGCCACTG	AscI
Cλ7forAsc	ACCGCCTCCACCG *GGCGCGCC* TTATTAAGAGCATTCTGCAGGGGCCACTG	AscI
Adapter for Vκ family 1 (1)	GGGAGGAT *GGAGAC* TGGGTC	BsmAI
Adapter for Vκ family 1 (2)	GGGAAGAT *GGAGAC* TGGGTC	BsmAI
Adapter for Vκ family 2	GGGAGAGT *GGAGAC* TGAGTC	BsmAI
Adapter for Vκ family 3 (1)	GGGTGCCT *GGAGAC* TGCGTC	BsmAI
Adapter for Vκ family 3 (2)	GGGTGGCT *GGAGAC* TGCGTC	BsmAI
Adapter for Vκ family 4	GGGAGTCT *GGAGAC* TGGGTC	BsmAI
Bridge for Vκ family 1 (1)	GGGAGGATGGAGACTGGGTCATCTGGATGTCTT *GTGCAC* TGTGACAGAGG	ApaLI
Bridge for Vκ family 1 (2)	GGGAAGATGGAGACTGGGTCATCTGGATGTCTT *GTGCAC* TGTGACAGAGG	ApaLI
Bridge for Vκ family 2	GGGAGAGTGGAGACTGGGTCATCTGGATGTCTT *GTGCAC* TGTGACAGAGG	ApaLI
Bridge for Vκ family 3 (1)	GGGTGCCTGGAGACTGGGTCATCTGGATGTCTT *GTGCAC* TGTGACAGAGG	ApaLI
Bridge for Vκ family 3 (2)	GGGTGGCTGGAGACTGGGTCATCTGGATGTCTT *GTGCAC* TGTGACAGAGG	ApaLI
Bridge for Vκ family 4	GGGAGTCTGGAGACTGGGTCATCTGGATGTCTT *GTGCAC* TGTGACAGAGG	ApaLI
Extender for all κ bridges	CCTCTGTCACA *GTGCAC* AAGACATCCAGATGACCCAGTCTCC	ApaLI
KaPCRt1	CCTCTGTCACA *GTGCAC* AAGAC	ApaLI
HuC κforAsc	ACCGCCTCCACCG *GGCGCGCC* TTATTAACACTCTCCCCTGTTGAAGCTCTT	AscI
Adapter for V_H1	CCAGACTGCAC *CAGCTG* CACC	PvuII
Adapter for V_H3 (1)	CCAGACTCCAC *CAGCTG* CACC	PvuII
Adapter for V_H3 (2)	CCAGACTCCAA *CAGCTG* CACC	PvuII
Adapter for V_H4 (1)	CCCGACTCCTG *CAGCTG* CACC	PvuII
Adapter for V_H4 (2)	CCCGACTCCTG *CAGCTG* CAGC	PvuII
Adapter for V_H6	CCTGACTGCTG *CAGCTG* TACC	PvuII
Adapter for V_H7	CCAGACTGTAC *CAGCTG* GACC	PvuII
Bridge for V_H1	GACTGCACCAGCTGAACTTCAGCCAT *GGCCGGCTGGGCC* TTCTACTAAACAGTC	SfiI
Bridge for V_H3 (1)	GACTCCACCAGCTGAACTTCAGCCAT *GGCCGGCTGGGCC* TTCTACTAAACAGTC	SfiI
Bridge for V_H3 (2)	GACTCCAACAGCTGAACTTCAGCCAT *GGCCGGCTGGGCC* TTCTACTAAACAGTC	SfiI
Bridge for V_H4	GACTCCTGCAGCTGAACTTCAGCCAT *GGCCGGCTGGGCC* TTCTACTAAACAGTC	SfiI
Bridge for V_H7	GACTGTACCAGCTGAACTTCAGCCAT *GGCCGGCTGGGCC* TTCTACTAAACAGTC	SfiI
Bridge for V_H6	GACTGCTGCAGCTGAACTTCAGCCAT *GGCCGGCTGGGCC* TTCTACTAAACAGTC	SfiI
Extender for all V_H	GACTGTTTAGTAGAA *GGCCCAGCCGGCC* ATGGCTGAAGTTCAG	SfiI
HCP2PCR2	GACTGTTTAGTAGAAGGCCCAGCC	
Adapter for V_H2 (1)	CACCAGCGTA *GGACC* AGACTCCTT	AvaII
Adapter for V_H2 (2)	CACCAGCACA *GGACC* AGACTCCTT	AvaII
Bridge for V_H2 (1)	AGCGTAGGACCGCTTTCTTTAAGAGTAATTTGGGCCAT *GGCCGGCTGGGCC* TAGACAAGTGTGGAAG	SfiI
Bridge for V_H2 (2)	AGCACAGGACCGCTTTCTTTAAGAGTAATTTGGGCCAT *GGCCGGCTGGGCC* TAGACAAGTGTGGAAG	SfiI
Extender for V_H2	CTTCCACACTTGTCTA *GGCCCAGCCGGCC* ATGGCCCAAATTACTCTTAAAGAAAGCG	SfiI
HC2PCRAva	CTTCCACACTTGTCTAGG	
NheICH1IgGfornested	GCCGATCTAG *GCTAGC* GGGAAGACCGATGGGCCCTTGG	NheI
Ch1IgGXTLfor	GCTGCTGAGGGAGTAGAGTGG	

Note: All sequences are written from 5′ to 3′. Enzymatic recognition sites are in italics. Arabic numbers in parentheses indicate pairings where multiple sequences occur in the germlines within a family.

the unbound adapters were washed away. The DNA–adapter complex bound on beads was then incubated in 250 μl of NEB restriction enzyme buffer containing 15 U BsmAI/μg ssDNA (for κ chains) or 50 U HinfI/μg ssDNA (for λ chains) for 1 h at the respective cleavage temperature (55°C for BsmAI and 50°C for HinfI) in order to release the light chain gene fragment into the supernatant.

(iii) (b). *ONCL cleavage of the heavy chains.* First, 0.1 M NaOH-denatured ssDNA was annealed to a set of adapters shown in Table 1 (35× molar excess for PvuII cleavage and 10× molar excess for AvaII cleavage) in presence of 1× NEB4 buffer (New England Biolabs) with controlled temperature decrease (from 90 to 50°C at 1°C per 45 s). After annealing, the unbound adapters were washed away. The

DNA–adapter complex bound on beads was then incubated in 250 µl of NEB restriction enzyme buffer containing 50 U PvuII/µg ssDNA or 80 U AvaII/µg ssDNA for 1 h at 50°C in order to release the heavy chain gene fragments in the supernatant.

(iv) *Bridge–extender ligation.* To create partially dsDNA, 500 pmol of a set of oligonucleotides were mixed with 500 pmol of the complementary set of oligonucleotides (Table 1, bridge and extender) at 95°C and then cooled to 16°C (at 1°C per 45 s) in 1× T4 DNA ligase buffer. Subsequently, 100 ng of the column-purified ssDNA was ligated to 100× molar excess of bridge–extender duplex. The ligation was performed in 75 µl using 200 U of T4 DNA ligase (NEB) and carried out for 16 h at 16°C. After ligation, the bridge–extender–DNA complex was purified using Qiagen PCR purification kit (Qiagen) and eluted in 50 µl EB buffer (10 mM Tris–HCl, pH 8.0).

(v) *Amplification and cloning of the bridge–extender DNA.* This step was performed for 13 cycles (light chains) or 15 cycles (heavy chains) in order to maintain maximal diversity in a 25 µl volume format using Advantage 2 DNA polymerase (Clontech) (1 min at 95°C, 1 min at 60–65°C and 2 min at 68°C). Sequences of the primers asserting restriction enzyme sites for directional cloning of the fragments are disclosed in Table 1 (KaPCRt1 and HuCκforAsc for κ light chains; ONPlePCR and Cλ2,7forAsc for λ light chains; HCP2PCR2 (PvuII) or HC2PCRAva (AvaII) and NheI-CH1IgGfornested for heavy chains).

Construction of the Fab library

For the construction of the primary light chain repertoires, the VκCκ and VλCλ PCR products, appended with the ApaLI and AscI restriction sites, were digested using 50 U/µg of each enzyme and agarose gel purified. Using T_4 DNA ligase (NEB), 1.5 µg of each resulting DNA fragment was ligated into 2.5 µg of similarly cut phagemid vector pMID21 (Figure 1). Subsequently, 400 ng of desalted λ-ligation mixture and 500 ng of the κ mixture were electroporated separately into the *Escherichia coli* strain TG1 using 10 ng of ligation mixture per electroporation event. The Fab library was obtained by cloning of the V_H repertoire into each of the light chain sublibraries. The V_HCH1 PCR products, appended with the SfiI and NheI restriction sites, were digested using 50 U/µg of each enzyme and agarose gel purified. Using T_4 DNA ligase (NEB), 6.7 µg of the V_HCH1 DNA fragments were ligated into 5 µg of the similarly cut λ or κ sublibrary DNA. Finally, 1250 ng of desalted λHC-ligation mixture and 1250 ng of the κHC one were electroporated into the *E.coli* strain TG1 using 25 ng of ligation mixture per electroporation event.

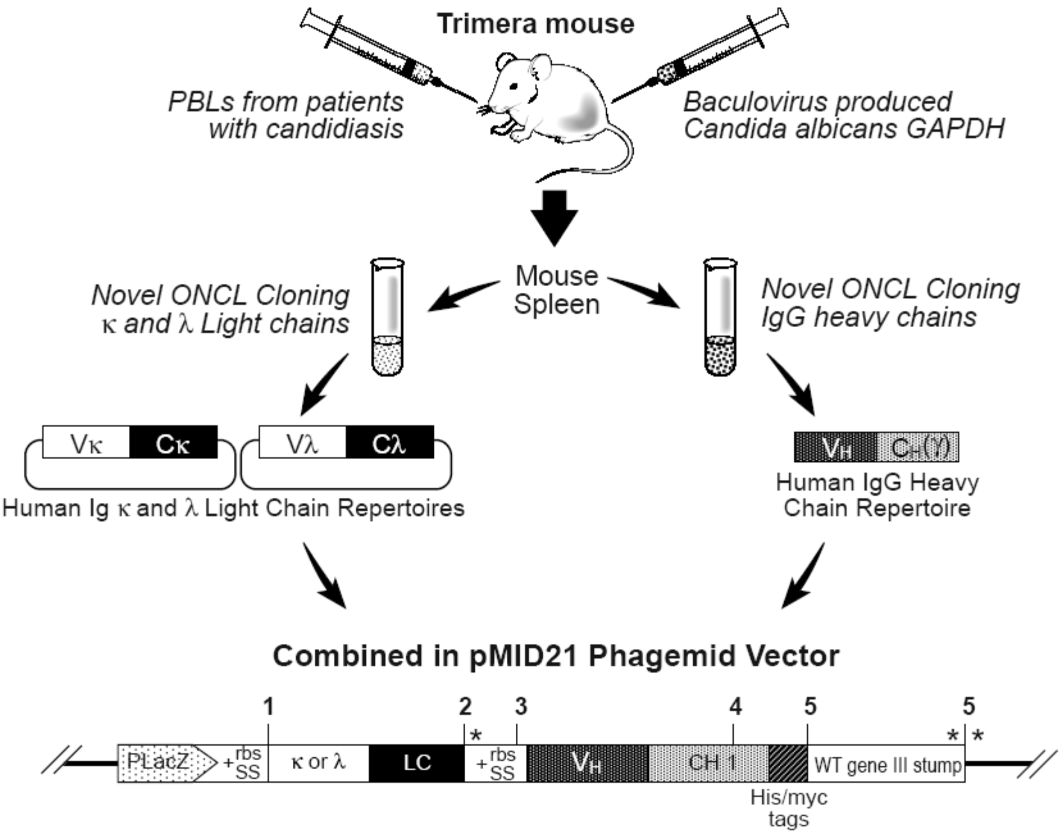

Figure 1. Overview of the library construction—combination of Trimera mouse technology, ONCL method and phage-display. Lower part shows a schematic representation of phagemid vector pMid21 used for display of antibody Fab fragments. This vector is derived from pCES1 (8). The polylinker region comprises two signal sequences, Cκ domain, ribosome binding site (rbs), CH1 domain, hexahistidine tag and a c-myc-derived sequence (His/myc tags). Light chain genes can be cloned as ApaLI–AscI (1,2) fragments and variable heavy chain genes as SfiI–NheI (3,4) fragments. The MluI (5) restriction sites flanking both sites of the *M13* gene III stump enable the removal of the stump anchor domain, which allows the production of soluble Fab fragments. Expression of the bicistronic operon is under control of the LacZ promoter (pLacZ).

Selection of the library

The rescue of the phagemid particles with helper phage M13–KO7 was performed according to Marks *et al*. (3) on a 2.3-l scale. For selection, 10^{13} phages were used with GAPDH antigen immobilized on immunotubes (Maxisorp tubes, Nunc) (3) or with soluble biotinylated GAPDH (31). The antigen concentration for selection on immunotubes was 70 μg/ml during all three rounds and 580 and 200 pmol for selection with the biotinylated antigen for round 1 and rounds 2 and 3, respectively.

Screening and sequencing of clones

Phage-displaying Fabs were produced from 94 individual clones as previously described (3). Culture supernatants were tested by ELISA with biotinylated antigen indirectly captured via immobilized BSA–streptavidin. Coating of the ELISA plates and the assay itself were performed as previously described (8). Fabs from clones giving a positive signal in ELISA (more than two times > background) were amplified using 5′ and 3′ backbone primers, and the PCR products were sequenced for both light and heavy chains. To allow soluble Fab expression, the positive phage antibodies were then recloned in order to remove the Fab-fused gene III anchor fragment (Figure 1). In practice, 1 μg of plasmid DNA was digested using 10 U/μg MluI restriction enzyme. The gene III anchor domain was then removed by agarose gel purification. Resulting phagemid vector was then ligated using T_4 DNA ligase (NEB) and transformed into *E.coli* TG1 strain.

Large-scale induction of soluble Fab fragments from individual clones was first performed on a 50 ml scale in 2× TY containing 100 μg/ml ampicillin and 2% glucose. After overnight growth at 30°C, 500 ml of 2× TY with 100 μg/ml ampicillin and 0.1% glucose were seeded with this overnight culture to reach an OD_{600} of 0.05 (1/100). After growth at 37°C to an OD_{600} of 0.8, the medium was supplemented with 1 mM isopropyl-1-thio-β-D-galactopyranoside and cells were incubated at 30°C for additional 4 h. Bacteria were then pelleted by centrifugation (10 min at 2934 *g*); periplasmic fractions were prepared by resuspending the cell pellet in 5.3 ml of ice-cold TES (0.2 M Tris, 0.5 mM EDTA and 0.5 M sucrose, pH 8.0). After an incubation of 10 min on ice, TES concentration was decreased by adding 1 vol of TES/H_2O (1:3) for additional 20 min incubation on ice. The periplasmic fraction was then collected by spinning the cell debris at 8000 *g* for 20 min at 4°C. Crude extracts were then dialyzed overnight over PBS at 4°C.

V_H distribution analysis

This analysis was performed on the V_H gene samples captured and amplified following the procedures described above. However, the amplification step was performed with identical but unbiotinylated primers. Thirty nanograms of this V_H gene PCR product were cloned using TOPO TA Cloning® for sequencing (Invitrogen) according to the manufacturer's instructions. The sequences of 115 clones were determined.

RESULTS

Immunization of Trimera mice with *C.albicans* GAPDH

The rationale for building a phage-display library from immunized Trimera mice is that display on phage and selection of the IgG human immune response in Trimera mice will allow selection of the desired antibodies at a higher rate than by direct screening of hybridomas from these mice. Thus, Trimera mice were prepared by transplantation of human PBLs obtained from four individual leukopheresis donors into irradiated BALB/c mice, radio-protected with bone marrow from SCID mice. Soluble *C.albicans* GAPDH (baculovirus produced) was used to boost the Trimera mice immune response (Figure 1). Human cell engraftment measured as the percentage of human $CD45^+$ cells in Trimera mice spleens showed that mouse spleens contained 0.8–28% of human $CD45^+$ cells. The presence of human *C.albicans* GAPDH-specific IgG antibodies was confirmed in immunized Trimera mice sera by ELISA on days 14–17 after transplantation. The specific human IgG activity (nanograms anti-*C.albicans* GAPDH-specific antibodies per milligram human IgG) ranged between 300 and 1800 ng/mg in different Trimera mice as compared with ∼40 ng/mg in donors' sera. This suggests an induction of the immune response to GAPDH in the Trimera mice.

Library construction

Capture of V-gene diversity. A Fab phage antibody library was built from the spleen RNA of 12 responding mice. The 12 RNA samples were processed separately. The capture of V genes was performed in two main steps. First, decapped full-length mRNA extracted from spleen cells of Trimera mice was tailed at the 5′ end with a synthetic RNA oligonucleotide before the reverse transcription was performed. For specific capture and amplification of antibody genes, 3′ end primers complementary to the constant regions of the antibody genes and 5′-biotinylated primers complementary to the attached synthetic sequence (Figure 2A) were used. Priming at this synthetic 5′ sequence avoids the use of primers within the variable regions of the antibody genes that could generate biases against rare subclasses of V genes or ones mutated at the priming sites. Only IgG-derived V_H segments were amplified from newly synthesized cDNAs using a primer located in the CH1 of IgG1-4. From this step onwards, the 12 samples, previously processed separately, were pooled in equal amounts.

The ONCL method. The V_H and V_L genes are cloned using a novel technology, referred to as 'ONCL', which is depicted in Figure 2A. First, dsDNA biotinylated on one strand is immobilized on streptavidin-coated magnetic beads. ssDNA is then prepared by alkaline denaturation. Short dsDNA segments are created by annealing of the oligonucleotide to the immobilized ssDNA providing a restriction site that can be recognized by the related restriction enzyme. High temperatures (>50°C) were used for the cleavage in order to avoid secondary structures in the ssDNA that could result in cleavage at other instances of the site recognized by the restriction enzyme. The cleaved ssDNA is then prepared for cloning by ligation of a partially duplexed synthetic DNA adapter, referred to as a bridge–extender. The bridge–extenders encode sequences of 5′ human antibody V genes that supply a useful restriction enzyme recognition site and bring the captured sequences into proper register. The 3′ terminal single-stranded part of the bridge–extender is complementary to the 5′ end of the cleaved ssDNA fragments. Finally, the ligated product is amplified using primers at the 5′ end of ligated DNA (non-V gene related) and primers at the 3′-constant region of antibody genes. Each primer introduces a restriction enzyme recognition site so that the Ig repertoire can be cloned into the display vector in the proper orientation and reading frame.

In order to cover as many human V-region gene segments as possible, a set of several $V\lambda$ (4), $V\kappa$ (6) or V_H (9) oligonucleotides annealing to framework 1 of antibody genes were designed (Table 1). The enzyme cleavage location was chosen very near the 5′ end of the V-region to avoid altering the original antibody sequences as presented in the example of ONCL for $V\lambda$ family 1 genes (Figure 2B). The oligonucleotides were designed so that they allow efficient capture of all commonly used V-gene segments. Their design was based on the most recent sequence information provided by V-base (http://vbase.mrc-cpe.cam.ac.uk/). For the light chains, primers for only the main families 1, 2 and 3 were designed, representing ~90% of the light chain repertoire (32). An analysis of the matching efficiency of each oligonucleotide to germ-line genes was performed for the heavy and light chains. The set of oligonucleotides we designed allows the capture of all V_H germ-line gene segments, except segments 1-03 and 4-34 (which contain neither PvuII nor AvaII sites), representing 1 and 2.5% of the re-arranged heavy chain repertoire, respectively (32).

The cleavage results of the spleen-derived $V\lambda$, $V\kappa$ and V_H repertoires are presented in Figure 3. λ Light chains were captured using a set of four oligonucleotides containing a HinfI enzymatic cleavage site. Following digestion of the single-stranded λ light chains immobilized on beads by HinfI (left panel A, column 2), a cleaved fragment of ~600 bp appeared in the supernatant fraction (column 3). Gel densitometry analysis revealed that 70% of captured λ light chain molecules were recovered (compare column 2 with column 3). K light chains were captured using a set of six

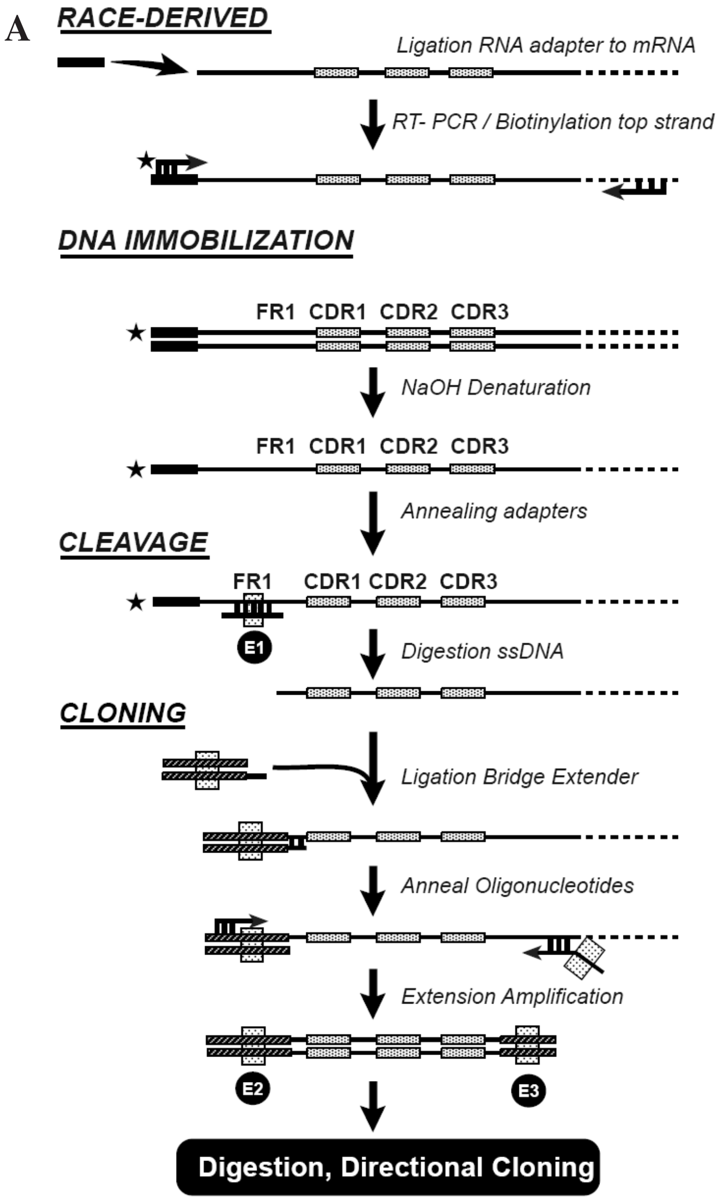

Figure 2. The ONCL technology. (**A**) Schematic outline of the ONCL technology. RACE: decapped mRNA is ligated to a RNA adapter, then converted to cDNA by RT–PCR. cDNA is then amplified using a 5′-biotinylated primer complementary to the adapter sequence combined with a 3′-primer either complementary to the human IgG-derived heavy chain constant region or to the light chain constant regions. DNA immobilization: double-stranded (ds) biotinylated RACE-derived PCR products are bound to streptavidin-coupled magnetic beads. After the DNA–beads complex is immobilized on a magnetic stand, ssDNA is prepared by NaOH denaturation. Cleavage of ssDNA: annealing of the adapter oligonucleotides to the top strand retained on the beads creates a dsDNA region accessible for the restriction enzyme (E1). Cleaved single-stranded V gene is released from the beads and can now be used in the next step. Preparation of the V-gene ssDNA for cloning: The ssDNA is then ligated to a 100× excess of partially dsDNA made through the annealing of two oligonucleotides. The lower-strand tail of this oligonucleotide duplex is complementary to the 5′ end of the ssDNA, allowing their recognition during the ligation procedure. The ligated product is then amplified using primers appended with appropriate restriction sites (plain black arrows), allowing the directional cloning of the V genes into the phagemid vector. E represents restriction enzymes, vertical rectangles represent restriction sites, asterisks represent biotin, dark rectangles represent the tailed sequence and ‘base-adorned’ arrows represent base pairing of the oligonucleotides with the ssDNA. (**B**) Example of ONCL method: λ1 V-gene capture and cloning. (1) Annealing of adapter for family Vλ1 to Vλ1 genes, immobilized on beads—cleavage by HinfI; (2) release of cleaved Vλ1 genes in supernatant; (3) annealing and ligation of cleaved Vλ1 genes to hybridized Vλ1 bridge and extender (ApaLI site underscored); (4) amplification of ligated DNA using ONPlePCR and Cλ2,7forAsc primers; (5) directional cloning into pMid21 of Vλ1 genes via ApaLI and AscI. X, non-V-gene related amino acid sequence.

oligonucleotides containing a BsmAI enzymatic cleavage site. Following digestion of the single-stranded κ light chains immobilized on beads by BsmAI (right panel A, column 2), a cleaved fragment of ∼600 bp appeared in the supernatant fraction (column 3). Gel densitometry analysis revealed that 90% of captured κ light chain molecules were recovered (compare column 2 with column 3). All heavy chain families except family 2 were captured using a set of seven oligonucleotides containing a PvuII enzymatic cleavage site (V_H families 1, 3, 4, 5, 6 and 7), while V_H family 2 was captured using a set of two oligonucleotides carrying an AvaII restriction recognition site. Following digestion of the single-stranded heavy chains immobilized on beads by PvuII enzyme (left panel B, columns 1 and 2), a cleaved fragment of 400 bp is released in the supernatant fraction (column 3). ssDNA left on the beads was then submitted to an AvaII cleavage (right panel B, columns 1 and 2), and a fragment of 400 bp was cleaved and collected in the supernatant (column 3). Gel densitometry analysis of both gels revealed that, in total, 85% of captured heavy chain molecules were cleaved. PvuII-cleaved V genes were combined with AvaII-cleaved genes in the ratio 80:15, which is the observed ratio of the genes so cleaved.

All single-stranded cleaved fragments (λ, κ and heavy chains) were then ligated to partially double-stranded oligonucleotides (bridge–extenders), of which the 3′ end is complementary to the 5′ end of the cleaved products (Table 1). In a final amplification (13–15 cycles) of the ligated V genes, primers appended with restriction recognition sites enable the directional cloning of the V genes into the expression vector (ApaLI–AscI for light chains and SfI–NheI for heavy chains).

Phagemid library construction. A two-step cloning procedure was used to construct the library. The ONCL-cleaved Vκ and Vλ repertoires were cloned separately to maintain the V-gene

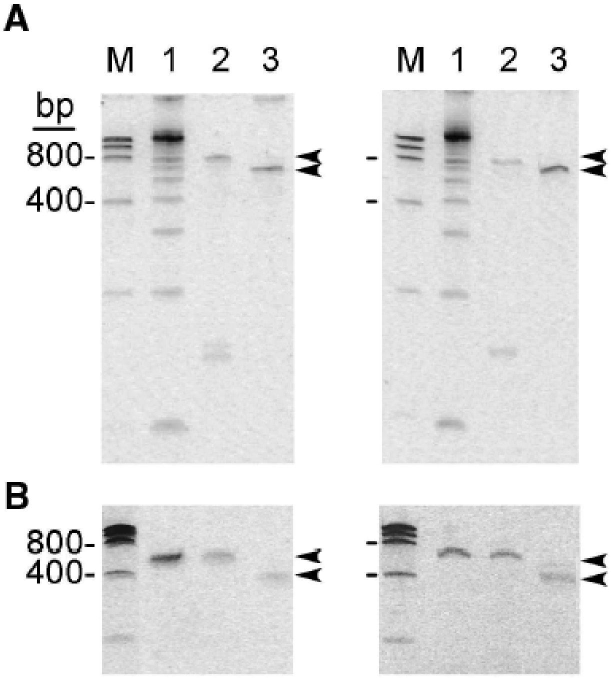

Figure 3. ONCL capture of V-gene repertoires. (**A**) Analysis of the cleavage efficiency of λ (left panel) and κ (right panel) light chains using the ONCL cleavage protocol outlined in Figure 2A. In lanes 2, ssDNA left on beads after the HinfI ONCL cleavage (λ chains, left panel) or the BsmAI cleavage (κ chains, right panel) and in lanes 3 cleaved ssDNA fragments released in the supernatant after the cleavage. The smaller band in both lanes 2 represents an artifact as it appears even in the uncleaved material (data not shown). Equivalent fractions of each samples were analyzed on 5% TBE-urea PAGE gels (Bio-Rad) together with the low DNA mass markers (M) and 100 bp (lanes 1) (Invitrogen). (**B**) Analysis of the cleavage efficiency of heavy chains after PvuII (left panel) and AvaII (right panel) ONCL cleavages. ssDNA immobilized on beads before any ONCL cleavage (lanes 1, left and right panels), ssDNA left on beads, respectively, after the PvuII ONCL cleavage (lane 2, left panel) or the AvaII one (lane 2, right panel), and cleaved ssDNA fragments released in the supernatant after the cleavage (lanes 3) are analyzed. Equivalent fractions of each sample were analyzed on 5% TBE-urea PAGE gels (Bio-Rad) together with the low DNA mass markers (M) (Invitrogen).

diversity, yielding one-chain libraries of typical size for libraries made by cloning of PCR fragments: 7.8×10^9 individual clones for the κ light chains and 2.3×10^9 clones for the λ ones into pMID21 phagemid vector. The heavy chain repertoire was then introduced into both light chain libraries. A Fab phage-display library carrying κ chains was then created of 5×10^9 transformants; the λ chain library contains 3.2×10^{10} Fab clones. PCR screening of 48 randomly selected clones showed that all carry full-length Fab-encoding inserts.

Quality control of the library by DNA sequencing. We evaluated the quality of the library by sequencing 84 clones. Our results showed that 70% of the Fab sequences are correct, while single or multiple mutations are present in 30% of the sequences. We found that this was mainly due to incorrect synthetic oligonucleotides or misannealing causing frameshifts.

The V_H distribution. To validate our technology, we first studied the distribution of the different V_H families in our final library. As a reference, we took the frequency of the V_H families in the IgG repertoire of healthy donors described by de Wildt *et al.* (32), where the authors showed that the V_H 1, 3 and 4 families are the most commonly represented in PBLs. We initially determined the sequences of the heavy chains of 35 clones randomly picked from our library. The sequences were aligned with the germ-line V_H gene database in order to determine their V_H family affiliation. We observed that 17% are representative for family V_H 1, 6% for V_H 2, 9% for V_H 3, 51% for V_H 4, 14% for V_H 5 and 3% for V_H 6. When compared with the distribution in healthy donors, our library had an over-representation of family 4, while family 3 is surprisingly under-represented, suggesting biases induced either through ONCL technology or caused by the modification of the human PBL expression in the Trimera mouse system. In order to understand the origin of this uncommon V_H distribution, we determined the sequences of 73 additional heavy chains of the ONCL library together with 115 heavy chains of the amplified V genes before the ONCL cleavage and aligned them to the germ-line V_H gene database. As depicted in Figure 4A, the V_H distributions observed before (RACE-derived clones) and after the ONCL cleavage are comparable, demonstrating that the V_H distribution we observed is not the consequence of our ONCL technology but most probably arose through the Trimera mouse system or the *C.albicans* infection. Furthermore, as shown in Figure 4B, some V_H gene segments (4-61, 4-39 and 4-59) seem to be over-represented when compared with 292 V_H sequences collected by Tomlinson *et al.* (33), showing that the pattern of usage of the segments in Trimera mice spleens (or in *C.albicans* infected patient PBLs) appears to differ from the usage in healthy donor samples.

Isolation and characterization of specific anti-GAPDH Fab antibodies

Selection and ELISA screening. The performance of the library was evaluated by selecting for binders to baculovirus-produced *C.albicans* GAPDH antigen. Two selection strategies were employed. The first used biotinylated GAPDH antigen bound to streptavidin-coupled magnetic beads, and the second was performed on GAPDH-coated immunotubes. Three rounds of selection were then performed for each strategy. From rounds 2 and 3 of each selection campaign, 96 individual phage clones were screened for reactivity with the biotinylated and coated GAPDH by ELISA. Biotinylated BSA bound to streptavidin was used as a negative control in the ELISA. The frequencies of positive phage ELISA clones at round 3 were 2 and 51%, for the first and the second method, respectively.

The number of distinct Fabs was analyzed by DNA sequence determination of both light and heavy chains. Two distinct Fab sequences were found for the first method of selection, and three distinct sequences were found for the second method. Their CDR amino acid sequences are presented in Table 2 and are compared with human germ-line genes. The selected anti-GAPDH antibodies belong to families V_H 1, 3 and 5; two of them are related to V_H segment 1-02, one to V_H segment 3-23, one to V_H segment 3-30 and one to V_H segment 5-51. A comparison of mutations occurring at the amino acid sequence level is shown for clones B1 and C1 in Table 3.

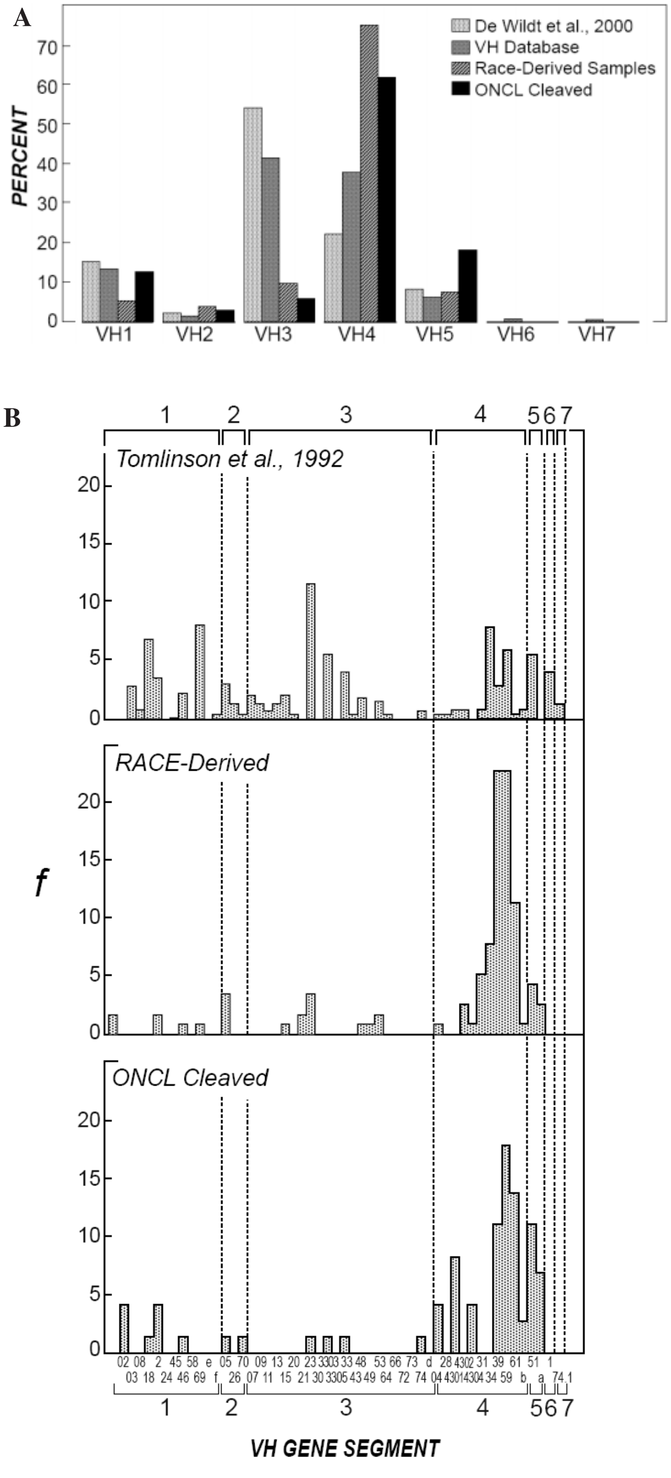

Figure 4. V_H family distribution analysis. (**A**) Frequency of V_H family use in healthy donor samples compared with Trimera mice samples (RACE-derived and ONCL-cleaved samples). V_H family distribution was compiled for the healthy donor samples from the 265 V_H described by de Wildt *et al.* (32), was studied in 1254 collected heavy chain sequences (V_H database) (R. C. Ladner, unpublished data) in 115 clones of RACE-derived V_H genes, and in 73 clones of the V_H repertoire from the immune library. (**B**) Use of human germline V-gene segments. Frequencies of use of human V_H segments was compiled from the 292 rearranged V_H genes in the database described by Tomlinson *et al.* (33), observed in 115 clones of Trimera mice-amplified and RACE-derived V_H genes, and in 73 clones of the V_H repertoire from the immune library. Frequencies (*f*) are plotted as % of total.

The two antibodies differ by only three amino acids, located in the heavy chain framework 1, CDR 2 and CDR 3.

Specificity analysis of the selected monoclonal Fabs. The specificity of candidate clones selected from the library was further confirmed by ELISA as both phage-displayed Fab product and soluble Fab product, directly collected from cell supernatant or FPLC-purified (Figure 5). To allow soluble Fab production, the positive phage antibodies were recloned as described in Materials and Methods. Soluble Fabs were then produced at high yields from *E.coli* periplasm. The average yield of soluble recombinant protein after metal chelate chromatography purification was ~100–400 μg purified Fab/L bacterial culture using shaker flasks. The specificity of the anti-GAPDH clones for binding to their target was then checked using ELISA (Figure 5). Two of the five specific phage antibodies (B1 and C1) were found reactive on GAPDH antigen as soluble Fab.

In summary, we have validated the novel ONCL technology first by the successful construction of an immune phage-display antibody library, then by the comparison of the IgG captured repertoires with IgG repertoires cloned by standard technologies. Moreover, we have shown that this library is functional through the isolation of five anti-GAPDH antibodies.

DISCUSSION

This report describes a novel approach to capture unbiased cDNA libraries, the ONCL method. Previously published methods for capturing Ig gene fragments have biases that ONCL avoids. We report an evaluation of ONCL through the construction of an immune phage-display library and selection of specific Fabs. We validated this new method by analyzing the antibody sequences of the library. We demonstrated that the ONCL cleavage of DNA in single-strand form is very effective, yielding a correct amplification product of the desired length, without any contaminating DNA populations. ONCL is the first technology that enables cleavage of linear ssDNA through the assistance of oligonucleotides and restriction enzymes to capture a gene repertoire where the gene family contains suitable restriction sites in all members.

ONCL combines numerous advantages compared with other protocols for antibody library construction currently in use. In this strategy, using the 5′ RACE as the first amplification step circumvents the biases introduced by specific PCR primers when located in the immunoglobulin 5′ variable region itself. Although antibody diversity resides mainly in the CDR regions, several reports have shown diversity within the framework regions, which may also be important for the binding to the target (18,19). Priming in a non-V-gene sequence potentially allows the capture of a wider repertoire. As expected, the antibody repertoire we captured was found to be diverse both in CDR and in framework regions. In the small sampling of RACE-derived samples that we examined, we saw the antibody repertoire was assembled from all the major V_H families (Figure 4B). The complexity of the repertoire might be affected by the 50–70 rounds of amplification, required to obtain enough material for the cleavage. In comparison, PCR-made phage-display libraries have been built after 60–80 PCR cycles (3,8–14) and diversity in those

Table 2. Deduced protein sequences of anti-GAPDH antibody fragments: CDR domains, V-gene family and germ-line (derivation)

Anti-GAPDH clone	LC-CDR1	LC-CDR2	LC-CDR3	Family/germline
R2 G5	RASQSISSYLN	AASSLQS	QQSYSTPRT	*K1/DPK9*
R2 C12	RSSQSLVYSDGNTYLN	KVSNRDS	MQGTHWPRT	*K2/DPK18*
R3 B1	SGSSSNIGTNYVY	RNNQRPS	AAWDDSLGGRV	*L1/DLP3*
R3 C1	SGSSSNIGTNYVY	RNNQRPS	AAWDDSLGGRV	*L1/DLP3*
R3 F5	RASQSISSSYLA	GASSRAT	QQYDSSSVT	*K3/DPK22*
	HC-CDR1	HC-CDR2	HC-CDR3	
R2 G5	DFGIT	RISPTDSYTMYSPSFQG	HRGAPYDYDSGGPYDYYGMDV	*VH5/5-51*
R2 C12	SYGMH	VTAHDGTNKYYADSVKG	VAGAYGENSWYFDV	*VH3/3-30*
R3 B1	GYYIY	WINPDSGGTNYAQKFHD	ATVSMTRGLFYYLED	*VH1/1-02*
R3 C1	GYYIY	WINPDSGGTNYARKFHD	ATVSMIRGLFYYLED	*VH1/1-02*
R3 F5	RYRMA	SIVPSGGRTFYADSVKG	NARRAFPSMDV	*VH3/3-23*

CDR, complementarity determining region.

Table 3. Deduced protein sequences of anti-GAPDH antibody fragments: comparison of heavy chain protein sequence from clones B1 and C1

	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4
	1 2 3	4	5	6	7 8 9		
	123456789012345678901234567890	1ab2345	67890123456789	012abc3456789012345	67890123456789012abc345678901234		
B1	EVQLVQSGAEVKKPGASVKVSCKSSGYTFT	G--YYIY	WVRQAPGQGLEWLG	WINP--DSGGTNYAQKFHD	RVTMTRDTSVSTAYLELSSLRSDDTAVYYCAR	ATVSMTRGLFYYLED	WGQ
C1	EVQLVQSGAEVKKPGASVRVSCKSSGYTFT	G--YYIY	WVRQAPGQGLEWLG	WINP--DSGGTNYARKFHD	RVTMTRDTSVSTAYLELSSLRSDDTAVYYCAR	ATVSMIRGLFYYLED	WGQ

Mutations are underlined. FR, framework region; and CDR, complementarity determining region.

libraries is well established. Since approximately the same number of PCR cycles is performed in the ONCL protocol, we expect the diversity to be at least equivalent. Furthermore, in conventional PCR-derived libraries, the use of V-gene-specific primers can affect the complexity of the repertoire by compromising the amplification of low expressed genes or genes that differ from germline in the priming region. In contrast, ONCL uses non-V-gene related primers in all amplification steps. This allows the maintenance of the initial gene proportions in the repertoire and therefore increases the chance of capturing rare mRNAs. As a main result, the probability of encompassing the entire antibody diversity is enhanced and complexity is very likely to be higher than in PCR-made libraries. This is important when analyzing RNA from heterogeneous B-cell populations. When we compared the heavy chain repertoire present in the library with the rearranged V_H segment gene database described by Tomlinson *et al.* (33), we observed that human V_H segments not frequently used were captured, such as 1-02, 3-21, 3-49, 4-30-4 and 5-a segments (Figure 4B). This usage of infrequently seen V_H gene segments could be driven by the *C.albicans* anti-GAPDH immune response, note that two of the five selected anti-GAPDH antibodies shared the 1-02 V_H segment.

Although the restriction enzymes used in ONCL may recognize only four bases, the cleavage has an effective specificity equal to the length of the adapter oligonucleotides. The annealing of the adapter can tolerate one or two base errors as long as the restriction site is maintained. The annealing and ligation of the bridge–extender can also tolerate one or more errors. Thus, the cleavage is precise while preserving diversity in the framework.

As with PCR technology, ONCL gives preference for complete 5′ regions because the oligonucleotide adapter will not anneal to truncated antibody gene fragments. This enables sequences that have deletions to be removed from the amplified DNA before displaying. Also, after RACE, all V-gene sequences that we analyzed showed full-length coding regions, including leader sequences. In contrast to PCR technology, this new technology offers the advantage of monitoring the cleavage efficiency and, therefore, the V-gene capture efficiency through the set of designed adaptors. Moreover, as ONCL requires DNA in a single-stranded format and its hybridization to appropriate oligonucleotide adapters, ONCL excels classical strategies that depend on PCR and restriction enzyme cleavage that would fail to succeed for the cloning of genes containing multiple and identical restriction enzyme sites.

A possible limitation of this novel cleavage approach is that the set of designed adapters might not anneal and capture all V genes. However, in our experiments, the heavy chain

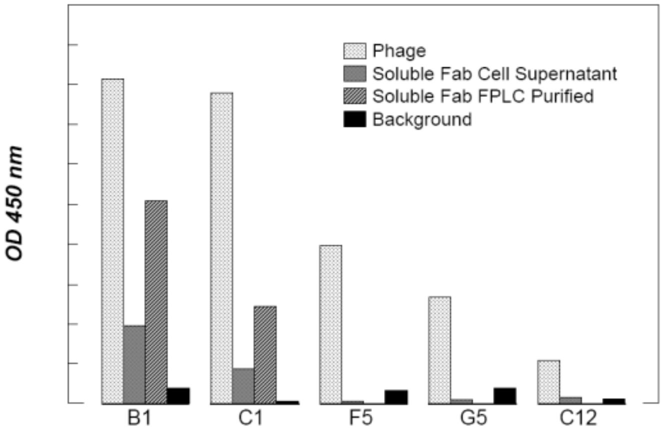

Figure 5. Specificity of selected anti-GAPDH antibody fragments determined by ELISA. Characterization of phage and soluble Fab anti-GAPDH antibodies by ELISA. The assay was performed by immobilizing biotinylated GAPDH on a polystyrene plate. Phage-displayed antibodies reactive with the coated antigen were detected with peroxidase-conjugated anti-M13 antibody (Amersham), while the detection of soluble Fabs was performed using 9E10 anti-c-myc monoclonal antibody (final concentration 2 μg/ml) followed by peroxidase-conjugated rabbit anti-mouse antibody (1:1000). The results of the assay are shown as absorbance at 450 nm. In the case of the data obtained with the phage preparation and the sFab cell supernatant, results were not normalized for protein concentrations.

sequencing data demonstrated the maintenance of almost all V_H segments when captured by the 5′ RACE after the ONCL cleavage (Figure 4B), confirming the correct design of the cleavage adapters. The slight differences observed between both repertoires are likely to be the result of a limited sequence sampling (Figure 4B). As two V_H germ-line gene sequences lack PvuII restriction sites (1-03 and 4-34), these V_H members are not likely to be captured. However, these two V_H genes are normally used at a very low frequency by the human repertoire, so the overall quality of the V_H repertoire will not be affected (32).

The ONCL cleavage is performed at high temperatures (>50°C) to maintain DNA in a single-stranded form. One might think that using these high temperatures could affect the enzymatic properties of type II restriction enzymes, for which lower temperature are recommended. However, we showed that efficiency of cleavage was maintained even for temperatures above 50°C.

The library was found to contain 70% of antibodies with a full read-through sequence and 30% with stop codons, mostly derived from incorrectly synthesized oligonucleotides, misannealing or from the PCR steps. When constructing libraries using standard PCR procedures, these types of mutations also arise. Likewise, the fact that we are using synthetic oligonucleotides in the bridge–extender ligation of the ONCL-cleaved V genes (Figure 2) also generated a few incorrect sequences that we found to be mainly due to the oligonucleotide synthesis itself. Nevertheless, these are compensated for by the construction of a very large size immune library of 3.3×10^{10} transformants.

We observed an over-representation of V_H family 4 and an under-representation of V_H family 3. The most frequent heavy chain germ-line gene V3-23 was almost absent in our samples, whereas three V_H4 family members 4-61, 4-39 and 4-59 were frequently observed (Figure 4B). Since the human PBLs used to engraft the Trimera mice were isolated from four patients, it is very unlikely that V3-23 would be under-represented and V_H4-61, 4-39 and 4-59 over-represented in all four donor samples, although we cannot fully exclude this possibility. Furthermore, one of the five selected anti-GAPDH antibodies uses the V_H3-23 segment, reinforcing our hypothesis that this bias is related to the use of the Trimera mouse system. Although human antibodies specific for several viral pathogens were successfully isolated using phage combinatorial libraries constructed from individuals with a positive serum titer against a given antigen (34), the use of the Trimera mouse system was guided by the additional idea to use it as a tool to boost the human IgG immune response against a specific antigen. ELISA data performed with stimulated mouse sera confirm this was effective. Moreover, two reports support the differences we observed in the V_H segment usage found in the IgG repertoire we isolated from GAPDH-immunized Trimera mice when compared with IgG repertoires derived either from normal or autoimmune individuals (32,35). Therefore, we think that this preferential V_H gene use is most likely due to B-cell triggering in the Trimera mouse system. However, we cannot exclude the possibility that the use of a novel antibody cloning technique (ONCL), which overcomes the variability in the B-cell repertoire, may have resulted in a more accurate representation of human V_H families as compared with previously reported technologies (32,33).

The human immune response, both in humans and in mice, results in antibodies that bind a rather limited number of protective epitopes (10), and immunization leads to the recognition only of a limited number of immunodominant epitopes. The diversity of immunoglobulin genes in an immune library is therefore expected to be limited compared with a non-immune library (IgM repertoire). In our case, the source of antibody-producing B cells was peripheral blood lymphocytes, which are mainly IgM-positive cells coming from *C.albicans*-infected patients. Although the frequency of IgG B cells activated against GAPDH was therefore expected to be rather low, the engrafted Trimera mice showed a clear IgG anti-GAPDH serum antibody titer, indicating the immune response directed to *C.albicans* GAPDH was enhanced in the Trimera mice. In a parallel work, the ONCL method has also been successfully used to build a non-immune library (36).

In addition, we have established the functionality of the ONCL library through the isolation of five specific anti-GAPDH antibodies. The two different selection strategies that we used gave different results, correlating with the findings of Lou *et al.* (37). The sequences of two of those five antibodies show significant identity (B1 and C1). This suggests that the B cells producing them were derived from the same pre-existing B-cell clone and indicates that the obtained anti-GAPDH antibodies are originating from a human immune (IgG) repertoire. However, although the proof-reading activity of Advantage 2 polymerase should reduce the mutation rate by 3-fold compared with a regular *Taq* polymerase, we cannot exclude mutations due to PCR amplification.

In summary, we have established a new cloning technology, the ONCL method, for making antibody libraries and validated it by isolating *C.albicans*-specific antibodies from immunized mice carrying human B-lymphocytes. The cornerstone of the technology, oligonucleotide-directed cleavage of ssDNA, is an alternative to PCR for the directional cloning of groups of related mRNAs, such as those of immunoglobulins and T-cell receptors. Besides the demonstrated application for the cloning of human antibody libraries as established here, we anticipate that the concept of oligonucleotide-assisted ssDNA cleavage will be applicable in cDNA cloning of repertoires of related genes, enzymatic restriction of DNA molecules containing multiple identical restriction sites, general strategies of mutagenesis, antibody affinity maturation (chain or CDR shuffling) and library design.

ACKNOWLEDGEMENTS

The authors thank K. Williams and F. Whelihan for helpful revision of the manuscript. The authors thank K. Rookey and H. G. Nastri for their contributions to the ONCL concept. The authors also thank all their colleagues at Dyax, in both Cambridge and Liege, for many contributions and discussions throughout the course of this work. Funding to pay the Open Access publication charges for this article was provided by Dyax.

Conflict of interest statement. None declared.

REFERENCES

1. Winter,G., Griffiths,A.D., Hawkins,R.E. and Hoogenboom,H.R. (1994) Making antibodies by phage-display technology. *Annu. Rev. Immunol.*, **12**, 433–455.
2. Hoogenboom,H.R. and Winter,G. (1992) By-passing immunisation. Human antibodies from synthetic repertoires of germline V_H gene segments rearranged *in vitro*. *J. Mol. Biol.*, **227**, 381–388.
3. Marks,J.D., Hoogenboom,H.R., Bonnert,T.P., McCafferty,J., Griffiths,A.D. and Winter,G. (1991) By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.*, **222**, 581–597.
4. Clackson,T., Hoogenboom,H.R., Griffiths,A.D. and Winter,G. (1991) Making antibody fragments using phage-display libraries. *Nature*, **352**, 624–628.
5. Hoogenboom,H.R., Griffiths,A.D., Johnson,K.S., Chiswell,D.J., Hudson,P. and Winter,G. (1991) Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res.*, **19**, 4133–4137.
6. Barbas,C.F.,III, Kang,A.S., Lerner,R.A. and Benkovic,S.J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc. Natl Acad. Sci. USA*, **88**, 7978–7982.
7. Griffiths,A.D., Malmqvist,M., Marks,J.D., Bye,J.M., Embleton,M.J., McCafferty,J., Baier,M., Holliger,K.P., Gorick,B.D., Hughes-Jones,N.C. *et al.* (1993) Human anti-self antibodies with high specificity from phage-display libraries. *EMBO J.*, **12**, 725–734.
8. de Haard,H.J., van Neer,N., Reurs,A., Hufton,S.E., Roovers,R.C., Henderikx,P., de Bruine,A.P., Arends,J.W. and Hoogenboom,H.R. (1999) A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J. Biol. Chem.*, **274**, 18218–18230.
9. Vaughan,T.J., Williams,A.J., Pritchard,K., Osbourn,J.K., Pope,A.R., Earnshaw,J.C., McCafferty,J., Hodits,R.A., Wilton,J. and Johnson,K.S. (1996) Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage-display library. *Nat. Biotechnol.*, **14**, 309–314.
10. Amersdorfer,P., Wong,C., Smith,T., Chen,S., Deshpande,S., Sheridan,R. and Marks,J.D. (2002) Genetic and immunological comparison of anti-botulinum type A antibodies from immune and non-immune human phage libraries. *Vaccine*, **20**, 1640–1648.
11. Barbas,C.F.,III, Collet,T.A., Amberg,W., Roben,P., Binley,J.M., Hoekstra,D., Cababa,D., Jones,T.M., Williamson,R.A., Pilkington,G.R. *et al.* (1993) Molecular profile of an antibody response to HIV-1 as probed by combinatorial libraries. *J. Mol. Biol.*, **230**, 812–823.
12. Sioud,M. and Hansen,M.H. (2001) Profiling the immune response in patients with breast cancer by phage- displayed cDNA libraries. *Eur. J. Immunol.*, **31**, 716–725.
13. Hanes,J., Jermutus,L., Weber-Bornhauser,S., Bosshard,H.R. and Pluckthun,A. (1998) Ribosome display efficiently selects and evolves high-affinity antibodies *in vitro* from immune libraries. *Proc. Natl Acad. Sci. USA*, **95**, 14130–14135.
14. Lerner,R.A., Barbas,C.F.,III, Kang,A.S. and Burton,D.R. (1991) On the use of combinatorial antibody libraries to clone the 'fossil record' of an individual's immune response. *Proc. Natl Acad. Sci. USA*, **88**, 9705–9706.
15. Mullis,K., Faloona,F., Scharf,S., Saiki,R., Horn,G. and Erlich,H. (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb Symp. Quant. Biol.*, **51**, 263–273.
16. Saiki,R.K., Scharf,S., Faloona,F., Mullis,K.B., Horn,G.T., Erlich,H.A. and Arnheim,N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350–1354.
17. Larrick,J.W., Danielsson,L., Brenner,C.A., Abrahamson,M., Fry,K.E. and Borrebaeck,C.A. (1989) Rapid cloning of rearranged immunoglobulin genes from human hybridoma cells using mixed primers and the polymerase chain reaction. *Biochem. Biophys. Res. Commun.*, **160**, 1250–1256.
18. de Haard,H.J., Kazemier,B., van der Bent,A., Oudshoorn,P., Boender,P., van Gemen,B., Arends,J.W. and Hoogenboom,H.R. (1998) Absolute conservation of residue 6 of immunoglobulin heavy chain variable regions of class IIA is required for correct folding. *Protein Eng.*, **11**, 1267–1276.
19. Knappik,A. and Pluckthun,A. (1995) Engineered turns of a recombinant antibody improve its *in vivo* folding. *Protein Eng.*, **8**, 81–89.
20. Frohman,M.A., Dush,M.K. and Martin,G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA*, **85**, 8998–9002.
21. Loh,E.Y., Elliott,J.F., Cwirla,S., Lanier,L.L. and Davis,M.M. (1989) Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science*, **243**, 217–220.
22. Akowitz,A. and Manuelidis,L. (1989) A novel cDNA/PCR strategy for efficient cloning of small amounts of undefined RNA. *Gene*, **81**, 295–306.
23. Maleszka,R. and Stange,G. (1997) Molecular cloning, by a novel approach, of a cDNA encoding a putative olfactory protein in the labial palps of the moth *Cactoblastis cactorum*. *Gene*, **202**, 39–43.
24. Chenchik,A., Diachenko,L., Moqadam,F., Tarabykin,V., Lukyanov,S. and Siebert,P.D. (1996) Full-length cDNA cloning and determination of mRNA 5′ and 3′ ends by amplification of adaptor-ligated cDNA. *Biotechniques*, **21**, 526–534.
25. Bradbury,A., Persic,L., Werge,T. and Cattaneo,A. (1993) Use of living columns to select specific phage antibodies. *Biotechnology (NY)*, **11**, 1565–1569.
26. Schramm,G., Bruchhaus,I. and Roeder,T. (2000) A simple and reliable 5′-RACE approach. *Nucleic Acids Res.*, **28**, E96.
27. Roeder,T. (1998) Solid-phase cDNA library construction, a versatile approach. *Nucleic Acids Res.*, **26**, 3451–3452.
28. Eren,R., Lubin,I., Terkieltaub,D., Ben-Moshe,O., Zauberman,A., Uhlmann,R., Tzahor,T., Moss,S., Ilan,E., Shouval,D. *et al.* (1998) Human monoclonal antibodies specific to hepatitis B virus generated in a human/mouse radiation chimera: the Trimera system. *Immunology*, **93**, 154–161.
29. Reisner,Y. and Dagan,S. (1998) The Trimera mouse: generating human monoclonal antibodies and an animal model for human diseases. *Trends Biotechnol.*, **16**, 242–246.
30. Ilan,E., Eren,R., Lubin,I., Nussbaum,O., Zauberman,A. and Dagan,S. (2002) The Trimera mouse: a system for generating human monoclonal antibodies and modeling human diseases. *Curr. Opin. Mol. Ther.*, **4**, 102–109.
31. Villamon,E., Gozalbo,D., Martinez,J.P. and Gil,M.L. (1999) Purification of a biologically active recombinant glyceraldehyde 3-phosphate dehydrogenase from *Candida albicans*. *FEMS Microbiol. Lett.*, **179**, 61–65.
32. de Wildt,R.M., Tomlinson,I.M., van Venrooij,W.J., Winter,G. and Hoet,R.M. (2000) Comparable heavy and light chain pairings in normal and systemic lupus erythematosus IgG(+) B cells. *Eur. J. Immunol.*, **30**, 254–261.

33. Tomlinson,I.M., Walter,G., Marks,J.D., Llewelyn,M.B. and Winter,G. (1992) The repertoire of human germline V_H sequences reveals about fifty groups of V_H segments with different hypervariable loops. *J. Mol. Biol.*, **227**, 776–798.
34. Williamson,R.A., Burioni,R., Sanna,P.P., Partridge,L.J., Barbas,C.F.,III and Burton,D.R. (1993) Human monoclonal antibodies against a plethora of viral pathogens from single combinatorial libraries. *Proc. Natl Acad. Sci. USA*, **90**, 4141–4145.
35. Griffiths,A.D., Williams,S.C., Hartley,O., Tomlinson,I.M., Waterhouse,P., Crosby,W.L., Kontermann,R.E., Jones,P.T., Low,N.M., Allison,T.J. *et al.* (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires. *Embo J*, **13**, 3245–3260.
36. Hoet,R.M., Cohen,E.H., Kent,R.B., Rookey,K., Schoonbroodt,S., Hogan,S., Rem,L., Frans,N., Daukandt,M., Pieters,H. *et al.* (2005) Generation of high-affinity human antibodies by combining donor-derived and synthetic complementarity-determining-region diversity. *Nat. Biotechnol.*, **23**, 344–348.
37. Lou,J., Marzari,R., Verzillo,V., Ferrero,F., Pak,D., Sheng,M., Yang,C., Sblattero,D. and Bradbury,A. (2001) Antibodies in haystacks: how selection strategy influences the outcome of selection from molecular diversity libraries. *J. Immunol. Methods*, **253**, 233–242.