Hybridization-based antibody cDNA recovery for the production of recombinant antibodies identified by repertoire sequencing

Javier Valdés-Alemán1,2,†, Juan Téllez-Sosa1, Marbella Ovila-Muñoz1, Elizabeth Godoy-Lozano1, Daniel Velázquez-Ramírez1,4, Humberto Valdovinos-Torres1, Rosa E Gómez-Barreto1, and Jesús Martínez-Barnetche1,*

1Departamento de Inmunología; Centro de Investigación Sobre Enfermedades Infecciosas; Instituto Nacional de Salud Pública; Cuernavaca, México; 2Programa de Ingeniería en Biotecnología; Universidad Politecnica del Estado de Morelos; Jiutepec, México

Current affiliation: †Janelia Farm Research Campus; Howard Hughes Medical Institute; Ashburn, VA USA; 2Universidad del Istmo; Juchitán de Zaragoza; Oaxaca, México

Keywords: high-throughput sequencing, repertoire, CDRH3, data mining, enrichment

Abbreviations: HEL, hen egg lysozyme; CDRH3, heavy chain complementarity determining region 3; FBS, fetal bovine serum

High-throughput sequencing of the antibody repertoire is enabling a thorough analysis of B cell diversity and clonal selection, which may improve the novel antibody discovery process. Theoretically, an adequate bioinformatic analysis could allow identification of candidate antigen-specific antibodies, requiring their recombinant production for experimental validation of their specificity. Gene synthesis is commonly used for the generation of recombinant antibodies identified in silico. Novel strategies that bypass gene synthesis could offer more accessible antibody identification and validation alternatives. We developed a hybridization-based recovery strategy that targets the complementarity-determining region 3 (CDRH3) for the enrichment of cDNA of candidate antigen-specific antibody sequences. Ten clonal groups of interest were identified through bioinformatic analysis of the heavy chain antibody repertoire of mice immunized with hen egg white lysozyme (HEL). cDNA from eight of the targeted clonal groups was recovered efficiently, leading to the generation of recombinant antibodies. One representative heavy chain sequence from each clonal group recovered was paired with previously reported anti-HEL light chains to generate full antibodies, later tested for HEL-binding capacity. The recovery process proposed represents a simple and scalable molecular strategy that could enhance antibody identification and specificity assessment, enabling a more cost-efficient generation of recombinant antibodies.

Introduction

The adaptive immune system is capable of generating a vast antibody response against a great variety of possible pathogens through clonal selection. The antibody repertoire is dynamically shaped in response to external stimuli. Natural infection or immunization can evoke a strong polarization of the repertoire by selecting high affinity antigen-specific clonal groups. The high-throughput sequencing of the antibody transcripts has enabled a more thorough analysis of the immunoglobulin diversity1,2 as a means to describe the complexity and dynamics of the antibody-mediated defense process, which has led to a better understanding of the adaptive immune response3,5 and an improved antibody discovery process.6,7

Used as tools for diagnosis, therapy and research, recombinant antibodies are indispensable in the biomedical field. Antibodies can be generated against different pathogens and many other molecules of interest, but the production of new recombinant antibodies against specific antigens is directly limited by the discovery strategy used to identify such antibodies. Traditionally, approaches such as hybridomas8 and phage display have been implemented to address this issue.9 More recently, techniques such as single B cell expression cloning, B cell immortalization10-13 and deep sequencing of the antibody repertoire are improving the analysis of the human antibody response. The analysis of the antibody repertoire through high-throughput sequencing has been used to study the diversity of the immune response, and for the identification of antibodies of interest.6,14-16 With an adequate bioinformatic analysis, it is possible to identify putative antigen-specific antibodies from the repertoire sequences based on certain parameters such as clonal frequency6 or structural relations with highly mutated antibodies of known specificity, such as broadly neutralizing antibodies against HIV.17 The production of recombinant antibodies based on the bioinformatic analysis of the repertoire relies mainly on the synthesis of DNA of specific light and heavy variable regions; however, the cost of the synthesis of DNA is a direct constraint affecting the evaluation throughput.6

*Correspondence to: Jesús Martínez-Barnetche; Email: jmbarnet@insp.mx
Submitted: 10/08/2013; Revised: 12/01/2013; Accepted: 12/04/2013
http://dx.doi.org/10.4161/mabs.27435
The CDRH3 is the region with the highest diversity within the heavy chain variable region (VH) and a major determinant of antigen specificity. Selecting the CDRH3 is an effective way to molecularly recover single clonal groups. In this study, we provide a proof-of-principle of targeting the CDRH3 by using restrictive PCR along with a hybridization-based recovery, allowing clonal groups with different relative abundances to be enriched and produced recombinantly without the need of chemical synthesis of the whole gene.

## Results

### Evidence of selection through enrichment

An immunoglobulin sequencing experiment was performed by immunizing two BALB-c mice with HEL at days 0 and 10. Spleens were collected at day 15, which corresponds to the peak of germinal center formation, somatic hypermutation and class switching. Spleen total RNA was used to generate IGHV amplicons of the IgG compartment followed by high throughput sequencing with the 454-Roche platform. A total of 45,996 and 36,512 reads were obtained from library A (LA) and B (LB), respectively. By clustering sequences according to VDJ rearrangement, CDRH3 length and identity, 783 (LA) and 902 (LB) clonal groups were identified (see Materials and Methods: Bioinformatics analysis and candidate selection).

To simplify the bioinformatic analysis, only the frequency of each clonal group was considered in the selection; the five most frequent from each library were selected for recovery (2G1–2G5, 3G1–3G5). The same amplicon libraries used for pyrosequencing were also used as templates for the amplification of the rearrangements found in the clonal groups of interest. Structural features of the five most frequent clonal groups in each mouse and the CDRH3 used for primers design are described in Table 1.

The whole enrichment process comprises two main selection points (Fig. 1). The first selection consisted of the PCR amplification of the desired rearrangements using IGHV and IGJH family-specific primers, for which stringent PCR

### Table 1. Primer sequences and V/J usage

| Clonal group | Probe sequence (5′-3′) | IGHV and IGJH usage‡ |
|--------------|------------------------|----------------------|
| 2G1          | /B/CAAAGTA CCCCGTGTAA CCAACATT | IGHV14-4*02 IGJH2*01 |
| 2G2          | /B/GTAAGTCG TGGAAGCAA TCCCCCAT | IGHV14-4*02 IGJH2*01 |
| 2G3          | /B/CCAGGAG TAATGAAACAT TACAG | IGHV14-4*02 IGJH3*01 |
| 2G4          | /B/AAACCCCTC CGTACCGGG CTTT | IGHV1–18*01 IGJH3*01 |
| 2G5          | /B/TAACGGGG CCGTATTACC TCCT | IGHV1–63*02 IGJH1*01 |
| 3G1          | /B/TACCCGAT GTTACCCCTC TCCTTTC | IGHV1–9*01 IGJH2*01 |
| 3G2          | /B/CTTCTGAT AATCATACCC ATCCCTTCTC | IGHV4–1*02 IGJH3*01 |
| 3G3          | /B/GTAAGGA GTCAATAC TAGTCG | IGHV1–85*01 IGJH3*01 |
| 3G4          | /B/AAGGAGG TCCCTCTGA TAAC | IGHV14–4*02 IGJH3*01 |
| 3G5          | /B/AACTACC GTGTAATCT CCCCTCTTG | IGHV1–76*01 IGJH2*01 |

/B/: Biotin molecule at 5′ end; bold sequence: sequence within the IGJH segment; underlined sequence: sequence within the IGHV segment. ‡ Determined by IMGT/V-QUEST.
conditions were used to minimize the amplification of undesired but structurally similar rearrangements. Ideally, only sequences with the same IGHV-IGHJ usage to that of the target group would reach this step. The PCR products generated at this point served as the input material for the second selection in which a biotinylated CDRH3-specific oligonucleotide was used for hybridization-based enrichment. After this step, most of the recovered sequences should have the desired rearrangement and CDRH3.

As a preliminary assessment of the enrichment process, a PCR checkpoint was performed immediately after the second selection (Fig. 2). The enrichment products were used as templates in PCR with the IGHV-IGHJ and IGHV-CDRH3 primers pairs, comparing the enriched samples and the negative controls (mock enrichments). As this system enables the enrichment of the desired sequences instead of a complete isolation of them, undesired rearrangements were expected to inevitably yield an amplification product for the mock enrichments. Therefore, it was important to determine the number of amplification cycles at which the unwanted sequences (background seen in the mock enrichments) reach a significant amount to minimize the abundance of irrelevant sequences (Fig. 2A). Additionally, the amplification pattern observed when using IGHV-IGHJ and IGHV-CDRH3 primers pairs should be similar, since it is expected that the majority of the sequences amplified by the IGHV-IGHJ pair are the ones amplified by the IGHV-CDRH3 pair. Nevertheless, a notable difference in the abundance of PCR product between the enriched sample and mock enrichment was observed for all clonal groups (Fig. 2B). This amplification difference is the evidence of enrichment of the desired molecules.

Validation of the enrichment of desired clonal groups by sequence analysis

Once the proper number of amplification cycles was determined for each group, the IGHV-IGHJ amplicon generated under such conditions was cloned into the expression vector. Only those clones with a CDRH3-specific insert detected through colony PCR were Sanger-sequenced. Assuming that those sequences amplified with their CDRH3 primer were also subject to enrichment in the recovery process, the positive to negative CDRH3-specific colony PCR ratio was used as an indicator of the enrichment efficiency (Fig. 3A).

This assumption, however, could consider CDRH3 primer mismatch to highly similar but not identical sequences. Therefore, 6 to 15 colonies from each of the eight clonal groups were sequenced and aligned against their corresponding target sequence. In eight of the ten groups, sequences with a correct rearrangement and identical CDRH3 were recovered (Fig. 3B), with an overall sequence identity of ≥98.6% ± 1.1. Partially complementary sequences with one or two mismatches in the CDRH3 were also recovered for these eight groups. Sequences with three or more mismatches in the CDRH3 were observed only in four of the groups. Nevertheless, some of these mismatches were synonymous, coding for the correct amino acid (Fig. S1).

Recovery of groups 3G2 and 3G5 was unsuccessful; only partially complementary sequences or wrong rearrangements were obtained.

Expression of recovered antigen-specific recombinant antibodies

One sequence for each of the eight clonal groups recovered was selected to be recombinantly expressed in HEK 293T cells. Although the recovery process considered only VH sequences, interchangeable murine light-chain variable regions (VL) were used to generate complete recombinant chimeric IgG.18,19 Because we did not have VL massive sequencing results readily available, the VH recovered were paired with four different VL (HH10, HH22, HH25, HH27A).
HH5, D1, D20 and F10.6.6 that had been previously reported to belong to HEL-binding antibodies. A total of 32 monoclonal antibodies (plus controls) were generated and evaluated by ELISA to assess expression and antigen-binding capacity. One irrelevant VH (from an unimmunized mouse) was paired with the four VL as negative controls to discard any evident specificity contribution by the VL chosen (data not shown).

The VH/VL combinations 2G3/HH10, 3G1/HH5 and 3G1/F10.6.6 were positive to the HEL-binding ELISA (Fig. 3C). In this case, the 3G1 VH showed binding capacity when paired separately with two different VL. Therefore, two of the eight VH evaluated showed binding capacity to HEL when paired to some of the VL chosen. The antigen-binding capacity of the antibodies obtained, however, does not rely on the enrichment process.
but depends directly on an effective bioinformatic analysis and immunization protocol.

**Discussion**

Massive sequencing provides a broad approach to study the repertoire dynamics in an immune response to an antigen, as well as the identification of antigen-specific sequences. The in silico candidate selection of the target sequences is commonly followed by synthetic production of the desired genes. The cost of gene synthesis, however, limits their experimental specificity validation and, in some cases, may restrict the selection by relying on consensus sequences. In this study, we propose a strategy for the recovery of heavy chain variable region cDNA to enrich the sequences of clonal groups of interest and produce them recombinantly, aiding experimental validation for the in silico analysis without the use of synthetic genes.

Two main alternative strategies, one mechanical and the other molecular, have been reported that bypass DNA synthesis when producing recombinant antibodies from digital repertoires. In the first case, a micromanipulator was used to retrieve specific sequencing beads coated with the desired DNA used in the 454 sequencing platform. This strategy requires a sophisticated manipulating device, and can only be used for the 454 sequencing system.

In the second strategy, a molecular recovery of the specific heavy chain antibody sequences was implemented to target the CDRH3 through an overlap PCR. However, the use of a CDRH3 primer in PCR can artificially add the desired sequence to a slightly different CDRH3 with which the primer can still hybridize and extend. Because the IGHV-CDRH3 and CDRH3-IGHJ fractions are first amplified and later joined, it is possible for two fractions that do not belong to each other to be joined, generating artificial sequences. This particular detail did not represent a problem because this strategy was tested with samples in which diversity was restricted only to the CDRH3. The recovery strategy proposed in our study is not solely based in PCR and is capable of retrieving complete biological variable region sequences. Theoretically, its application could be extended to other massive sequencing platforms.

The recovery process proposed in this study consists of two main selection points, one for the selection of the desired rearrangement through PCR, and the other for the selection of CDRH3 through a hybridization-based enrichment. These two filters proved to be sufficient to recover eight of ten expanded clonal groups from two diverse samples. The success in the first selection point relies on the design of the primers and the amplification conditions. Although the set of primers used in this study was intended to select specific murine IGHV and IGHJ families, amplification was not entirely restricted to them because the primers sequence could also be completely or partially shared with other families. There are several primers set proposed for the selective amplification of murine IGHV and IGHJ subfamilies, however most do not cover the whole variable region, from the first to the last codon. Despite the fact that the primers set used in this study lacked strict subfamily selectivity, such primer restriction was proven not to be essential. Nevertheless, the design of an extended and more selective set of primers would improve the overall recovery outcome. A similar approach can be done to amplify human VH sequences to broaden the extent of this strategy through the design of a new set of primers.

The hybridization-based CDRH3 selection represents the main clonal selection step at which a CDRH3-specific probe was used to recover each group. Biotinylated oligonucleotides have been previously used for the enrichment of specific transcripts or other target molecules within a mixture, and for capturing CDRs to exploit sequence diversity for combinatorial biology approaches. However, as far as we know, there are no reports of a targeted recovery of cDNA of specific antibody variable regions with this strategy, possibly due to the challenge it represents.

Since the CDRH3 sequence serves as a tag mark to identify clonal groups, a unique probe had to be designed for each group. The coverage of the probe was limited mainly to the IGHD and junctions (N and P nucleotides) because this region is the most diverse and distinctive for each clonal group within the repertoire. Inclusion of a high number of positions into the IGHV and IGHJ segments, resulting in the use of large probes, yields partially complementary sequences and hinders the recovery process (data not shown). Therefore, clonal groups with short and low diversity CDRH3 may be more difficult to recover.

The CDRH3-specific PCR after the VH cloning served as an indicator of the enrichment efficiency. The initial representation of the clonal groups before the enrichment started at 2.1% and up to 9.5% in the filtered digital repertoire. Thus, the actual representation is even lower, considering the sequences that were taken out in the in silico quality filters. At the CDRH3-specific colony PCR, group 2G5 was enriched up to a 100% of the colonies evaluated, and the lowest was 31% for 2G4. The abundance achieved after the enrichment varied among the eight clonal groups recovered. However, sequences with an identical CDRH3 were recovered in each case by testing a few colonies.

The presence of an amplification product at the CDRH3-specific colony PCR (checkpoint) did not ensure an identical CDRH3. Therefore, the clones were sequenced, showing that identical and nearly identical CDRH3 were recovered. The sequence of group 2G3 that was used in the transfections contained one non-synonymous mismatch (W→L) at the CDRH3 and still showed antigen-binding capacity. This indicates that sequences with one or two mismatches at the CDRH3 should still be considered for further evaluation, unless an amino acid known to be essential is lost.

We managed to recover sequences with a correct rearrangement and identical CDRH3 to that of the target molecule, whereas further variability was found along the rest of the VH region. This particular variety, presumably within the same clonal group, could be especially convenient when comparing different idiotypes with a common clonal origin. Through this approach one could identify specific positions in the VH of matured antibodies that confer them higher affinity when compared with their precursors.

While the discovery of HEL-binding antibodies from the enriched VH sequences does not provide any information
regarding the recovery success, it does indicate if an appropriate bioinformatic selection was performed. The combinations 2G3/HH10, 3G1/HH5 and 3G1/F10.6.6 were positive to the HEL-binding ELISA. The fact that the same VH paired with two different VL retained antigen-binding capacity reaffirmed that generic VL may be used for screening, as it has been previously reported.\textsuperscript{18,19,33} However, a more diverse and better selected pool of VL could allow high-throughput screening. Possibly, some of the negative VH recovered sequences could show antigen-binding capacity if paired with other VL sequences. To overcome VH-VL mispairing, massive sequencing of native VH-VL pairs could also be incorporated.\textsuperscript{34} Considering that the bioinformatic analysis was scarce, that the identification was solely based on frequency and no other parameters (e.g., hypermutation rate, CDRH3 composition and framework/CDR mutation ratio), and that non-native light chains were used, the observed outcome was better than expected.

The antibody repertoire is a very diverse and complex mixture of closely related sequences. Hence, an efficient isolation of a single sequence from such a complex mixture is not straightforward. The method proposed in this study allowed the recovery of specific clonal groups with simple molecular techniques. This strategy can yet be further improved and expanded, even to similar immune diverse samples, such as the T cell receptor repertoire.\textsuperscript{35} Coupled with improved bioinformatics methods to identify putative antigen-specific antibodies, this recovery method may also provide the proof of concept for scalable and automated devices for high-throughput antibody production.

In summary, ten presumably antigen-specific clonal groups were identified in silico from a VH antibody repertoire. Eight of these clonal groups were recovered through a PCR and hybridization-based strategy that enabled their recombinant expression. Sequences with the correct rearrangement and identical CDRH3 were recovered in each case. This alternative recovery strategy allows the enrichment of specific clonal groups without the need of synthetic DNA through the use of everyday molecular biology techniques.

\section*{Materials and Methods}

\subsection*{Mice immunization}

Adult (6–8 wk old) male BALB-c mice were immunized twice intraperitoneally with HEL (Cat. L6876, Sigma). The first immunization was performed with 30 \( \mu \)g of HEL with incomplete Freund’s adjuvant. Ten days later, a second immunization was performed with 5 \( \mu \)g of HEL with the same adjuvant. The spleen was removed 15 d after the first immunization and frozen in liquid nitrogen until further processing.

\subsection*{Libraries preparation and massive sequencing}

Total RNA was extracted with Trizol from the immunized mice spleens. Splenic RNA (1.5 \( \mu \)g) was used to generate VH amplicons by a 5' RACE-PCR approach. Briefly, total RNA was used to generate cDNA with a d(T)30 primer. In the 3' end of the cDNA, a universal primer containing the adaptor A sequence (454-Roche) was incorporated using the protocol described by Matz et al.\textsuperscript{36} IgG VH amplicons were generated by PCR with IgG-specific barcoded primers fused with adaptor B sequence (454-Roche) to allow discrimination of sequencing reads according to each mouse (5'-AdaptorB/4mer-454key/6mer-barcode/ CAGGGG CCGTGGATA GACHGATGG-3'). Amplicons were quantified with the Agilent Bioanalyzer and pooled in a 1:1 molar ratio. Amplicons were sequenced with the GS FLX Titanium platform (454-Roche) in a 2-region gasket pooled with other barcoded libraries following the standard amplicon sequencing protocol. To have good quality coverage and depth in the CDRH3 region, sequencing was performed from IGHC to IGHV direction using primer B.

\subsection*{Bioinformatic analysis and candidate selection}

The default signal processing pipeline for amplicon was used. The resulting SFF files were filtered using PRINTSEQ\textsuperscript{37} to filter out reads > 200 bp or with average Q values > 30, and were split according library barcode. We have developed a pipeline for automated analysis of clonal diversity in Ig Rep-Seq data (data not shown). To define clonal groups, IGHV and IGHJ segments for each read are mapped with IgBLAST,\textsuperscript{38} and the CDRH3 sequence is extracted with HMMER (http://hmmer.janelia.org/). For all the reads sharing the same VDJ rearrangement, the algorithm recursively clusters CDRH3 having \( \geq 98\% \) identical length and \( \geq 98\% \) sequence identity with USEARCH.\textsuperscript{39} The result is the consensus of each clonal group in FASTA format, complemented with the number of reads belonging to each clonal group that is used as proxy of clonal frequency. Since the libraries are derived from mRNA, clonal groups having several reads could be derived from clonal expansion of few plasma cells. Higher frequency clonal groups from each mouse were selected as putative anti-HEL specific clonal groups as described previously.\textsuperscript{6}

\subsection*{CDRH3 probes design}

One CDRH3 probe was designed for each of the ten clonal groups intended to be recovered. They were designed to cover mainly the CDRH3 region and a few nucleotides from the end of IGHV and the beginning of IGHJ segments, according to the IMGT/V-QUEST annotation.\textsuperscript{40} Ideally, no more than five positions outside the CDRH3 were included in each side, unless necessary to avoid significant secondary structures. These probes contained the complementary sequence of the CDRH3 so they could also be used in PCR as reverse primers along with a forward primer at an upstream position. Each probe was synthesized with a biotin molecule at its 5' end (IDT).

\subsection*{Rearrangement-specific PCR}

The reactions were performed with the following final concentrations: 1.6 fm VH PCR product as template, 0.5 mM IGHV forward primer, 0.5 mM IGHJ reverse primer, 1X PCR buffer, 0.2 mM dNTPs, 2.5 mM MgSO\textsubscript{4} and 0.025 U/\( \mu \)L Platinum Taq DNA polymerase High Fidelity (Invitrogen). Three forward 5'-phosphorylated primers for the amplification of murine IGHV families IGHV1 (5'-AATTCAGAGG TCCAGCTGCA GCAGTCCTGA-3'), IGHV4 (5'-AATTCAGAGG TGAAGCTTCT CGAG-3'), IGHV14 (5'-AATTCAGAGG TTCAAGCTGCA GCAGTCCTGG-3') and IGHV15 (5'-AATTCAGAGG AGACGGTGAC CG-3') were used separately. Three reverse primers were used for the amplification of the IGHJ families: IGHJ1 (5'-CTAGCTGAGG AGACGGTGAC CG-3'), IGHJ2 (5'-CTAGCTGAGG AGACGGTGAC AGTGGTG-3') and IGHJ3 (5'-CTAGCTGAGG AGACGGTGAC TGGTG-3').
and IGHJ3 (5′-CTAGCTGCAG AGACAGTGAC CAGA-3′). These primers contained the partial recognition sites for EcoRI (forward) or NheI (reverse), and a phosphate at the 5′ end to allow cohesive-end cloning through Exonuclease III digestion (Sigma). Stringent amplification conditions were used through a touchdown PCR scheme. Amplification started with 95 °C for 5 min; 35 cycles of melting at 95 °C for 30 s, annealing at X °C for 30 s, and extension at 72 °C for 30 s; and a final extension of 72 °C for 5 min. The annealing temperature (X) decreased 2 °C (X-2 °C) at cycle six, 1 °C (X-3 °C) at cycle 11 and 2 °C (X-5 °C) at cycle 16. The initial annealing temperature (X) for the primer pairs IGHV1-IGHJ1, IGHV1-IGHJ2, IGHV1-IGHJ3, IGHV14-IGHJ2 was 81 °C; for IGHV4-IGHJ3 was 75 °C; for IGHV1-IGHJ3 was 71 °C; and for IGHV3-IGHJ3 was 64 °C. The PCR products were gel-purified before any further processing.

Clonal group cDNA enrichment

Streptavidin-coated magnetic beads were used to recover the hybridized target-molecule/biotinylated-probe complexes. Five µL of magnetic beads (Enrichment beads, 454-Roche sequencing system) per sample were pre-washed twice with 200 µL of enhancing buffer (454-Roche sequencing system) every time, retrieving the beads with a magnetic base (Cat. S1506S, New England Biolabs), and finally resuspended in 10 µL of enhancing buffer per sample. For every clonal group intended to recover, 7.5 ng of the corresponding rearrangement-specific PCR product were mixed with annealing buffer (454-Roche sequencing system) up to 26 µL of magnetic beads were added to every sample and incubated at 95 °C for 5 min, followed by immediate cooling in ice for 2 min. Four µL of the CDRH3-specific biotinylated probe at 5 pmol/µL were added, proceeding to an incubation 5 °C above the corresponding Tm for 8 min and cooling in ice for 2 min. Ten µL of pre-washed magnetic beads were added to every sample and incubated in gentle rotation at room temperature for 7 min. The buffer was removed from the magnetic beads in a magnetic base once the solution had completely cleared (≥90 s). The remaining beads were washed ten times with 500 µL of enhancing buffer. Thirty µL of previously warmed water at 95 °C were added to the beads and were incubated at 95 °C for 5 min. The beads were immediately transferred to a magnetic base and the enrichment product solution was recovered once it had completely cleared (90 s). Along with every enriched sample, a mock enrichment (negative control) was performed, for which the biotinylated probe was replaced with annealing buffer.

Enrichment testing and VH cloning

The enriched VH sequences were cloned into the expression vector pVAGO-CHG1, derived from pFUSEss-CHIg-hG1 (InvivoGen), which contains the human IgG1 constant region and an omega-cassette that confers resistance to chloramphenicol. To generate the inserts, the VH sequences were PCR-amplified from the enrichment products using the corresponding IGHV and IGHJ primers. The amount of template used represented 10% of the final volume of the reaction; however, various separate reactions were done to cover increasing amplification cycles, starting at 13 and up to 20. The amplification started at 95 °C for 5 min; 13 to 20 cycles of 95 °C for 30 s, X-5 °C for 30 s, 72 °C for 30 s; and a final extension of 72 °C for 5 min. The value of X and chemical composition of the reaction were as mentioned in the rearrangement-specific PCR methods. Along with every sample, the negative controls of the enrichment were also tested by PCR and analyzed by electrophoresis in a 1% agarose gel. The amplification products generated from an enriched sample were visually compared with the ones generated from the corresponding mock enrichments. The final number of cycles elected was the highest one at which a notable product was generated from the enriched sample, but none or very scarce was obtained from the mock enrichment. Once the final number of amplification cycles was defined for every sample, the reactions were repeated only for the chosen cycles and using the Direct Universal Cloning System (Sigma). The new PCR products were digested with Exonuclease III, and the pVAGO-CHG1 with EcoRI and NheI (Fermentas). Digestion with Exonuclease III and the cloning procedures were as indicated by the manufacturer. The introduction of recombinant plasmids in Escherichia coli strain TOP10 was performed by electroporation. The TOP10 strain was grown at 37 °C in Luria-Bertani medium and, when required, chloramphenicol was added to a final concentration of 30 µg/mL.

Transformants were analyzed by colony PCR for inserts using in-vector primers, and for the specific CDRH3 by using the IGHV and CDRH3 (biotinylated) primers. Only those colonies positive for both PCRs were Sanger-sequenced. The clone with the closest sequence to its target was used for transfections. The recovered VH sequences were paired with four VL sequences: HH10 [GenBank: M356671], D1.3 [PDB: 1FDL_L], HH5 [PDB: 1BQL_L] and F10.6.6 [PDB: 2Q76_A], previously reported as positive for HEL binding. The VL sequences were synthesized (gBlocks, IDT) and cloned into the expression vector pVAGO-CLhk, derived from pFUSE2ss-CLIg-hk (InvivoGen), which contains the human kappa constant region and an omega-cassette that confers resistance to chloramphenicol.

Recombinant antibody expression

HEK 293T cells (Thermo Scientific) were plated at a density of 125 000 cells per well in 24-well plates with DMEM complete medium supplemented with 100 U/mL penicillin, 100 µg/mL streptomycin (Life Technologies) and 10% FBS) and incubated at 37 °C and 5% CO2. After 24 h, the cells reached 70–80% confluence and the growth medium was replaced with low-FBS DMEM (2% FBS). For each well, 2 µg of total DNA were used for transfections. Endotoxin-free plasmids were mixed in a 1:3 mass ratio of the heavy and light chains, respectively. The DNA was mixed with 2.5 µL of 2 M CaCl2 and water to a 20 µL volume. Finally, 20 µL of 2X HBS (180 mM NaCl, 50 mM acid-free HEPES, 2 mM Na2HPO4, pH 7.1) were added and the transfectant was applied to the cells. The growth medium was replaced with low-glucose DMEM (12–725F, Lonza) without FBS, 24 h after transfection. The supernatants were harvested and passed through 0.2 µm filters (Advantec) 72 h after transfection. Sodium azide was added to a final concentration of 0.01% to each sample, which were then stored at 4 °C.

IgG and HEL-binding ELISA
ELISAs were performed in 96-well plates (Thermo Scientific) coated overnight (4°C) with a goat anti-human IgG primary antibody (109-005-098, Jackson ImmunoResearch) at 10 μg/mL for the detection of IgG; or HEL (L6876, Sigma) at 50 μg/mL for the detection of HEL-binding antibodies. The plates were washed three times with PBS/0.1% Tween-20 and blocked with PBS/3% BSA (IgG ELISA) or PBS/5% skimmed milk (HEL ELISA) for 1 h at room temperature. After washing the plates three more times, 50 μL of the supernatants containing IgG were transferred to the wells and incubated for 2 h at room temperature with gentle shaking. Bound IgG were detected with a goat anti-human IgG antibody coupled with HRP (ab98567, Abcam). The assay was developed with OPD and H₂O₂, and 15 min later the absorbance was read at 490 nm.

Disclosure of Potential Conflicts of Interest
No potential conflict of interest was disclosed.

Acknowledgments
This work was supported by Consejo Nacional de Ciencia y Tecnología [S0008–14120, CB-2009–01–13376].

Supplemental Materials
Supplemental materials may be found here:
www.landesbioscience.com/journals/mabs/article/27435

References
1. Benichou J, Ben-Hamo R, Louzoun Y, Efroni S. Rep-Seq: uncovering the immunological repertoire through next-generation sequencing. Immunology 2012; 135:183-91; PMID:22043864; http://dx.doi.org/10.1111/j.1365-2567.2011.03527.x
2. Fischer N. Sequencing antibody repertoires: the next generation. MAbs 2011; 3:17-20; PMID:21899780; http://dx.doi.org/10.4161/mabs.3.1.14169
3. Arnaout R, Lee W, Cahill P, Honan T, Sparrow T, Weiand M, Nusbaum C, Rajewsky K, Korall SB. High-resolution description of antibody heavy-chain repertoires in humans. PLoS ONE 2011; 6:e22365; PMID:21828618; http://dx.doi.org/10.1371/journal.pone.0022365
4. Boudinot P, Marriott-Ferrandiz ME, Pasquier LD, Benmansour A, Cazenave PA, Six A. New perspectives for large-scale repertoire analysis of immune receptors. Mol Immunol 2008; 45:2437-45; PMID:18279958; http://dx.doi.org/10.1016/j.molimm.2007.12.018
5. Glavinle J, Zhai W, Berka J, Telman D, Huerta G, Mela GR, Ni L, Mei L, Sundar PD, Day GM, et al. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. Proc Natl Acad Sci U S A 2009; 106:20216-21; PMID:19785695; http://dx.doi.org/10.1073/pnas.0907771106
6. Reddy ST, X, Miklos AE, Hughes RA, Kang, SH, Hsu KH, Chrysostomou C, Hanick -Smith IV, Iverson BL, Tucker PW, et al. Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. Nat Biotechnol 2010; 28:965-9; PMID:20802495; http://dx.doi.org/10.1038/nbt.110; PMID:23500657; http://dx.doi.org/10.1016/j.nbt.2009.09.017
7. Ravn U, Didelot G, Venet S, Ravn U, Buonois V, Gueneau F, Calloud S, Kosco- Vilbois M, Fischer N. Transfering the characteristics of naturally occurring and biased antibody repertoires to human antibody libraries by trapping CDRH1 sequences. PLoS One 2012; 7:e43471; PMID:12950753; http://dx.doi.org/10.1371/journal.pone.0043471
8. Ravn U, Gueneau F, Baerlocher L, Osteras M, Desmurs M, Malinge P, Magistrelli G, Farinelli L, Kosco- Vilbois MH, Fischer N. By-passing in vitro screening—next generation sequencing technologies applied to antibody display and in silico candidate selection. Nucleic Acids Res 2010; 38:e193; PMID:20849698; http://dx.doi.org/10.1093/nar/gkq289
9. Wu A, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, Chen X, Longo NS, Louder M, McKee K, et al.; NISC Comparative Sequencing Program. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. Science 2011; 333:1593-602; PMID:21839983; http://dx.doi.org/10.1126/science.1207552
10. Smith-Gill SJ, Hamel PA, Lovbe TB, Dorrington KJ. Contributions of immunoglobulin heavy and light chains to antibody specificity for lysozyme and two haptons. J Immunol 1987; 139:4155-44; PMID:2447160
11. Kurosawa N, Yoshioka M, Fujimoto R, Yamagishi F, Isobe M. Rapid production of antigen-specific monoclonal antibodies from a variety of animals. BMC Biol 2012; 10:80; PMID:23007270; http://dx.doi.org/10.1186/1741-7007-10-80
12. Tiller T, Meffre E, Yurasov S, Tsuiji M, Nussenzweig MC, Wardemann H. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression cloning. J Immunol Methods 2008; 329:112-24; PMID:17996249; http://dx.doi.org/10.1016/j.jim.2007.09.017
13. Yoshioka M, Kurosawa N, Isobe M. Target-selective joint polymerase chain reaction: a robust and rapid method for high-throughput production of recombinant monoclonal antibodies from single cells. BMC Biotechnol 2011; 11:75; PMID:21774833; http://dx.doi.org/10.1016/j.mabs.2011.10.008
14. Saggie I, Wine Y, Shfelt-Carasso L, Nahary L, Georgiou G, Benhar I. Antibody isolation from immunized animals: comparison of phage display and antibody discovery via gene repertoire mining. Protein Eng Des Sel 2012; 25:539-49; PMID:22988130; http://dx.doi.org/10.1093/protein/gzs060
15. Vener S, Ravn U, Buonois V, Gueneau F, Calloud S, Kosco- Vilbois M, Fischer N. Transfering the characteristics of naturally occurring and biased antibody repertoires to human antibody libraries by trapping CDRH1 sequences. PLoS One 2012; 7:e43471; PMID:22950753; http://dx.doi.org/10.1371/journal.pone.0043471
16. Tiller T, Meffre E, Yurasov S, Tsuiji M, Nussenzweig MC, Wardemann H. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression cloning. J Immunol Methods 2008; 329:112-24; PMID:17996249; http://dx.doi.org/10.1016/j.jim.2007.09.017
17. Chardès T, Villard S, Ferrière G, Pichaczky M, Cerutti M, Devauchelle G, Pau B. Efficient amplification and direct sequencing of mouse variable regions from any immunoglobulin gene family. FEBS Lett 1999; 452:386-94; PMID:10386627; http://dx.doi.org/10.1016/S0014-5793(99)01064-9
18. Dittramajumdar AK, Jacobson DP, Hood LE, Osman GE. Rapid cloning of any rearranged mouse immunoglobulin variable genes. Immunogenetics 1996; 43:141-51; PMID:8550098; http://dx.doi.org/10.1007/BF00176675
19. Esoño S, Frobert Y, Grassi J, Crémion C, Bosquet D. A general method allowing the design of oligonucleotide primers to amplify the variable regions from immunoglobulin cDNA. J Immunol Methods 2003; 279:251-66; PMID:12969565; http://dx.doi.org/10.1016/S0022-1759(03)00242-4
20. Heinrichs A, Miletin C, Gherardi E. Universal cloning and direct sequencing of rearranged antibody V genes using C region primers, biotin-captured cDNA and one-side PCR. J Immunol Methods 1995; 178:241-51; PMID:7836876; http://dx.doi.org/10.1016/S0022-1759(94)00262-U
21. Roharsi S, Ganju P, Sehgal D. Systematic design and testing of nested (RT-)PCR primers for specific amplification of mouse rearranged/expressed immunoglobulin variable region genes from small number of B cells. J Immunol Methods 2008; 339:205-19; PMID:18926828; http://dx.doi.org/10.1016/j.jim.2008.09.017
22. Seijen AM, Seijen HG, Bos NA. Systematic design of mouse Vh gene family-specific oligonucleotides. J Immunol Methods 2001; 254:161-8; PMID:11406161; http://dx.doi.org/10.1016/S0022-1759(01)00396-9
29. Wang Z, Raifu M, Howard M, Smith L, Hansen D, Goldsby R, Ratner D. Universal PCR amplification of mouse immunoglobulin gene variable regions: the design of degenerate primers and an assessment of the effect of DNA polymerase 3′ to 5′ exonuclease activity. J Immunol Methods 2000; 233:167-77; PMID:10648866; http://dx.doi.org/10.1016/S0022-1759(99)00184-2

30. Mangiapan G, Vokurka M, Schouls L, Cadranel J, Lecossier D, van Embden J, Hance AJ. Sequence capture-PCR improves detection of mycobacterial DNA in clinical specimens. J Clin Microbiol 1996; 34:1209-15; PMID:8727904

31. Hansen-Hagge TE, Trefzer U, zu Reventlow AS, Kaltoft K, Sterry W. Identification of sample-specific sequences in mammalian cDNA and genomic DNA by the novel ligation-mediated subtraction (Limes). Nucleic Acids Res 2001; 29:E20; PMID:1166940; http://dx.doi.org/10.1093/nar/29.4.e20

32. Jirholt P, Ohlin M, Borrebaeck CA, Söderlind E. Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework. Gene 1998; 215:471-6; PMID:9714846; http://dx.doi.org/10.1016/S0378-1119(98)00317-5

33. Xu JL, Davis MM. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. Immunity 2000; 13:37-45; PMID:10933393; http://dx.doi.org/10.1016/S1074-7613(00)00006-6

34. DeKosky BJ, Ippolito GC, Deschner RP, Lavinder JJ, Wise Y, Rawlings BM, Varadarajan N, Giesecke C, Dörner T, Andrews SF, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. Nat Biotechnol 2013; 31:166-9; PMID:23334449; http://dx.doi.org/10.1038/nbt.2492

35. Morgan RA, Dudley ME, Wunderlich JR, Hughes MS, Yang JC, Sherry RM, Royal RE, Topalian SL, Kammula US, Restifo NP, et al. Cancer regression in patients after transfer of genetically engineered lymphocytes. Science 2006; 314:126-9; PMID:16946036; http://dx.doi.org/10.1126/science.1129003

36. Matz M, Shagin D, Bogdanova E, Britanova O, Lukyanov S, Diatchenko L, Chenchik A. Amplification of cDNA ends based on template-switching effect and step-out PCR. Nucleic Acids Res 1999; 27:1558-60; PMID:10037822; http://dx.doi.org/10.1093/nar/27.6.1558

37. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics 2011; 27:863-4; PMID:21278185; http://dx.doi.org/10.1093/bioinformatics/btr026

38. Ye J, Ma N, Madden TL, Ostrell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res 2013; 41:W34-40; PMID:23671333; http://dx.doi.org/10.1093/nar/gkt382

39. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010; 26:2460-1; PMID:20709691; http://dx.doi.org/10.1093/bioinformatics/btq461

40. Brochet X, Lefranc MP, Giudicelli V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-D-J and V-D-J sequence analysis. Nucleic Acids Res 2008; 36:W503-8; PMID:18503082; http://dx.doi.org/10.1093/nar/gkn1316

41. Leamon JHG, Crouchman TP. Improved antibiotic-resistance gene cassettes and omega elements for Escherichia coli vector construction and in vitro deletion/insertion mutagenesis. Gene 1995; 160:63-7; PMID:7628718; http://dx.doi.org/10.1016/0378-1119(95)00108-1