Protein Engineering, Design & Selection vol. 27 no. 10 pp. 301–307, 2014
Published online September 1, 2014 doi:10.1093/protein/gzu032

From deep sequencing to actual clones

Sara D'Angelo[1,4], Sandeep Kumar[2,3], Leslie Naranjo[2], Fortunato Ferrara[1], Csaba Kiss[2] and Andrew R.M. Bradbury[2,4]

[1]New Mexico Consortium, Los Alamos, NM, USA, [2]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA and [3]Present address: Compugen USA, Inc., San Francisco, CA 94080, USA

[4]To whom correspondence should be addressed.
E-mail: amb@lanl.gov (A.R.M.B.); sdangelo@lanl.gov (S.D.)

Received June 18, 2014; revised June 18, 2014; accepted July 3, 2014

Edited by Valerie Daggett

The application of deep sequencing to *in vitro* display technologies has been invaluable for the straightforward analysis of enriched clones. After sequencing *in vitro* selected populations, clones are binned into identical or similar groups and ordered by abundance, allowing identification of those that are most enriched. However, the greatest strength of deep sequencing is also its greatest weakness: clones are easily identified by their DNA sequences, but are not physically available for testing without a laborious multistep process involving several rounds of polymerization chain reaction (PCR), assembly and cloning. Here, using the isolation of antibody genes from a phage and yeast display selection as an example, we show the power of a rapid and simple inverse PCR-based method to easily isolate clones identified by deep sequencing. Once primers have been received, clone isolation can be carried out in a single day, rather than two days. Furthermore the reduced number of PCRs required will reduce PCR mutations correspondingly. We have observed a 100% success rate in amplifying clones with an abundance as low as 0.5% in a polyclonal population. This approach allows us to obtain full-length clones even when an incomplete sequence is available, and greatly simplifies the subcloning process. Moreover, rarer, but functional clones missed by traditional screening can be easily isolated using this method, and the approach can be extended to any selected library (scFv, cDNA, libraries based on scaffold proteins) where a unique sequence signature for the desired clones of interest is available.

Keywords: antibody/inverse PCR/deep sequencing/phage display/yeast display

Introduction

Next-generation sequencing (NGS) (Niedringhaus *et al.*, 2011) has been widely implemented in projects that go beyond genome sequencing, for which it was primarily developed: protein evolution (Schlinkmann *et al.*, 2012), computationally designed drug screening (Whitehead *et al.*, 2012), *in vitro* (Ravn *et al.*, 2010) and *in vivo* (Reddy *et al.*, 2010) antibody selections (reviewed in Georgiou *et al.*, 2014), interactome (Di Niro et al., 2010) analysis and immune repertoire profiling (Glanville *et al.*, 2011) have all benefited greatly from the ability to sequence massive numbers of clones.

In vitro display technologies (reviewed in (Rothe *et al.*, 2006; Bradbury *et al.*, 2011)) represent the common denominator for most of the selection/evolution approaches in that they allow the isolation of biomolecules with desired properties from large libraries, using multiple rounds of selection. Traditionally, candidate binders are identified by extensive individual clone screening using microtiter plates. This is an inefficient approach beset by redundancy for abundant clones, and sparse, or absent, representation of clones present at lower frequencies within the selected population. Our group (Di Niro *et al.*, 2010), along with others (Ravn *et al.*, 2010), has pioneered an alternative approach that exploits the use of NGS in *in vitro* display selection analyses. The entire selection output (usually 10^{5-6} clones) is analyzed by deep sequencing. Sequences are binned, ranked, and a rapid assessment of the abundance and identity of positive clones is easily obtained. In addition, rarer clones that would not have been identified by standard screening may be found (D'Angelo *et al.*, 2014) as well as potentially cross-reactive or polyspecific clones (Ferrara *et al.*, 2013).

One of the most challenging examples in the *in vitro* selection field is represented by antibodies. Their simplest recombinant format, the scFv (single-chain fragment variable) (Huston *et al.*, 1988), has been widely used to select target-specific binders (Bradbury *et al.*, 2011). In the scFv, the variable domains (VH and VL), responsible for antigen-binding activity, are connected by a flexible linker. Ideally, the complete VH and VL genes would be sequenced by NGS. However, the technology is presently limited by attainable read lengths. Consequently, when selection outputs of a scFv library are characterized by NGS, analysis is usually restricted to HCDR3 (Heavy Chain Complementarity Determining Region 3), the most variable of the six hyper-variable loops present in the VH and VL chains. HCDR3 shows wide variations in length, structure, shape and sequence (Morea *et al.*, 1998), as well as intrinsic conformational diversity (James *et al.*, 2003), reflecting the importance of HCDR3s in antibody-binding specificity (Xu and Davis, 2000). Although HCDR3 is the most diverse, variability is also found in the five other complementarity determining region (CDRs), as well as, to a lesser extent, the four framework regions flanking the CDRs (Fig. 1a). Altogether, these features contribute to antibody diversity and consequent antigen-binding activity.

While the complexity and depth of an antibody library/selection can be assessed from deep sequencing using appropriate algorithms (AbMining Toolbox (D'Angelo *et al.*, 2014), VDJFasta (Glanville *et al.*, 2009)), the rescue of identified clones is another matter. Deep sequencing provides vast amounts of useful information, but positive clones need to be separately isolated using the obtained sequence information. A number of rescue strategies have been reported for antibody-based constructs, including correlation with randomly picked clone (Ravn *et al.*, 2010), fragment assembly (Ravn *et al.*, 2010) and gene synthesis (Saggy *et al.*, 2012). The approach of random clone picking involves correlating the full (Sanger)

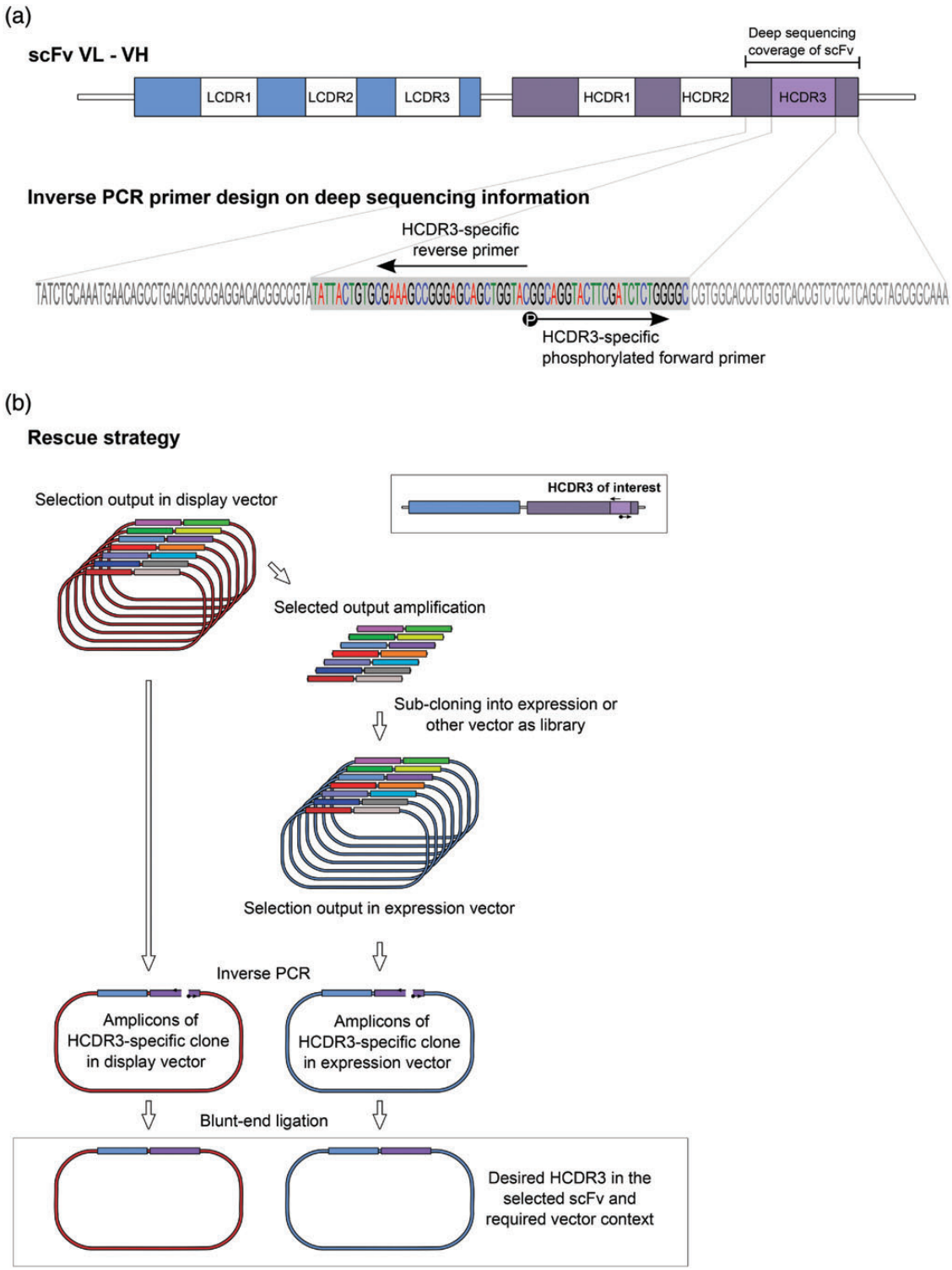

Fig. 1. Inverse PCR strategy applied to antibodies. (**a**) The structure of the scFv gene. The variable light (VL) and heavy (VH) chain CDRs are indicated by white boxes. The portion of the VH covered by deep sequencing is identified and magnified: the HCDR3, as identified by the AbMining tool, is boxed in gray and the inverse PCR primers are shown. A black circle on the forward primer indicates phosphorylation. (**b**) Schematic representation of the specific HCDR3 rescue strategy: the desired HCDR3-specific back-to-back primers are designed on the HCDR3 sequence as obtained from deep sequencing. Plasmid DNA obtained from the selection output is used as a template for inverse PCR allowing the isolation of the desired scFv molecule identified by *in vitro* selection. The amplification is carried out on the selected output either in its original display vector context (left) or after subcloning into a suitable expression vector (right). The final product is a plasmid carrying the specific scFv.

sequences of the positive clones from the screening assay with deep-sequencing analyses (Di Niro *et al.*, 2010; Ravn *et al.*, 2010), and does not necessarily provide the most abundant NGS clones. Fragment assembly requires multi-step polymerization chain reaction (PCRs, Ravn *et al.*, 2013), and is time consuming with the risk of artifactual recombination, while gene synthesis, requiring full-length sequencing, is expensive. Furthermore, the methods mentioned above are relatively low throughput

compared with the high throughput of *in vitro* antibody selections and NGS. Here, we present a rapid method to isolate clones of interest directly from a selected library using an inverse PCR (Hoskins *et al.*, 2005) and ligation (Fig. 1b), an approach successfully implemented to screen complementary (cDNA) libraries (Hoskins *et al.*, 2005) and to generate domain-focused sublibraries (Pedelacq *et al.*, 2011). As used here, it involves the synthesis of HCDR3-specific inverse PCR primers for each clone, and can be adapted to any type of non-scaffold- or scaffold-based library.

Materials and methods

Bacterial and yeast strains

DH5aF′: F′/endA1 hsdR17(rKmK+) supE44 thi-1 recA1 gyrA (Na1r) relA1 D(lacZYAargF) U169 (m80lacZDM15)

Omnimax (Life Technologies): F′ {proAB lacIq lacZM15 Tn10(TetR) (ccdAB)} mcrA (mrr hsdRMS-mcrBC) 80(lacZ)M15 (lacZYA-argF)U169 endA1 recA1 supE44 thi-1 gyrA96 relA1 tonA panD

BL21(DE3)Gold (Agilent): *E. coli* B F− ompT hsdS(rB− mB−) dcm+ Tetr gal λ(DE3) endA Hte

EBY100 (kindly provided by Prof. Dane Wittrup): MATa AGA1::GAL1-AGA1::URA3 ura3-52 trp1 leu2-delta200 his3-delta200 pep4::HIS3 prb11.6R can1 GAL

Lactobacillus acidophilus [American type culture collection (ATCC)4356] and *L. helveticus* (ATCC521) were obtained from ATCC.

scFv antibody selections

The targets for the scFv phage display selections were the full-length *in vivo* biotinylated His-tagged CDK2 protein (NP_001789.2), produced by SGC-Toronto and the *L. acidophilus*, grown under anaerobic conditions (5% CO_2, 37°C, in *Lactobacilli* MRS broth). The naïve scFv library described in Sblattero and Bradbury (2000) was used for two rounds of phage display against the two antigens. For the anti-CDK2 selections, two additional rounds of yeast display were performed. The detailed protocols for antibody selections against biotinylated proteins and whole bacterial cells are described in Ferrara *et al.* (2012) and Close *et al.* (2013), respectively.

Deep sequencing

The plasmid DNA of the anti-CDK2 second sort output and of the anti-*L. acidophilus* second phage output were extracted and used as template for the PCR targeting the HCDR3 region of scFvs. Briefly, a set of 18 forward primers mapping on the framework region upstream of the HCDR3 and carrying one of the Ion Torrent sequencing adaptors were used in combination with a barcoded reverse primer mapping on the common SV5 tag region of both display vectors and carrying the second adaptor required for sequencing. The primer sequences and method are described in detail in D'Angelo *et al.* (2014). Once amplified with the proofreading Phusion polymerase (NEB), gel extracted and quantified (Q-bit, HS-DNA kit, Invitrogen), the amplicon libraries were processed using the Ion Xpress Amplicon library protocol and then prepared for sequencing on the Ion 316 Chip (Life Technologies).

For sequence analysis, we used the AbMining Toolbox as described in D'Angelo *et al.* (2014). Briefly, the barcoded sequences were quality trimmed, parsed by barcode (each barcode identifying a specific selection output) and processed for the identification of the HCDR3. Identified HCDR3s were translated into amino acid sequences and clustered at Hamming distance 1 to minimize the effect of sequencing errors. Finally, HCDR3 clusters were ranked by abundance. For each unique aa HCDR3 sequence, the corresponding DNA consensus was obtained through AbMining.

Primer design and inverse PCR

The inverse PCR primers were designed on the DNA consensus sequence for the HCDR3 of interest as back to back primers directed outwards from the middle of the HCDR3. Standard rules were followed, when possible, for primer design: common annealing temperature, minimal self-annealing, and presence of a G/C-clamp at the 3′-end. The primers Tm, self-complementarity and GC content were checked with the Oligo Calc tool (Kibbe, 2007). For each HCDR3-specific primer pair, the forward primer was phosphorylated with T4 PNK (NEB)-0.25 U for 10 μg of primer, at 37°C for 30 min, followed by enzyme inactivation at 70°C for 5 min). Phosphorylation allows subsequent ligation of the blunt-end PCR product. The inverse PCR was carried out using a highly processive and high-fidelity polymerase with proof-reading activity (Phusion High Fidelity Polymerase, NEB) and 0.1 ng of template DNA (100–1000 times the diversity of the selection output). After amplification, the correct PCR product was gel extracted and purified (Qiaquick Gel extraction kit, Qiagen) to avoid contamination from the original plasmid template. An inverse PCR for the anti-CDK2 selection output was carried out directly on the plasmid prep obtained from the yeast sorted population. One hundred ng of the purified product were blunt-end ligated with T4 ligase for 2 h at 23°C and transformed into DH5aF′ bacterial cells. In contrast, the inverse PCR for the anti-*L. acidophilus* selections was carried on the plasmid prep of the phage selected scFv population subcloned into a modified pEP expression vector. Briefly, the phage plasmid prep output was digested with BssHII/NheI (NEB) for 4 h at 37°C, then gel extracted and ligated with 1 μg of a pEP vector carrying compatible ends. The bacteria transformed with the pEP subcloned library were harvested and plasmid DNA was extracted to be used as template for the HCDR3-specific inverse PCR. The blunt-end ligation of the HCDR3-specific inverse PCR was then transformed into BL21(DE3)Gold cells to allow subsequent expression of the soluble scFv.

Before carrying out binding assays, single clones for each transformation were analyzed by Sanger sequencing in order to confirm the presence of the correct HCDR3 and obtain the sequence of the full-length scFv. The anti-CDK2 sequenced plasmid clones were then re-transformed into the EBY100 yeast display strain (Yeast transformation kit, Sigma) for testing by flow cytometry.

Binding assays

The yeast cells transformed with the three anti-CDK2 monoclonal scFvs constructs were induced as described in Chao *et al.* (2006) and stained with 100 nM of biotinylated antigen. The antigens used were either CDK2 or the unrelated protein domain of USP11 (NP_004642.2, aa 61–285), as negative control. The binding was detected with streptavidin (APC-conjugated), while the display of each scFv on the yeast surface was measured by anti-SV5 (PE-conjugated) staining.

For the anti-*L. acidophilus* scFvs, *L. acidophilus* and *L. helveticus* fresh cells were grown in *Lactobacilli* MRS broth (5% CO_2, 37°C) and stained with the supernatant of the induced pEP cultures as described in Close *et al.* (2013). Briefly, the three anti-*L. acidophilus* and the negative control anti-lysozyme scFvs clones were expressed in 1 ml of auto-induction media in a 96-deep well plate. Following over night incubation with shaking (1000 rpm) at 30°C, 200 μl of the culture supernatants were added to 10^6–10^7 bacterial cells (either *L. acidophilus* or the negative control *L. helveticus*) and incubation was performed for 1 h at RT. After washing, the scFvs were fluorescently labeled using an anti-SV5 (PE-conjugated) antibody. After 1 h incubation at RT, cells were washed and analyzed. All flow cytometry analyses were performed with the FACS Aria flow cytometer (BD). The data analysis was performed with the FlowJo_V10 (Tree Star, Inc.; Ashland, OR) software.

Table I. Top ranking clones in anti-CDK2 and anti-*L. acidophilus* selections. AA sequence, absolute number of sequences and relative frequency of the HCDR3 are reported. The HCDR3 sequence includes the region between Cys92 and Trp103 (Kabat nomenclature)

CDK2 selection			
Rank	HCDR3	# sequences	% of total
1	CAKGFRAGDAFDIW	9233	23.8
2	CASQGFQGDAFDIW	1912	4.9
3	CASHSGNLGTNGVGDAFDIW	1867	4.8
4	CARGSSGSFDIW	1702	4.4
5	CARPYYGSGDAFDYW	1344	3.5
6	CAHSYGDPFDYW	1264	3.3
7	CARADWIDAFDIW	930	2.4
8	CARPLSGWYGDAFDIW	619	1.6
9	CARGGLTTFDIW	515	1.3
10	CARGGQLSSGYYFDAFDIW	505	1.3
	Total		51.3
L. acidophilus selection			
Rank	HCDR3	# sequences	% of total
1	CSTDDYGGNW	123968	56.9
2	CATGDAFDMW	6000	2.8
3	CARGSLGAFDIW	4080	1.9
4	CRHRXLRVVSW	708	0.3
5	CARDSMWVVAAKRKLHNCFDPW	383	0.2
6	CARHKIREFPHAFEIW	198	0.1
7	CARIGGGKRRSHFDYW	195	0.1
8	CSTDGLLVVSW	175	0.1
9	CARVPERGGDCYSFGIW	153	0.1
10	CARVGDGYNYAFDIW	124	0.1
	Total		62.4

Results

Antibody selections

In order to assess the validity of the inverse PCR for the rescue of antibodies identified after deep sequencing, two different selection outputs from the same naïve recombinant scFv library (Sblattero and Bradbury, 2000) were analyzed. In the first, a full-length CDK2 recombinant protein (NP_001789.2) underwent two rounds of phage display, followed by two rounds of yeast display sorting. We have previously shown (Ferrara *et al.*, 2012) that this approach significantly increases the number of identified target-specific antibodies. The target of the second selection was a gram positive bacterium, *L. acidophilus* (ATCC4356). In this case, two rounds of phage display were carried out (the method is described in Close *et al.* (2013)), and the clones from the second selection output (10^4) were analyzed directly by sequencing.

Deep-sequencing analysis and HCDR3 identification

We analyzed the selection outputs using Ion Torrent sequencing, which was able to comprehensively cover the HCDR3s of the selected scFvs, with a 210-bp average read length. The sequencing depth covered the starting output diversity 4–22 times.

We used the AbMining toolbox (D'Angelo *et al.*, 2014) to analyze the two selection outputs. The identified HCDR3s were clustered at a Hamming distance of 1 (HCDR3 differing by only one residue were classified as a unique sequence) and ranked by abundance. The top 10 ranking clones in each of the selections under analysis are reported in Table I, along with their relative abundances within the sequenced selection

Table II. HCDR3 DNA consensus and primer design. Underlined, the primer mapping on the HCDR3 sequence. The Tm for each forward and reverse primer, calculated according to the nearest-neighbor method, is reported

CDK2 selection			
Rank	DNA consensus	Forward primer Tm	Reverse primer Tm
1	TATTACTGTGCGAAGGGATTTCGGGCTGGTGATGCTTTTGATATCTGGGGC Y Y C A K G F R A G D A F D I W G ATAATGACACGCTTCCCTAAAGCCCGACCACTACGAAAACTATAGACCCCG	52.0	52.0
2	TATTACTGTGCGAGCCAGGGCTTTCAGGGAGATGCTTTTGATATCTGGGGC Y Y C A S Q G F Q G D A F D I W G ATAATGACACGCTCGGTCCCGAAAGTCCCTCTACGAAAACTATAGACCCCG	50.0	53.0
3	TATTACTGTGCGAGCCATTCCGGGAATTTAGGTACTAATGGTGTAGGCGATGCTTTTGATATCTGGGGC Y Y C A S H S G N L G T N G V G D A F D I W G ATAATGACACGCTCGGTAAGGCCCTTAAATCCATGATTACCACATCCGCTACGAAAACTATAGACCCCG	51.0	55.0
L. acidophilus selection			
Rank	DNA consensus	Forward primer Tm	Revrese primer Tm
1	TATTACTGTAGCACAGATGACTACGGTGGTAACTGGGGC Y Y C S T D D Y G G N W G ATAATGACATCGTGTCTACTGATGCCACCATTGACCCCG	50.0	50.0
2	TATTACTGTGCGACCGGGGATGCTTTTGATATGTGGGGC Y Y C A T G D A F D M W G ATAATGACACGCTGGCCCCTACGAAAACTATACACCCCG	49.4	51.7
3	TATTACTGTGCAAGAGGCTCTCTGGGGGCTTTTGATATCTGGGGC Y Y C A R G S L G A F D I W G ATAATGACACGTTCTCCGAGAGACCCCCGAAAACTATAGACCCCG	50.0	51.3

output. While in the case of *L. acidophilus*, it is clear that the selection was dominated by one clone (56.9% of the entire population), with two others just above background, in the case of CDK2, we observed a more diverse polyclonal population where the top 10 clones together made up 51.3% of the entire repertoire.

ScFv rescue by inverse PCR

The rescue strategy to isolate clones for which HCDR3 sequences have been identified is depicted in Fig. 1b. We sought to isolate scFv clones corresponding to the top three identified HCDR3s for each selection. The HCDR3 amino acid and the corresponding consensus DNA, sequences were obtained using AbMining software. The primers were designed on this consensus (Table II) and used in the inverse PCR with the selection output (either in the display vector or after subsequent recloning into suitable downstream expression vectors) as template (Fig. 1b). We opted for a high-fidelity polymerase for the inverse PCR in order to reduce the introduction of unwanted mutations. With the CDK2 output, the clones of interest were directly rescued from the original yeast display vector context. For the anti-*L. acidophilus* selections, the entire phage display selection output was first subcloned into a bacterial expression vector, and the plasmid DNA obtained from this library used as template for the inverse PCR. The latter approach allows the direct isolation of clones of interest within a desired vector context, without the need for subcloning individual clones after the first inverse PCR. In these experiments, the antibody with the lowest abundance tested was 1.9% (anti-*lactobacillus*), and in additional experiments (not shown), we have successfully isolated antibodies where the HCDR3 has an abundance of 0.5%.

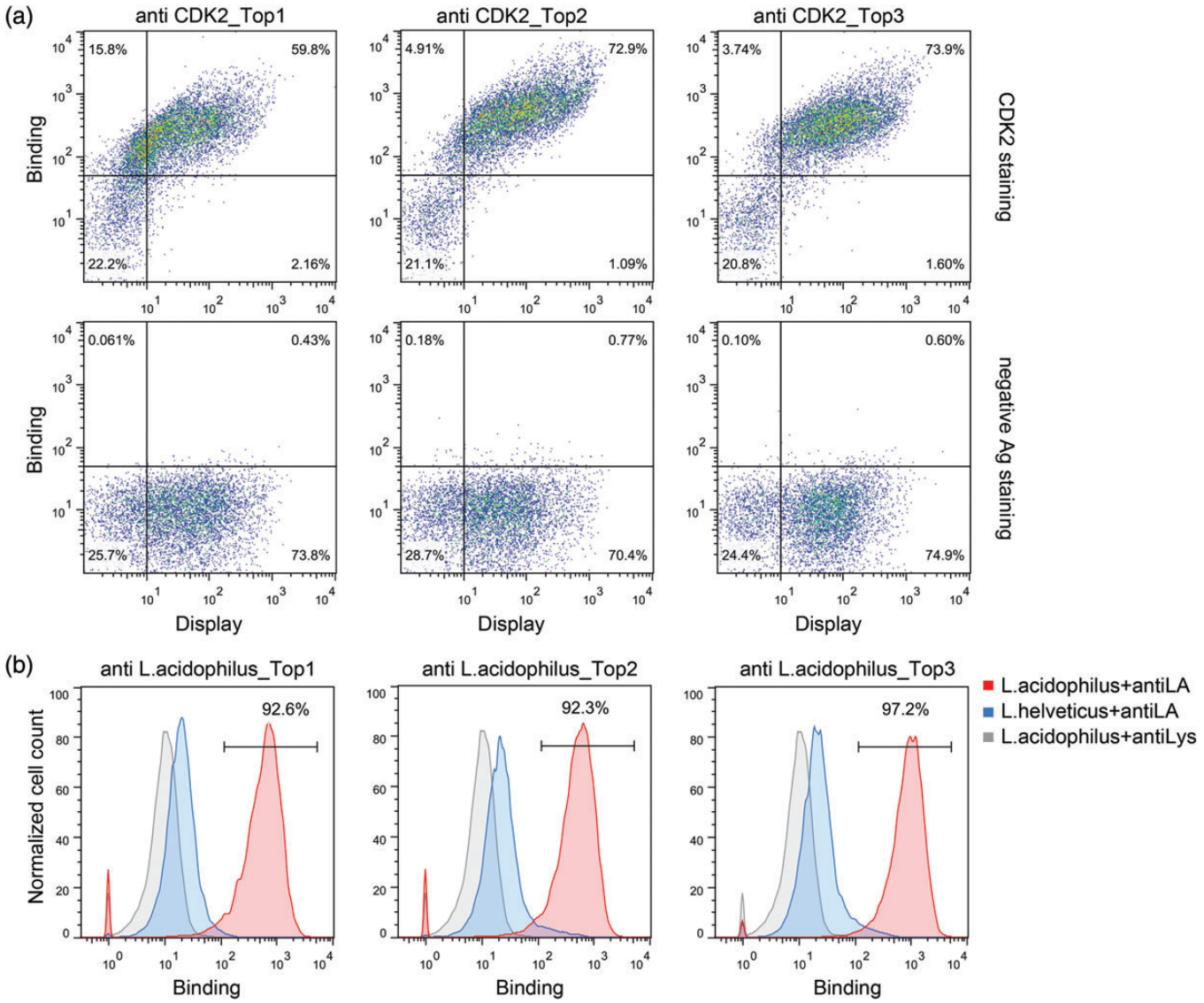

Fig. 2. Binding assessment of the top thre antibodies identified through deep sequencing. (**a**) Top three anti-CDK2 ranking antibodies: the dot plot charts in the top row show a clear shift of CDK2 binding clones into the top-right quadrant of the plot, representing yeast displaying the binding scFv. In the bottom row, no binding was observed against the negative control (a 225-aa domain of USP11, NP_004642.2), demonstrating the specificity of the scFvs. (**b**) Top three anti-*L. acidophilus* ranking scFvs. The histograms show a shift of the bacterial population toward the right end of the chart when bound by the three different anti-*L. acidophilus* scFvs selected. Specificity was assessed by testing the scFvs with the closely related *L. helveticus* species. Anti-lysozyme scFv (anti-Lys) served as negative control.

Binding assessment of rescued clones

Once confirmed by sequencing, the anti-CDK2 single clones were retransformed into yeast cells, while the positive anti-*L. acidophilus* clones were used directly in binding assays. The top three ranking clones from the anti-CDK2 selections were assayed by flow cytometry: scFv binders were displayed on the yeast surface and antigen binding detected with fluorescently labeled antigen (as shown in Fig. 2a). Flow cytometry was also used to assess antibody specificity for the *L. acidophilus* selection. The isolated scFvs specifically bind *L. acidophilus* and not the closely related *L. helveticus*, while the irrelevant anti-lysozyme scFv does not bind either of them (Fig. 2b).

Discussion

The depth and breadth of analysis that can be carried out on the output of selections from display libraries increased dramatically with the advent of NGS. This is only expected to increase as the number and length of reads continues to rise. However, isolating clones identified by NGS from a selection output continues to remain a challenge, particularly when full-length sequences are not available. Unlike previous technology in which large numbers of physical clones were sequenced, and testing their properties involved identifying the microtiter well that contained them, NGS provides only information. In order to test the functional properties of a clone identified by sequencing, it is necessary to use the sequence information to actually isolate the clone. For antibody constructs, this has been addressed using a number of different approaches. In fragment assembly (Ravn *et al.*, 2010, 2013), the HCDR3 sequence is used as an anchor sequence to amplify two fragments, upstream and downstream of the HCDR3, which are then PCR assembled and cloned into a vector of interest. This involves three individual PCR amplifications, as well as a separate cloning step, and runs the risk of additional PCR mutations, as well as artifactual misassembly. Gene synthesis (Saggy *et al.*, 2012) can be carried out where NGS is able to provide the full length of identified clones. However, even though the cost of gene synthesis is becoming significantly cheaper, it is prohibitive on a large scale, and usually takes weeks. A commonly used approach is to pick random clones, sequence them using Sanger sequencing and correlate picked clone sequences with those identified by NGS (Di Niro *et al.*, 2010; Ravn *et al.*, 2010). This is usually effective for abundant clones, but the most common clones identified by NGS cannot always be found (Di Niro *et al.*, 2010; Ravn *et al.*, 2010). While this can probably be overcome by sequencing more clones, this increases complexity and reduces throughput, and becomes impractical for rarer clones, where many more Sanger sequences are required. These methods are all relatively low throughput compared with the high throughput of *in vitro* antibody selections and NGS. The approach described here allows the direct isolation of NGS identified clones of interest from a selected library using an inverse PCR (Hoskins *et al.*, 2005) and ligation (Fig. 1b). Although never applied to NGS outputs, the inverse PCR has been used to generate domain-focused sublibraries (Pedelacq *et al.*, 2011) and to screen cDNA libraries (Hoskins *et al.*, 2005). As used here, it exploits the unique barcode nature of the HCDR3 sequence to synthesize pairs of outward facing primers that are used to amplify plasmids containing the HCDR3. In the examples above, we show the simplicity of the approach, demonstrating that clones identified for their abundance by NGS can be easily isolated and screened for activity. It is important to note that, since the VL and other HCDRs may vary, the HCDR3 does not uniquely identify a particular antibody, and that after carrying out a selection and clone isolation using this approach, a small library of different antibodies all recognizing the same target and containing the same HCDR3 may be isolated. We anticipate that as sequencing read length improves, it will be possible to define more accurately the true diversity of antibodies with identical HCDR3s.

An additional advantage of the approach described here, illustrated with the antibodies recognizing *L. acidophilus*, is that after selection, the complete output can be recloned into a derivative vector prior to the isolation of specific clones identified by NGS. The anti-*L. acidophilus* antibodies were initially selected by phage display, and the complete output cloned into a bacterial expression vector. It was from this secondary expression vector library that the specific antibodies were isolated, ready for immediate testing, without the need for additional recloning. It is clear that the same process can be applied to the same antibody cloned into numerous different vectors, using the same unique inverse primers.

This simple single-step inverse PCR procedure can be applied to the isolation of any gene of interest where unique fingerprints can be identified by NGS, and is likely to become an essential, straightforward, cheap addition to any NGS-based analysis of library selections.

Funding

This work was supported by the National Institutes of Health (1-U54-DK093500-01 to A.R.M.B.); and the Los Alamos National Laboratory LDRD program and NIH grant (1R01HG004852-01A1 to A.R.M.B.). Funding for open access charge: National Institutes of Health.

References

Bradbury,A.R., Sidhu,S., Dubel,S. and McCafferty,J. (2011) *Nat. Biotechnol.*, **29**, 245–254.

Chao,G., Lau,W.L., Hackel,B.J., Sazinsky,S.L., Lippow,S.M. and Wittrup,K.D. (2006) *Nat. Protoc.*, **1**, 755–768.

Close,D.W., Ferrara,F., Dichosa,A.E., *et al.* (2013) *BMC Microbiol.*, **13**, 270.

D'Angelo,S., Glanville,J., Ferrara,F., Naranjo,L., Gleasner,C.D., Shen,X., Bradbury,A.R. and Kiss,C. (2014) *MAbs*, **6**, 160–172.

Di Niro,R., Sulic,A.M., Mignone,F., *et al.* (2010) *Nucleic Acids Res.*, **38**, e110.

Ferrara,F., Naranjo,L.A., Kumar,S., Gaiotto,T., Mukundan,H., Swanson,B. and Bradbury,A.R. (2012) *PLoS ONE*, **7**, e49535.

Ferrara,F., Naranjo,L.A., D'Angelo,S, Kiss,C. and Bradbury,A.R. (2013) *J. Immunol. Methods*, **395**, 83–87.

Georgiou,G., Ippolito,G.C., Beausang,J, Busse,C.E, Wardemann,H. and Quake,S.R. (2014) *Nat. Biotechnol*, **32**, 158–168.

Glanville,J., Zhai,W., Berka,J., *et al.* (2009) *Proc. Natl Acad. Sci. USA*, **106**, 20216–20221.

Glanville,J., Kuo,T.C., von Budingen,H.C., *et al.* (2011) *Proc. Natl Acad Sci. USA*, **108**, 20066–20071.

Hoskins,R.A., Stapleton,M., George,R.A., Yu,C., Wan,K.H., Carlson,J.W. and Celniker,S.E. (2005) *Nucleic Acids Res.*, **33**, e185First published on 2005/12/06, .

Huston,J.S., Levinson,D., Mudgett-Hunter,M., *et al.* (1988) *Proc. Natl Acad. Sci. USA*, **85**, 5879–5883.

James,L.C., Roversi,P. and Tawfik,D.S. (2003) *Science*, **299**, 1362–1367.

Kibbe,W.A. (2007) *Nucleic Acids Res.*, **35**, W43–W46.

Morea,V., Tramontano,A., Rustici,M., Chothia,C. and Lesk,A.M. (1998) *J. Mol. Biol.*, **275**, 269–294.

Niedringhaus,T.P., Milanova,D., Kerby,M.B., Snyder,M.P. and Barron,A.E. (2011) *Anal. Chem.*, **83**, 4327–4341.

Pedelacq,J.D., Nguyen,H.B., Cabantous,S., *et al*. (2011) *Nucleic Acids Res.*, **39**, e125.
Ravn,U., Gueneau,F., Baerlocher,L., *et al*. (2010) *Nucleic Acids Res.*, **38**, e193.
Ravn,U., Didelot,G., Venet,S., Ng,K.T., Gueneau,F., Rousseau,F., Calloud,S., Kosco-Vilbois,M. and Fischer,N. (2013) *Methods*, **60**, 99–110.
Reddy,S.T., Ge,X., Miklos,A.E., *et al*. (2010) *Nat. Biotechnol.*, **28**, 965–969. First published on 2010/08/31.
Rothe,A., Hosse,R.J. and Power,B.E. (2006) *FASEB J.*, **20**, 1599–1610.
Saggy,I., Wine,Y., Shefet-Carasso,L., Nahary,L., Georgiou,G. and Benhar,I. (2012) *Protein Eng. Des. Sel.*, **25**, 539–549.
Sblattero,D. and Bradbury,A. (2000) *Nat. Biotechnol.*, **18**, 75–80.
Schlinkmann,K.M., Honegger,A., Tureci,E., Robison,K.E., Lipovsek,D. and Pluckthun,A. (2012) *Proc. Natl Acad. Sci. USA*, **109**, 9810–9815.
Whitehead,T.A., Chevalier,A., Song,Y., *et al*. (2012) *Nat. Biotechnol.*, **30**, 543–548.
Xu,J.L. and Davis,M.M. (2000) *Immunity*, **13**, 37–45.