

# Recombinant renewable polyclonal antibodies

Fortunato Ferrara<sup>1,†</sup>, Sara D'Angelo<sup>1,†</sup>, Tiziano Gaiotto<sup>1,†</sup>, Leslie Naranjo<sup>2</sup>, Hongzhao Tian<sup>2</sup>, Susanne Gröslund<sup>3</sup>, Elena Dobrovetsky<sup>3</sup>, Peter Hrabec<sup>2</sup>, Fridtjof Lund-Johansen<sup>4</sup>, Silvia Saragozza<sup>5</sup>, Daniele Sblattero<sup>5</sup>, Csaba Kiss<sup>2</sup>, and Andrew RM Bradbury<sup>2,\*</sup>

<sup>1</sup>New Mexico Consortium; Los Alamos, NM USA; <sup>2</sup>Los Alamos National Laboratory; Los Alamos, NM USA; <sup>3</sup>Structural Genomics Consortium, University of Toronto, Toronto, ON Canada; <sup>4</sup>University of Oslo; Oslo, Norway; <sup>5</sup>Università del Piemonte Orientale; Novara, Italy;

<sup>#</sup>Present address: National Institute for Biological Standards and Control (NIBSC), Potters Bar, Hertfordshire, UK

<sup>†</sup>These authors contributed equally to this work.

**Keywords:** polyclonal recombinant antibodies, yeast display, phage display, antibody validation

**Abbreviations:** CTBP, C-terminal binding protein; ELISA, enzyme linked immunosorbant assay; HCDR3, Heavy chain complementarity determining region 3; HPA, Human Protein Atlas; scFv, single chain Fv; PrESTs, Protein epitope signature tag; rrpAbs, recombinant renewable polyclonal antibodies; SDS–PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TEV, tobacco etch virus

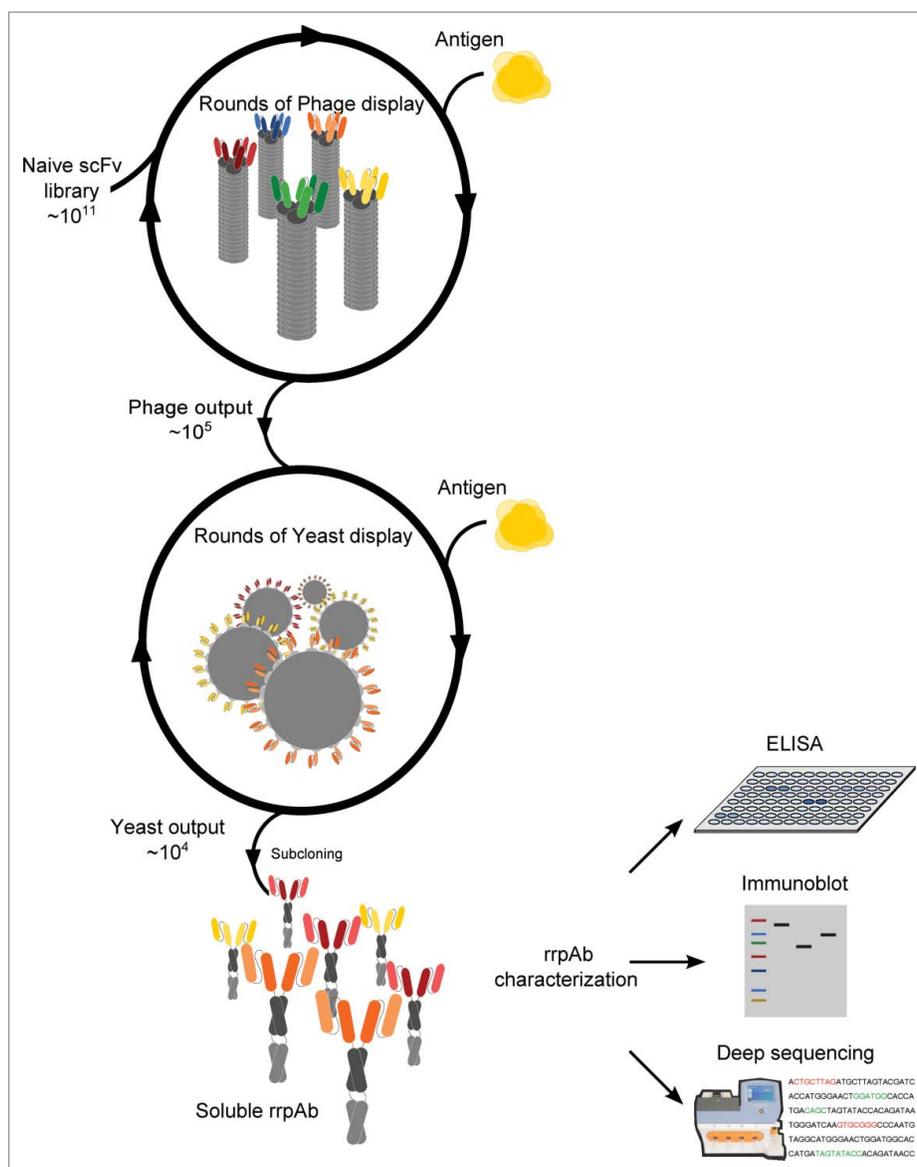
Only a small fraction of the antibodies in a traditional polyclonal antibody mixture recognize the target of interest, frequently resulting in undesirable polyreactivity. Here, we show that high-quality recombinant polyclonals, in which hundreds of different antibodies are all directed toward a target of interest, can be easily generated in vitro by combining phage and yeast display. We show that, unlike traditional polyclonals, which are limited resources, recombinant polyclonal antibodies can be amplified over one hundred million-fold without losing representation or functionality. Our protocol was tested on 9 different targets to demonstrate how the strategy allows the selective amplification of antibodies directed toward desirable target specific epitopes, such as those found in one protein but not a closely related one, and the elimination of antibodies recognizing common epitopes, without significant loss of diversity. These recombinant renewable polyclonal antibodies are usable in different assays, and can be generated in high throughput. This approach could potentially be used to develop highly specific recombinant renewable antibodies against all human gene products.

## Introduction

Traditional polyclonal antibody preparations, generated by the immunization of rabbits, contain barely 0.5–5% of antibodies that will recognize the target.<sup>1</sup> This partly explains the variable quality<sup>2–5</sup> of most commercial polyclonal antibodies, only 20–50% of which actually recognize their targets specifically.<sup>6–8</sup> Although it may be possible to overcome this problem with affinity purification,<sup>1,9</sup> after which a majority of antibodies should recognize the target of interest, the finite and non-renewable nature of polyclonal antibodies limits the amount of specific antibody that can be isolated. Furthermore, as demand increases for highly specific antibodies that are able, for instance, to discriminate between closely related proteins, the potential amount of available antibody becomes substantially less. This represents a barrier to the provision of high-quality, specific antibodies for studies of the ~20,000 protein coding genes identified in the human genome project,<sup>10</sup> relatively few of which have known functions.<sup>11</sup>

The Antibodypedia database<sup>12</sup> describes >1,000,000 antibodies, some with associated experimental data, purportedly recognizing over 90% of all human gene products. While this facilitates the choice of suitable antibodies, it does not overcome the quality problems described above. Furthermore, the choice is often confusing: a majority of commercial antibodies is directed to a small number of popular targets,<sup>2</sup> (e.g., ERBB2 has >1200) and the same original antibodies are frequently sold by different providers under different labels,<sup>12,13</sup> rendering the number of truly unique antibodies significantly lower. This situation reflects the “catch-22” of commercial antibodies: while those against popular targets sell well, those against unknown ones do not, even though the best way for an unknown target to become popular is for new tools, such as antibodies, to become broadly accessible.<sup>14</sup> Research, consequently, tends to be focused in those areas where antibodies are available. These problems prompted the establishment of the Human Protein Atlas (HPA),<sup>15</sup> a program to identify and generate high-quality antibodies against every human gene product. Antibodies produced by the HPA are

\*Correspondence to: Andrew RM Bradbury; Email: amb@lanl.gov  
Accepted: 11/13/2014



**Figure 1.** Selection scheme. After antibody selection by phage display, the selection outputs are transferred to a yeast display vector where additional fluorescence activated cell sorting is carried out. Antibodies are then recloned for secretion and tested for specificity.

affinity purified, polyclonal antibodies generated by the immunization of rabbits with protein epitope signature tags (PrESTs), which are protein fragments identified informatically as being unique in the proteome. However, antibodies produced in this way, although functional, are not renewable, and once a batch has run out, it needs to be replaced via immunization of more rabbits.

In contrast, *in vitro* display methods (e.g., phage/yeast display) yield recombinant antibodies that are usually deconvoluted to single clones for further testing. An advantage of producing antibodies recombinantly rather than by immunization is that the selection conditions can be modified to generate antibodies against specific target forms. These include antibodies recognizing one protein, but not a closely related one,<sup>16</sup> spliced

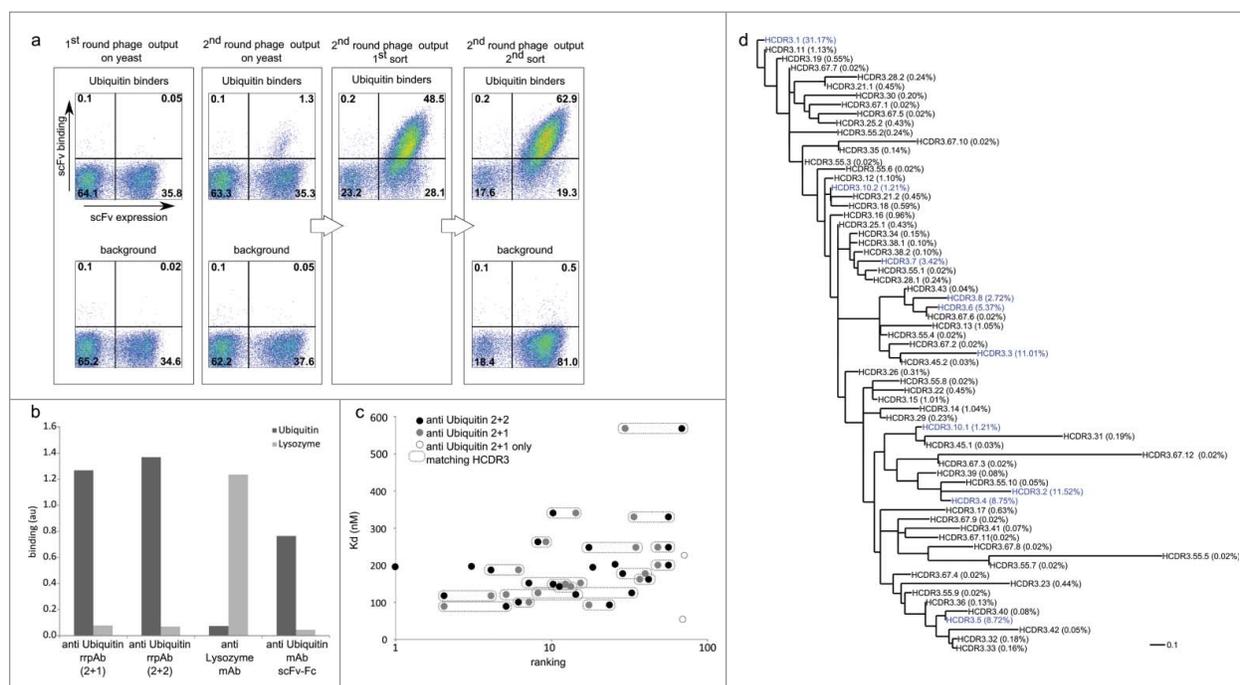
variants,<sup>17,18</sup> specific epitopes,<sup>19</sup> or active conformations.<sup>20-22</sup> Furthermore, affinities or specificities can be improved by *in vitro* evolution,<sup>23,24</sup> and selected antibodies can be fused to additional functional elements, such as enzymes or antibody constant regions. Here, we describe the use of display methods to generate recombinant renewable polyclonal antibodies, as a valid alternative to the non-renewable polyclonal antibody products presently on the market.

## Results

We have previously shown that combining phage and yeast display can significantly increase the number of specific monoclonal antibodies selected against targets from naïve phage antibody libraries, with essentially all antibodies recognizing the target after 2 rounds of phage selection and 2 rounds of yeast sorting.<sup>25</sup> Based on this observation, we opined that the antibody mixtures obtained after such selections could be effective as recombinant renewable polyclonal antibodies (rrpAbs). We tested this hypothesis using the scheme outlined in Fig. 1, carrying out one or 2 rounds of phage selection against ubiquitin, and cloning the output into a yeast display vector. As the goal was to keep “polyclonality” as high as possible to generate and select an antibody population with broad diversity, the target (ubiquitin) concentration used was 400 nM for phage selection and 200 nM for yeast sorting.

Both the *first* and the *second* phage display outputs were used for the subcloning of antibodies into the yeast display system (Fig. 2A). The first round input gave only 0.05% of yeast-displayed antibodies that recognized ubiquitin, with a background of 0.02%, which confirmed previous estimates of positive clone abundance after a single phage selection round.<sup>26</sup> After the second round of phage selection, 1.3% of clones bound the target, with a background of 0.05%.

Given these results, 2 rounds of phage selection were chosen as the yeast display input to facilitate sorting. After one round of yeast sorting (termed 2+1, for 2 rounds of phage selection and one round of yeast sorting), the percentage of positive clones increased to 48.5%. This rose to 62.9% after an additional sorting round (2+2). No further improvement was seen with additional sorting rounds. In both cases, 10,000 positive yeast cells



**Figure 2.** Characterizing the system and establishing optimal selection conditions. **(a)** Flow cytometry plot, in which each dot represents an individual yeast clone displaying antibodies. The number of positive antibody clones is represented by those found in the top-right quadrant that both express scFv and bind target. Only after 2 rounds of phage selection are there enough positive clones to sort effectively. The binding population increases significantly after one or 2 rounds of fluorescence activated cell sorting. **(b)** ELISA testing of normalized amounts of soluble polyclonal antibodies obtained after one or two rounds of sorting. Two monoclonal scFVs in the minibody format (anti-lysozyme and anti-ubiquitin) were used as controls. **(c)** Plot of antibody affinity versus abundance ranking position for the antibodies described in Table 2 after one (gray dots) or two (black dots) sorts. With the exception of the 2 clones indicated by gray circles, antibodies were common to both selections (dashed connectors). **(d)** Phylogram of anti-ubiquitin HCDR3 identified in 2+2 through deep sequencing. Ranking position (HCDR3.x) and abundance within the polyclonal population (%) are shown. The 10 most abundant HCDR3 clones are shown in blue. The x-axis shows evolutionary distance as substitutions per site; scale bar depicts 3 amino acid changes.

were sorted. Both of these sorting outputs were subcloned into a yeast secretion vector, in which the scFv was fused to a rabbit Fc domain, and assessed for antigen recognition and specificity. As shown in Fig. 2B, they were very similar in their reactivity, with 2+2 rounds showing slightly higher and more specific signals.

We assessed the diversity of the 2 populations by Ion Torrent sequencing, using the AbMining Toolbox<sup>27</sup> software package, and found 98 (1 sort) and 67 (2 sorts) unique HCDR3 clusters with a Hamming distance of 1 (i.e., bins in which the individual members differ by no more than a single amino acid from the commonest HCDR3 in that bin). 58 HCDR3 clusters were common to both rounds (Table 1), and a total of 107 different HCDR3s were identified. The affinities of 25 randomly selected different clones (Table 2), which comprised the 10 most abundant and 15 additional HCDR3s, were determined for clones isolated from both sorting rounds. All but 2 clones (gray circles in Fig. 2C) were found in both populations, albeit at different rankings and abundances. Antibody affinities ranged from 51.5–565 nM, spanning the antigen concentrations used for phage selection (400 nM) and yeast sorting (200 nM), and the mean affinity of all measured antibodies was 188 nM, which was very close to the concentration of antigen used for sorting. Importantly, all isolated antibody clones that we tested, including one present at an abundance of only 0.04% and ranked 70th (in the

2+1 selection), recognized ubiquitin (affinity 224.5 nM). None of the tested antibodies showed non-specific binding to an irrelevant negative control, indicating that essentially all selected antibodies recognize the target of interest. Interestingly, there was no significant correlation between abundance ranking and affinity (2-sided Kendall's tau;  $p = 0.08$  for the 2+1 selection and  $p = 0.1$  for the 2+2 selection).

The antibody with the highest measured affinity (51.5 nM) was found after a single sort where it was the 68th most abundant clone. Along with the sorting data, this shows that the polyclonal pool, as well as the monoclonals comprising it, binds ubiquitin specifically. The diversity of the polyclonal population is represented in Fig. 2D, showing all the unique HCDR3s present in the 2+2 selection output in a phylogram rooted on the most abundant clone. The sequence of the top HCDR3 (HCDR3.1 in the figure) is distant from the sequence of the other most abundant clones (HCDR3.2-HCDR3.10.2, shown in blue), demonstrating the diversity of the polyclonal population.

Polyclonal antibodies can be easily obtained by immunization. However, because such antibody mixtures are not renewable, valuable specificities can be easily lost, making scientific reproducibility difficult. To assess the renewability of these yeast-expressed polyclonals, we amplified the secretory yeast pool by growing to saturation after 100-fold dilutions 4 times, for a total

**Table 1.** Deep sequencing characterization of selected anti-ubiquitin scFvs shows the number of unique HCDR3 binned at a Hamming distance of 1 is 98 after one, and 67, after 2 rounds of sorting, with 58 common to both

Sort output	Total # sequences	# unique HCDR3 (Hamming = 1)	# common HCDR3	
2+1	13216	107	98	58
2+2	12778		67	

$10^8$ -fold amplification. After each growth, antibodies were harvested, tested for functionality, and their genes sequenced by Ion Torrent. The top 20 ranked HCDR3s (Fig. 3A) maintain their approximate percentage abundance during the cloning process from display to secretion, as well as over the full one hundred million-fold amplification, with 90% show a less than 2-fold variation after  $10^8$  amplification (shaded box in Fig. 3A). Of the remaining 2 antibodies, the worst was reduced by only 13-fold, representing a maximal growth disadvantage of  $\sim 9\%$  for each division. Furthermore, functionality is maintained through the amplification cycles (Fig. 3B–C), demonstrating that for all practical purposes, this recombinant polyclonal antibody can be repeatedly amplified without losing representation or functionality. Furthermore, it is also important to note that when the antibodies were cloned from the display vector to the yeast secretion vector, the HCDR3 diversity was also maintained.

We next selected antibodies against 8 additional human targets (proteins produced by the Structural Genomics Consortium – Table S1), using 2 rounds of phage and 2 rounds of yeast sorting (with 100 nM antigen). Outputs were recloned for secretion

**Table 2.** Random clones from the selection output tested for binding affinity. The different HCDR3s were identified through Sanger sequencing of the full length scFv. The ranking refers to the position of the specific HCDR3 in the deep sequencing characterization of the polyclonal output after one or 2 sorts

HCDR3	Rank 2+1	Rank 2+2	$K_d$ (nM)
CAKGIAAVDYW	1	1	192.5
CASLRSAIYYDSSGRDAFDIW	4	2	114.5
CAKVRGGDW	3	3	193.6
CARGGQLSSGYFFDAFDIW	6	4	183.9
CARYYYDSSGYADAFDIW	2	5	85.7
CAKGVSGMDVW	7	6	97.1
CTNKGGAFDIW	15	7	148.4
CAKGLGYGMDVW	9	8	259.6
CAKPGNGAFDIW	14	10	337.4
CARAYSSSWYFPDYW	12	10	145.6
CAKGIAAFDYW	13	11	139.5
CARDVGEVPPDLW	5	14	117.7
CARGPGGVGATVLDYW	34	17	244.6
CAKGGFGAFDIW	18	18	191.0
CARDGVPETYSKGGSCWTFDAFDIW	17	23	89.5
CARGVTPDYW	25	25	199.0
CARGGGIAARW	39	28	174.0
CAHMYYYDSSGYFFDYW	8	32	122.0
CARDIGSSPTPFWDYW	36	41	158.7
CARRASAFDIW	33	55	326.9
CAKGGGGFDYW	47	55	244.9
CARGRMDVW	47	55	196.9
CARNANYGYW	29	67	565.1
CAKGAGSLDYW	70	NA	224.5
CVKTIMGAFDIW	68	NA	51.5

in the scFv-Fc format.<sup>28</sup> Binding specificity was first tested by yeast display (Fig. 4A), with all antibodies tested against all targets. All polyclonals appeared to bind their targets specifically, with the exceptions of the antibodies selected against CTBP1 and CTBP2 (proteins that are 78% homologous), which showed some cross-reactivity. These 2 polyclonals were rendered highly specific for their target proteins by carrying out 2 rounds of negative sorting, in which yeast displaying antibodies that did *not recognize* the cross-reacting proteins were isolated, followed by an additional round of positive sorting, for yeast displaying antibodies that did recognize the target protein (Fig. 4B). Testing of the 6 original polyclonals and the improved CTBP polyclonals against all other targets by ELISA showed excellent specificity, with good signal-to-noise ratios (Fig. 4C).

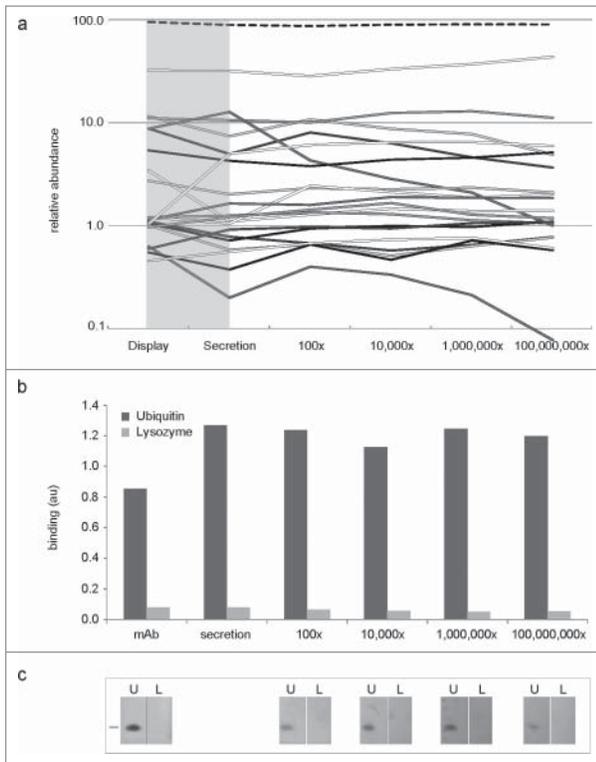
In addition to ELISA, functionality of the polyclonals was assessed by Western blotting and protein arrays. This was considerably facilitated by the use of the rabbit Fc domain in the scFv-Fc fusions, which allowed the use of standard secondary reagents. Initial testing of the selected polyclonals by Western blot (Fig. 4D) showed that 7 of 8 polyclonals were functional. The polyclonal against CDK2 was further assessed for binding on a HEK293 cell lysate, and shown to recognize a specific band of the correct size (Fig. 5A), with no background. The same anti-CDK2 rrpAb was tested in a protein microarray against 71 different proteins and was found to be highly specific, more so than a corresponding commercial polyclonal (sc-163) (Fig. 5B).

The diversity of these polyclonals was also assessed by deep sequencing. Table 3 shows the analysis of the polyclonals at the highest stringency parameters. The number of unique HCDR3s range from 74 to 460 for the 6 polyclonals we analyzed. In the case of CTBP2, subtraction of those antibodies also recognizing CTBP1 reduced the diversity somewhat, from 140 to 96 clusters, making the rrpAb overall less diverse, but more specific.

## Discussion

With the human genome project complete, there is an unmet need for high quality antibodies recognizing the products of the  $\sim 20,000$  identified genes. This requires the ability not to only generate high quality effective antibodies in high throughput and high volumes, but also to ensure minimal batch-to-batch variation. As described in a number of recent publications,<sup>2–5</sup> most commercial antibodies do not fulfill these ideals. Furthermore, commercial antibodies tend to be highly focused on “popular” targets, with over 1000 against p53,<sup>7</sup> for example.

One critical issue is antibody format. Monoclonal antibodies are renewable, but they are expensive to generate and functionally unpredictable. Some are suitable for Western blots, while others



**Figure 3.** Polyclonal amplification and renewability. **(a)** Percentage clone abundance of the commonest 20 clones assessed by deep sequencing, after cloning from display to secretion vector and after 4 one hundred-fold amplifications **(b–c)**: Maintenance of antibody functionality after each one hundred-fold amplification in ELISA **(b)** and Western blotting **(c)**.

are useful only in ELISAs, immunohistochemistry, immunoprecipitation or immunofluorescence. In contrast, polyclonal antibodies generated against proteins have an intrinsic advantage as they recognize multiple epitopes, and they thus have the potential to function in many assays. Disadvantages include their tendency to non-specific binding, and the lack of renewability and reproducibility, as represented in lot-to-lot variation.<sup>29</sup> Valuable specificities often die with the animal that produced them. The Human Protein Atlas has overcome some of these problems by affinity-purifying polyclonals developed in rabbits on the immunizing antigen. While this reduces non-specific binding, it does not overcome the problem of renewability, or the need for immunization.

An ideal research antibody would be highly specific for its target, with neither cross-reactivity nor non-specific binding, usable in all assays, infinitely renewable and definable by sequence. Furthermore, should any properties be less than ideal, e.g., unintended cross-reactivity, it would be possible to improve the antibody relatively easily, either by removing undesirable specificities, or adding desirable ones. Finally, the generation of such research antibodies would be relatively straightforward and possible on the scale needed for proteomic research.

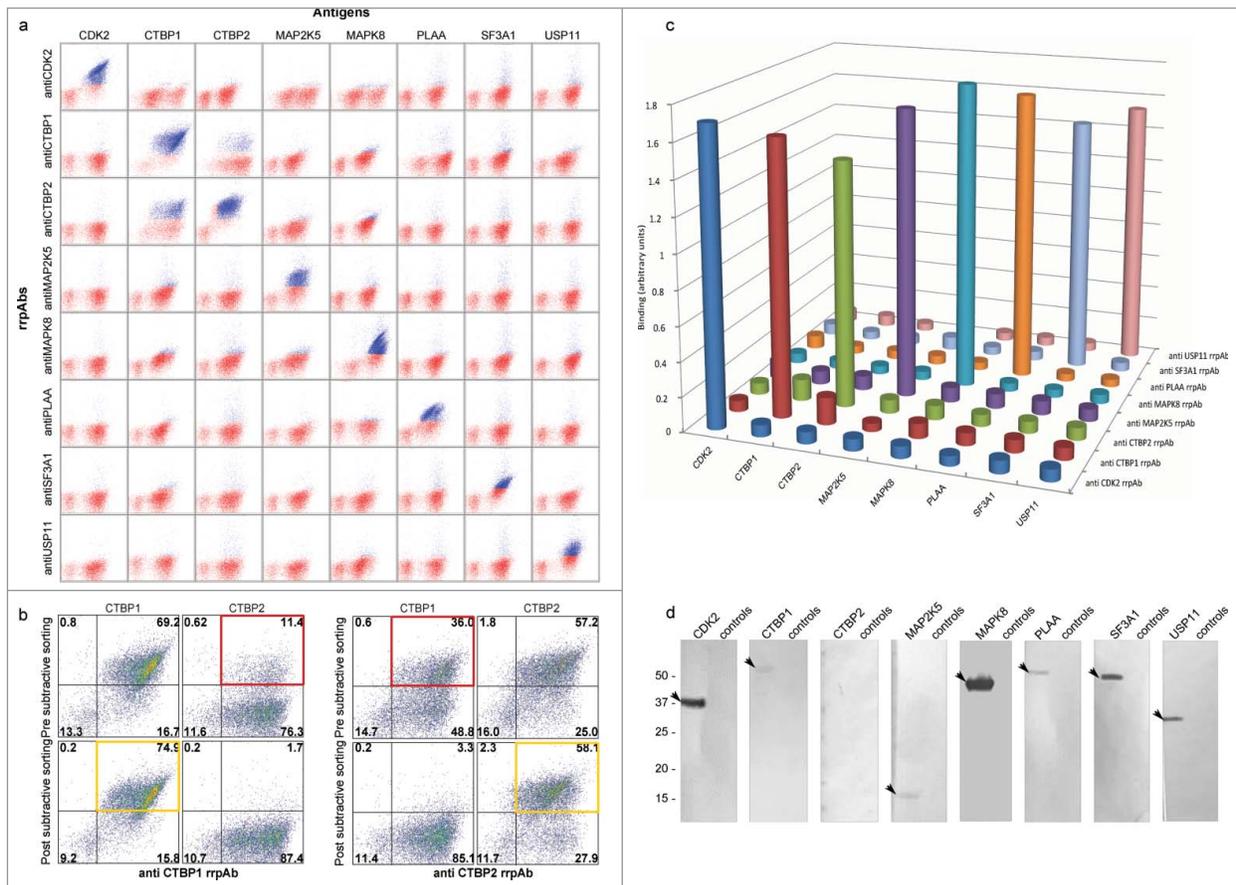
We believe that the development of renewable recombinant polyclonal antibodies (rrpAbs) described here fulfills most of

these criteria. High specificity is a result of the fact that essentially all antibodies recognize the target. This mirrors findings that affinity-purified polyclonal antibodies, in which the vast majority of antibodies recognize the target of interest, give far less background than traditional unfractionated polyclonals. By virtue of their polyclonal nature, these antibodies function in ELISA, Western blotting and protein arrays.

The ability to completely inhibit the Gal1 promoter, and hence the potentially toxic effects of antibody expression during yeast display and secretion, normalizes cell growth and likely allows the enormous amplification demonstrated here. While this was previously shown for a single antibody artificially spiked at 5% into a library,<sup>30</sup> we now show that it is also true for a collection of antibodies all directed toward a single target. The availability of the recombinant polyclonal pool within the context of a yeast display vector makes it very straightforward, as demonstrated with the CTBP1/2 example, to modify the polyclonal properties to eliminate antibodies likely to recognize common epitopes (Fig. S1) and amplify those recognizing unique epitopes (green areas). However, unlike natural polyclonal preparations, in which, at best, each positive or negative selection does no more than purify the antibodies of interest present in the mixture, the ability to amplify recombinant antibodies allows the subset of improved antibodies to be readily produced in far greater quantities. Notwithstanding their potential for amplification, these antibodies cannot be easily defined at the sequence level. This is due to their high diversity, subtle changes in abundance and present limits to next-generation sequencing. However, they do provide an extremely rich source of specific monoclonal antibodies that can be easily isolated after deep sequencing using inverse PCR.<sup>31</sup> Using appropriate selection strategies ensures that all antibodies isolated using this inverse PCR approach have the properties desired, including the complete antibody gene sequence.

The combination of phage and yeast display described here harnesses the advantages of both platforms. Our initial phage antibody library<sup>32</sup> has an estimated diversity of  $\geq 3 \times 10^{11}$ . By carrying out 2 rounds of phage display, we capture the antibodies showing some reactivity for the target. This reduces the diversity to  $10^5$ – $10^6$ , which is compatible with yeast display cloning, and allows further precise sorting to restrict reactivity to those clones recognizing the target of interest. The diversity of the polyclonals after selection was assessed by deep sequencing, with analysis restricted to the HCDR3 because of read length limitations. The number of different HCDR3s selected against the test antigens ranges from 74 to 460 (Table 3), with the actual number of different antibodies likely to be significantly higher when different  $V_L$  chains and additional  $V_H$  mutations are taken into account.

In the single antigen (ubiquitin) in which we assessed the binding abilities of individual monoclonals, all antibodies tested recognized ubiquitin, even down to the least abundant tested antibody, which was present at only 0.04%. This suggests that the vast majority are specific for the target. This depth of diversity allows the removal of cross-reactive clones without significantly affecting the number of different clones: even when antibodies cross-reactive with CTBP1 (or CTBP2) are eliminated, over 90



**Figure 4.** Polyclonal antibody functionality. **(a)** All yeast displayed polyclonal antibodies are specific when tested against all other antigens, as represented by the dot plots on the diagonal, all of which are highly positive (blue dots), with the exception of antibodies against CTBP1 and CTBP2 which show some cross-reactivity. **(b)** Although CTBP1 and CTBP2 are 78% homologous, specific polyclonals recognizing each can be selected by carrying out negative selections. The desired and unwanted specificities are highlighted in the yellow and red quadrants, respectively. **(c-d)** After subcloning for secretion, all polyclonals are highly specific for their target antigens in ELISA **(c)** and Western **(d)**. In the Western blot assay, each rrpAb was tested on its specific antigen (left lane) and on the mix of all the other 7 test antigens (controls lane).

HCDR3 clusters remain. This great diversity may reflect the relatively high target concentrations we used for selection and sorting, which are expected to eliminate binding discrimination between, and so equalize binding of, low and high affinity antibodies. This may also explain the unexpected finding that there was no correlation between antibody affinity and abundance rank (Fig. 2C). Although we used high target concentrations to increase antibody variety, the unexpectedly high diversity obtained indicates that we could probably improve mean antibody affinity without compromising functional polyclonality, by reducing target concentrations during selection. In preliminary experiments (data not shown), we find a much closer correlation between antibody affinity and rank abundance when target concentrations are significantly reduced. Within the context of future genome scale projects, this will facilitate the ready isolation of high quality sequence defined monoclonals from these polyclonal pools without the need for extensive screening.

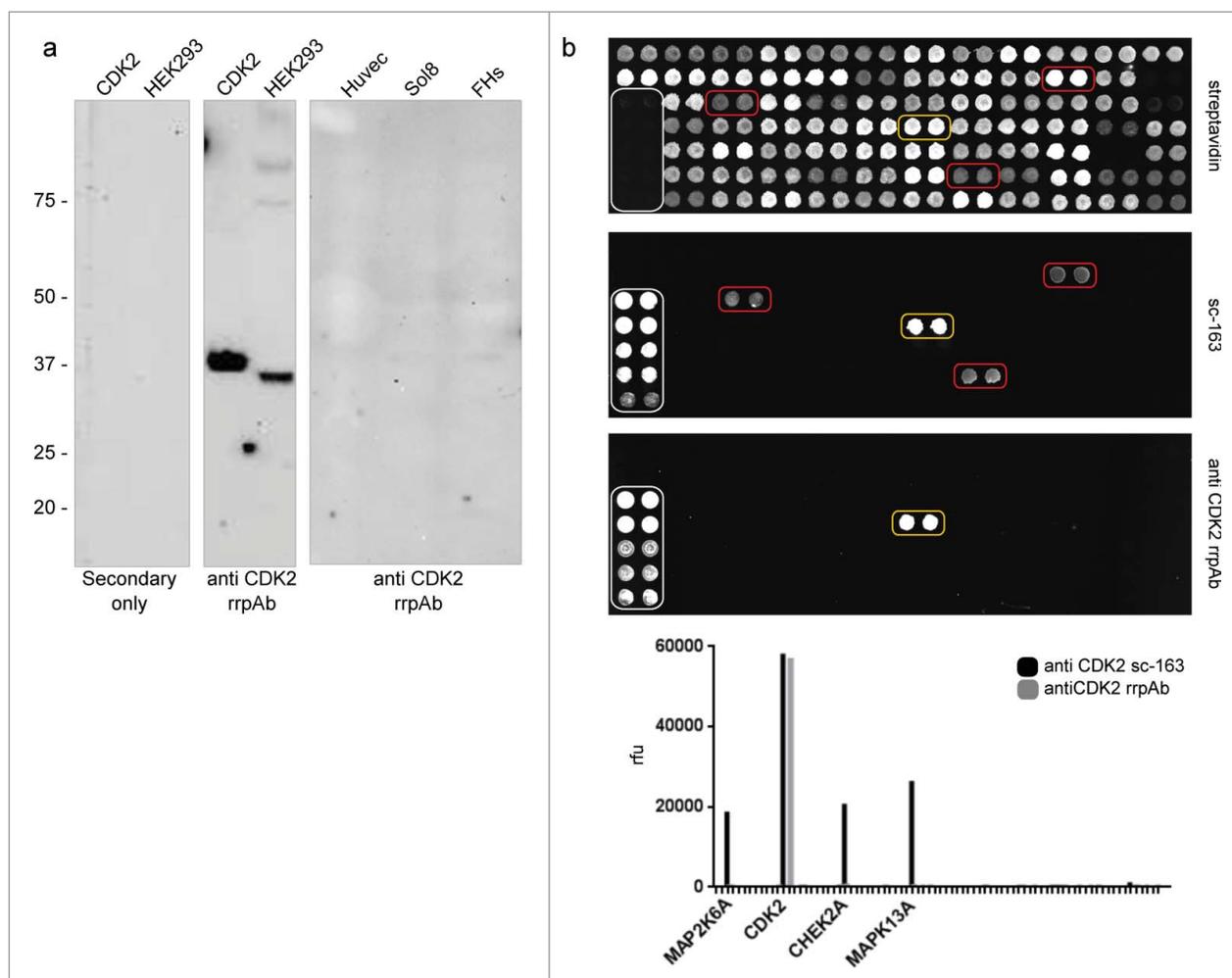
Conceptually, our approach is superficially similar to recombinant polyclonal antibodies,<sup>33</sup> created by the cloning of Fabs from selected *immune* phage libraries into mammalian expression

vectors.<sup>34</sup> However, by using *naïve* phage antibody libraries and yeast display, we are able to avoid the need for immunization and the creation of individual target specific phage display libraries, allowing the development of this approach at the proteomic scale. In conclusion, the approach described here should make the selection of highly specific antibodies against all proteins encoded by the human genome a tractable problem.

## Material and Methods

### Antigens

The 8 antigen domains of interest (Table S1) were subcloned into the expression vector p28BIOHTEV-LIC containing a TEV cleavable N-terminal Avi tag for *in vivo* biotinylation and a C-terminal hexahistidine tag for affinity purification. The proteins were produced in *E. coli* BL21 (DE3) cultivated overnight at 18°C in a LEX system (Harbinger Biotech, Markham, Ontario) and *in vivo* biotinylated by co-expression of BirA ligase and supplementation of 50 µg/L D-Biotin in the culture medium. The



**Figure 5.** Anti-CDK2 polyclonal validation. **(a)** rrpAb WB on cell lysates. The anti-CDK2 rrpAbs is functional in WB recognizing the target CDK2 as purified recombinant protein (lane CDK2) or on HEK293 lysate, recognizing both bands of the correct size (difference in size between the 2 recognized proteins is attributed to the expression tags in the recombinant format). **(b)** Comparison of the commercial anti-CDK2 polyclonal antibody (sc-163) with the anti-CDK2 rrpAb in an array assay. The polyclonal antibodies are challenged against 79 different proteins (detected with streptavidin- top panel). Binding specificity of the commercial (central panel) and of the rrpAb (lower panel) anti-CDK2 is reported in the bottom chart. White squares indicate the IgG calibrator used for normalization. Yellow and red squares indicate the specific and non-specific antigen binding, respectively.

proteins were then purified using a 2-step purification protocol on an ÄKTA Xpress (GE Healthcare), including an immobilized metal chromatography step (5 mL HiTrap Ni-chelating column, GE Healthcare) followed by Gel Filtration (HiLoad XK16/60 Superdex 200 column, GE Healthcare) and eluted in 20 mM HEPES, 300 mM NaCl, 5% glycerol, 0.5 mM TCEP, 0.5 mM EDTA, pH 7.5. Relevant fractions corresponding to mono-disperse peaks were pooled and analyzed by SDS-PAGE and mass spectroscopy to verify protein identity and degree of biotinylation.

#### Phage-display selection

Selection of scFv antibody fragments by phage display from our naïve library<sup>32</sup> was performed as previously described<sup>25</sup>. Briefly, selections were carried out in parallel in solution<sup>35</sup> using the automated Kingfisher magnetic bead system (Thermo Lab Systems): 100–400 nM of biotinylated antigen were used in each

selection cycle. The biotinylated antigens were incubated with the naïve library and the bound phages were captured using  $2 \times 10^7$  streptavidin magnetic beads (Dynabeads M-280).

#### Yeast display and sorting of scFvs

After one or 2 rounds of phage selection, the selected scFv antibodies were subcloned into the yeast display vector as previously described.<sup>25</sup> The selected scFv genes were amplified with specific primers that introduced an overlap with the yeast display vector pDNL6. The vector and the fragments were cotransformed into yeast cells so to allow cloning by gap repair.<sup>36</sup>

The yeast mini-libraries so obtained were further enriched for binders by one or 2 rounds of sorting using flow cytometry according to published methods.<sup>36</sup> After induction,  $2 \times 10^6$  yeast cells were stained with 100–200 nM of biotinylated antigen. Cells were labeled with streptavidin-AlexaFluor633 to detect binding of biotinylated target antigens and anti-SV5-PE to assess

**Table 3.** Deep sequencing characterization of selected scFv against 6 different targets. Total number of Ion Torrent sequences, number of unique HCDR3 and percentage of the top 1 and the top 10 most abundant clones in each selection output are reported. Symbols indicate identity of HCDR3 clones in the different selection outputs. The characterization of the anti-CTBP2 rrpAb is shown, before and after negative sorting with the highly homologous CTBP1 protein

	Yeast output	Tot. # sequences	# unique HCDR3	Top1 (%)	Top10 (%)
CDK2	2+1	40169	373	17.8 <sup>S</sup>	57.7
	2+2	32126	132	16.2 <sup>S</sup>	76.7
CTBP2	2+1	35372	234	21.5 <sup>*</sup>	79.4
	2+2	39519	140	25.2 <sup>*</sup>	85.8
	1st negative	23652	284	42.7 <sup>+</sup>	72.2
	2nd negative	15532	204	57.7 <sup>+</sup>	82.5
	2+3	29654	96	70.5 <sup>+</sup>	96.1
MAP2K5	2+1	38454	460	6.9 <sup>#</sup>	26.8
	2+2	24491	224	18.2 <sup>#</sup>	67.7
PLAA	2+1	71343	139	38.3 <sup>o</sup>	87.2
	2+2	40646	74	50.0 <sup>o</sup>	95.8
SF3A1	2+1	16224	127	29.7 <sup>oo</sup>	81.5
	2+2	4553	136	51.3 <sup>oo</sup>	85.6
USP11	2+1	161329	397	11.4	46.3
	2+2	38191	211	5.7	45.2

scFv display levels. For the second sort, neutravidin-Alexa-Fluor633 was used to replace conjugated streptavidin and eliminate the chance to sort for streptavidin binders. Yeast clones showing both antigen binding (AlexaFluor633 positives) and display (PE positives) were sorted. The collected cells were grown at 30°C for 2 d and induced for the next round of sorting at 20°C for 16 h.

All the described flow cytometry experiments were performed using the FACSaria (Becton Dickinson).

After the second sort, single clones were screened for specificity using the specific target and the remaining unrelated antigens and conjugated streptavidin as negative controls. Affinity measurements were carried out according to published methods.<sup>30</sup> All experiments with single clones were performed in a 96-well format using the LSRII (Becton Dickinson) flow cytometer.

#### Production of polyclonal antibodies (yeast)

After one and 2 rounds of yeast sorting, the antibody outputs were subcloned into the yeast expression vector pDNL9-RMR (Rabbit Minibody for Recombination) to allow the expression and secretion of scFvs as rabbit Fc fusions<sup>37</sup> into the culture supernatant. pDNL6 and pDNL9-RMR were designed to have compatible ends to promote in vivo homologous recombination of the PCR amplified yeast display library into the yeast expression vector. YVH10 yeast cells (kindly provided by Prof. Dane Wittrup, Massachusetts Institute of Technology) were used for expression of polyclonals. Yeast antibody expression was performed following secretion protocols depicted in Wentz and Shusta,<sup>38</sup> using SGT as induction media. Culture supernatants were used directly as a reagent in Western blot.

#### ELISA screening of antibodies

ELISA was performed by coating 96-well ELISA plates (Nunc) with 10 ug/ml neutravidin diluted in 1x phosphate-buffered saline (PBS), and incubating the plate overnight at 4°C. Wells were blocked with 2% MPBS (1x PBS, 2% skim milk (w/

v)) and incubated with 10 ug/ml of biotinylated proteins. After a rinse step, yeast supernatants diluted 1:5 in 1xPBS were added to the wells, followed by HRP-conjugated goat anti-rabbit antibody (Santa Cruz). Immunocomplexes were revealed by adding TMB (Thermo) and reading the plate at 450 nm. The ELISA results were normalized after quantifying the rrpAb present in the supernatant by directly coating the yeast supernatant on a plate and detecting it with HRP-conjugated goat anti-rabbit antibody.

#### Immunoblotting on purified antigens and cell extracts

Biotinylated proteins were separated by SDS-PAGE (Invitrogen) and transferred to nitrocellulose membranes (GE Healthcare). The membranes were blocked with 2% MPBS for 1 h at room temperature, then incubated 1h with yeast supernatants, diluted 1:1 in 2% MPBS. After extensive washes, the membranes were incubated with AP-conjugated goat anti-rabbit antibody (Santa Cruz), and the immune-complexes revealed by the chromogenic substrate NBT/BCIP (Thermo).

HEK293 cell lysates were also used to test the polyclonal antibodies. Thirty ug were used as source of antigens in the immunoblot assay. The detection of immune complexes (specific antigen and the binding activity of the expressed rrpAbs) was obtained with HRP-conjugated goat anti-rabbit antibody (Santa Cruz) and chemiluminescence substrates (Pierce).

#### Deep sequencing: Sample preparation and data analysis

Sample preparation and sequence analysis were performed as described.<sup>27</sup> Briefly, plasmid DNA was recovered from the sorted selection outputs and amplified by a primer set designed for Ion Torrent sequencing. The amplicons were sequenced on an Ion 316 chip. The sequencing results were analyzed by the AbMining toolbox software package using the default settings for quality filtering. The identified HCDR3s were clustered at Hamming distance 1 and analyzed further in MS Excel.

The protein phylogeny was built by PhyML [v3.0\_360–500M<sup>39,40</sup>] from the ClustalW alignment of HCDR3 and

rendered with the rainbowtree tool (<http://www.hiv.lanl.gov/content/sequence/RAINBOWTREE/rainbowtree.html>). For the inferred phylogeny to reflect length variation, gapped sites were treated as an additional character state, rather than as missing information.

### Protein arrays

A total of 79 different proteins, including negative and positive controls, were printed on FAST slide (Whatman), at 15°C and 50% humidity using the Calligrapher arrayer (BioRad). Slides were first dried overnight and then saturated with blocking solution (4% nonfat dry milk in PBS 0.1% Tween20 (PBST) for 1 h at RT. The proper spotting of antigens on the substrate was assessed by PE-conjugated streptavidin (Life Technologies, 1:400 dilution). PrG purified anti-CDK2 rrpAb and commercial sc-163 (Santa Cruz) were used at 1 µg/mL. Incubations were performed in binding buffer (2% nonfat dry milk in PBST) for 1h at RT. Slides were washed 3 times for 10 min with PBST. Secondary antibody, Cy5 conjugated anti-rabbit IgG (Listarfish), was used at 1:200 dilution for 1 h at RT in binding buffer. Final washing was performed as previously described. Slides were dried and scanned with a ScanArray Gx (Perkin Elmer). The inter-slide

step normalization was performed with a rabbit IgG calibrator (inter-slide calibration).

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Funding

This work was supported by the National Institutes of Health [5U54DK093500–02 to ARMB]. The Structural Genomics Consortium, a registered charity (number 1097737) receives funds from the Canada Foundation for Innovation, Eli Lilly Canada, GlaxoSmithKline, the Ontario Ministry of Economic Development and Innovation, the Novartis Research Foundation, Pfizer, AbbVie, Bayer, Takeda, Janssen, Boehringer Ingelheim and the Wellcome Trust.

### References

- Nilsson P, Paavilainen L, Larsson K, Odling J, Sundberg M, Andersson AC, Kampf C, Persson A, Al-Khalili Szigyarto C, Ottosson J, et al. Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* 2005; 5:4327-37; PMID:16237735; <http://dx.doi.org/10.1002/pmic.200500072>
- Bordeaux J, Welsh A, Agarwal S, Killiam E, Baquero M, Hanna J, Anagnostou V, Rimm D. Antibody validation. *Biotechniques* 2010; 48:197-209; PMID:20359301; <http://dx.doi.org/10.2144/000113382>
- Saper CB. A guide to the perplexed on the specificity of antibodies. *J Histochem Cytochem* 2009; 57:1-5; PMID:18854594; <http://dx.doi.org/10.1369/jhc.2008.952770>
- Couchman JR. Commercial antibodies: the good, bad, and really ugly. *J Histochem Cytochem* 2009; 57:7-8; PMID:18854593; <http://dx.doi.org/10.1369/jhc.2008.952820>
- Jensen BC, Swigart PM, Simpson PC. Ten commercial antibodies for alpha-1-adrenergic receptor subtypes are nonspecific. *Naunyn Schmiedebergs Arch Pharmacol* 2009; 379:409-12; PMID:18989658; <http://dx.doi.org/10.1007/s00210-008-0368-6>
- Slaastad H, Wu W, Goullart L, Kanderova V, Tjonnfjord G, Stuchly J, Kalina T, Holm A, Lund-Johansen F. Multiplexed immuno-precipitation with 1725 commercially available antibodies to cellular proteins. *Proteomics* 2011; 11:4578-82; PMID:21956872; <http://dx.doi.org/10.1002/pmic.201000744>
- Blow N. The generation game. *Nature* 2007; 447:741-3; PMID:17554312; <http://dx.doi.org/10.1038/447741a>
- Haab BB, Dunham MJ, Brown PO. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol* 2001; 2:RESEARCH0004; PMID:11182887; <http://dx.doi.org/10.1186/gb-2001-2-2-research0004>
- Silman I, Katchalski E. Water-insoluble derivatives of enzymes, antigens, and antibodies. *Annu Rev Biochem* 1966; 35:873-908; <http://dx.doi.org/10.1146/annurev.bi.35.070166.004301>
- Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. GENCODE: the reference human genome annotation for the ENCODE project. *Genome Res* 2012; 22:1760-74; PMID:22955987; <http://dx.doi.org/10.1101/gr.135350.111>
- Lane L, Argoud-Puy G, Britan A, Cusin I, Duek PD, Evalet O, Gateau A, Gaudet P, Gleizes A, Masselot A, et al. neXtProt: a knowledge platform for human proteins. *Nucleic Acids Res* 2012; 40:D76-83; PMID:22139911; <http://dx.doi.org/10.1093/nar/gkr1179>
- Alm T, von Feilitzen K, Lundberg E, Sivertsson A, Uhlen M. A chromosome-centric analysis of antibodies directed toward the human proteome using Antibodypedia. *J Proteome Res* 2014; 13:1669-76; PMID:24533432; <http://dx.doi.org/10.1021/pr4011525>
- Voskuil J. Commercial antibodies and their validation. *F1000Res* 2014; 3:232; PMID:25324967; <http://dx.doi.org/10.12688/f1000research.4966.2>
- Edwards AM, Isserlin R, Bader GD, Frye SV, Willson TM, Yu S. Too many roads not taken. *Nature* 2011; 470:163-5; PMID:21307913; <http://dx.doi.org/10.1038/470163a>
- Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zvalhen M, Kampf C, Wester K, Hober S, et al. Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* 2010; 28:1248-50; PMID:21139605; <http://dx.doi.org/10.1038/nbt1210-1248>
- Ayris J, Woods T, Bradbury A, Pavlik P. High-throughput screening of single-chain antibodies using multiplexed flow cytometry. *J Proteome Res* 2007; 6:1072-82; PMID:17330944; <http://dx.doi.org/10.1021/pr0604108>
- Giovannoni L, Viti F, Zardi L, Neri D. Isolation of anti-angiogenesis antibodies from a large combinatorial repertoire by colony filter screening. *Nucleic Acids Res* 2001; 29:E27; <http://dx.doi.org/10.1093/nar/29.5.e27>
- Villa A, Trachsel E, Kaspar M, Schliemann C, Somavilla R, Rybak JN, Rosli C, Borsi L, Neri D. A high-affinity human monoclonal antibody specific to the alternatively spliced EDA domain of fibronectin efficiently targets tumor neo-vasculature in vivo. *Int J Cancer* 2008; 122:2405-13; PMID:18271006; <http://dx.doi.org/10.1002/ijc.23408>
- Klooster R, Eman MR, le Duc Q, Verheesen P, Verrips CT, Roovers RC, Post JA. Selection and characterization of KDEL-specific VHH antibody fragments and their application in the study of ER resident protein expression. *J Immunol Methods* 2009; 342:1-12; PMID:19041652; <http://dx.doi.org/10.1016/j.jim.2008.10.009>
- Gao J, Sidhu SS, Wells JA. Two-state selection of conformation-specific antibodies. *Proc Natl Acad Sci U S A* 2009; 106:3071-6; PMID:19208804; <http://dx.doi.org/10.1073/pnas.0812952106>
- Thomsen ND, Koerber JT, Wells JA. Structural snapshots reveal distinct mechanisms of procaspase-3 and -7 activation. *Proc Natl Acad Sci U S A* 2013; 110:8477-82; PMID:23650375; <http://dx.doi.org/10.1073/pnas.1306759110>
- Hu X, Kang S, Lefort C, Kim M, Jin MM. Combinatorial libraries against libraries for selecting neopeptide activation-specific antibodies. *Proc Natl Acad Sci U S A* 2010; 107:6252-7; PMID:20308586; <http://dx.doi.org/10.1073/pnas.0914358107>
- Schier R, McCall A, Adams GP, Marshall KW, Merritt H, Yim M, Crawford RS, Weiner LM, Marks C, Marks JD. Isolation of picomolar affinity anti-c-erbB-2 single-chain Fv by molecular evolution of the complementarity determining regions in the center of the antibody binding site. *J Mol Biol* 1996; 263:551-67; PMID:8918938; <http://dx.doi.org/10.1006/jmbi.1996.0598>
- Boder ET, Midelfort KS, Wittrup KD. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc Natl Acad Sci USA* 2000; 97:10701-5; PMID:10984501; <http://dx.doi.org/10.1073/pnas.170297297>
- Ferrara F, Naranjo LA, Kumar S, Gaiotto T, Mukundan H, Swanson B, Bradbury AR. Using phage and yeast display to select hundreds of monoclonal antibodies: application to antigen 85, a tuberculosis biomarker. *PLoS One* 2012; 7:e49535; PMID:23166701; <http://dx.doi.org/10.1371/journal.pone.0049535>
- de Wildt RM, Mundy CR, Gorick BD, Tomlinson IM. Antibody arrays for high-throughput screening of antibody-antigen interactions. *Nat Biotechnol* 2000; 18:989-94; PMID:10973222; <http://dx.doi.org/10.1038/79494>
- D'Angelo S, Glanville J, Ferrara F, Naranjo L, Gleasner CD, Shen X, Bradbury AR, Kiss C. The antibody mining toolbox: an open source tool for the rapid analysis of antibody repertoires. *MAbs* 2014; 6:160-72;

- PMID:24423623; <http://dx.doi.org/10.4161/mabs.27105>
28. Li SL, Liang SJ, Guo N, Wu AM, Fujita-Yamaguchi Y. Single-chain antibodies against human insulin-like growth factor I receptor: expression, purification, and effect on tumor growth. *Cancer Immunol Immunother* 2000; 49:243-52; PMID:10941907; <http://dx.doi.org/10.1007/s002620000115>
  29. Hjelm B, Forsstrom B, Lofblom J, Rockberg J, Uhlen M. Parallel immunizations of rabbits using the same antigen yield antibodies with similar, but not identical, epitopes. *PLoS One* 2012; 7:e45817; PMID:23284606; <http://dx.doi.org/10.1371/journal.pone.0045817>
  30. Feldhaus MJ, Siegel RW, Opresko LK, Coleman JR, Feldhaus JM, Yeung YA, Cochran JR, Heinzelman P, Colby D, Swers J, et al. Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat Biotechnol* 2003; 21:163-70; PMID:12536217
  31. D'Angelo S, Kumar S, Naranjo L, Ferrara F, Kiss C, Bradbury AR. From deep sequencing to actual clones. *Protein Eng Des Sel* 2014; 27:301-7; PMID:25183780; <http://dx.doi.org/10.1093/protein/gzu032>
  32. Sblattero D, Bradbury A. Exploiting recombination in single bacteria to make large phage antibody libraries. *Nat Biotechnol* 2000; 18:75-80; PMID:10625396; <http://dx.doi.org/10.1038/71958>
  33. Sarantopoulos S, Kao CY, Den W, Sharon J. A method for linking VL and VH region genes that allows bulk transfer between vectors for use in generating polyclonal IgG libraries. *J Immunol* 1994; 152:5344-51
  34. Sharon J, Sarantopoulos S, Den W, Kao CY, Baecher-Allan CM, Santora KE, Sompuram SR, Petersen-Mahrt S, Williams BR. Recombinant polyclonal antibody libraries. *Comb Chem High Throughput Screen* 2000; 3:185-96; <http://dx.doi.org/10.2174/1386207003331643>
  35. Marks JD, Bradbury A. Selection of human antibodies from phage display libraries. *Methods Mol Biol* 2004; 248:161-76
  36. Boder ET, Wittrup KD. Optimal screening of surface-displayed polypeptide libraries. *Biotechnol Prog* 1998; 14:55-62; PMID:10858036; <http://dx.doi.org/10.1021/bp970144q>
  37. Di Niro R, Ziller F, Florian F, Crovella S, Stebel M, Bestagno M, Burrone O, Bradbury AR, Secco P, Marzari R, et al. Construction of miniantibodies for the in vivo study of human autoimmune diseases in animal models. *BMC Biotechnol* 2007; 7:46-55; PMID:17678525; <http://dx.doi.org/10.1186/1472-6750-7-46>
  38. Wentz AE, Shusta EV. Enhanced secretion of heterologous proteins from yeast by overexpression of ribosomal subunit RPP0. *Biotechnol Prog* 2008; 24:748-56; PMID:18396911; <http://dx.doi.org/10.1021/bp070345m>
  39. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003; 52:696-704; <http://dx.doi.org/10.1080/10635150390235520>
  40. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010; 59:307-21; PMID:20525638; <http://dx.doi.org/10.1093/sysbio/syq010>