share or publish the new visualization on the web.

Trackster's use of data subsets to reduce analysis computation time is applicable to a wide set of genomic tools. For instance, genomic interval operations (such as intersect and subtract), transcript assembly and quantification, and human variation analysis (such as SNP calling) are compatible with Trackster's analysis approach. However, tools (such as some peak callers) that use data from many or all genomic regions to build a global model require additional support to work with Trackster. These tools must be run once in full to generate the model, and then the model can be stored in Galaxy and reused in Trackster. Transcript quantification in Cufflinks benefits from a global model, and Trackster makes use of it when it available. Alternatively, a tool (for example, a read mapper) may require all input data because it is not possible to identify, before runtime, a subset of input data needed to produce correct output in a particular genomic region. For such tools, dynamic filtering can be used to simulate running a tool using different parameters. In this approach, a tool's parameters are relaxed so that many potential outputs are produced and attribute values are attached to output data. Filtering can then be used to observe the data that would be produced for particular parameter values.

Visualization and data analysis tools are used in nearly all high-throughput sequencing experiments, yet too often they are not well integrated. Coupling visualization and analysis tools into a visual analysis environment where analysis output can be generated and visually assessed in real time is a powerful approach for computational science. Trackster provides an environment for interactive visual analysis that is widely applicable to many different high-throughput sequencing experiments. General visual analysis techniques that can be performed in Trackster include tool parameter-space visualization and exploration, systematic sweeps of parameter values and dynamic filtering. Trackster makes visual analysis possible for a wide variety of tools by leveraging the Galaxy framework, thereby tapping into the large collection of tools already integrated into Galaxy and providing a simple path for integrating additional tools into Trackster. This approach to tool integration enables popular, production-level tools, such as Cufflinks in our example, to be integrated into Trackster without modification to the tools themselves. In our experiment, Trackster's visual analysis features made it possible to use interactive visualization to improve Cufflinks' transcript assemblies via parameter-space exploration and to remove assembly artifacts using dynamic filtering. Trackster also supports collaborative visual analysis via web-based, fully functional shared visualizations that can be modified, extended, re-shared and published.

ACKNOWLEDGMENTS

Efforts of the Galaxy Team (E. Afgan, D. Baker, D. Blankenberg, N.C., J.G., G. Von Kuster, R. Lazarus, K. Li) were instrumental in making this work happen. This project was supported by American Recovery and Reinvestment Act (ARRA) funds through grant number HG005542 from the National Human Genome Research Institute, National Institutes of Health, as well as grants HG005133, HG004909 and HG006620 and National Science Foundation grant DBI 0543285. Additional funding is provided, in part, under a grant from the Pennsylvania Department of Health using Tobacco Settlement Funds. The Department specifically disclaims responsibility for any analyses, interpretations or conclusions.

Note: Supplementary information is available at http://www.nature.com/doifinder/10.1038/nbt.2404.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Jeremy Goecks[1], Nate Coraor[2], The Galaxy Team[3], Anton Nekrutenko[2] & James Taylor[1]

[1]Departments of Biology and Mathematics & Computer Science, Emory University, Atlanta, Georgia, USA. [2]Center for Comparative Genomics and Bioinformatics, Penn State University, University Park, Pennsylvania, USA. [3] The Galaxy Project, http://galaxyproject.org.
e-mail: james.taylor@emory.edu or anton@bx.psu.edu

1. Trapnell, C. *et al. Nat. Biotechnol.* **28**, 511–515 (2010).
2. Trapnell, C. *et al. Nat. Protoc.* **7**, 562–578 (2012).
3. Nielsen, C.B., Cantor, M., Dubchak, I., Gordon, D. & Wang, T. *Nat. Methods* **7**, S5–S15 (2010).
4. Rutherford, K. *et al. Bioinformatics* **16**, 944–945 (2000).
5. Kent, W.J. *Genome Res.* **12**, 996–1006 (2002).
6. Robinson, J.T. *et al. Nat. Biotechnol.* **29**, 24–26 (2011).
7. Fiume, M., Williams, V., Brook, A. & Brudno, M. *Bioinformatics* **26**, 1938–1944 (2010).
8. Jankun-Kelly, T.J. & Kwan-Liu, M. *IEEE Transactions on Visualization and Computer Graphics* **7**, 275–287 (2001).
9. Pretorius, A.J., Bray, M.A.P., Carpenter, A.E. & Ruddle, R.A. *IEEE Transactions on Visualization and Computer Graphics* **17**, 2402–2411 (2011).
10. Goecks, J., Nekrutenko, A., Taylor, J. & the Galaxy Team. *Genome Biology* **11**, R86 (2010).
11. Blankenberg, D. *et al. Curr. Protoc. Mol. Biol.* **89**, 19.10.1–19.10.21 (2010).
12. Ron, D. & Walter, P. *Nat. Rev. Mol. Cell Biol.* **8**, 519–529 (2007).
13. Walter, P. & Ron, D. *Science* **334**, 1081–1086 (2011).
14. Mori, K. *J. Biochem.* **146**, 743–750 (2009).
15. Calfon, M. *et al. Nature* **415**, 92–96 (2002).
16. Yanagitani, K., Kimata, Y., Kadokura, H. & Kohno, K. *Science* **331**, 586–589 (2011).
17. Guo, H., Ingolia, N.T., Weissman, J.S. & Bartel, D.P. *Nature* **466**, 835–840 (2010).
18. Reid, D.W. & Nicchitta, C.V. *J. Biol. Chem.* **287**, 5518–5527 (2012).

Proteomics-directed cloning of circulating antiviral human monoclonal antibodies

To the Editor:

In the May 2012 issue of your journal, we described an approach that uses proteomics and next-generation sequencing to identify antigen-specific antibodies directly from the serum of immunized animals, and we applied it to clone circulating antibodies to five different antigens from the serum of rabbits and mice[1]. Many technologies for isolating antibodies have been developed and applied to gain insight into the specific human antibody response to various pathogens, but none have directly addressed the serological response at the proteomic level[2–5]. In addition, recent evidence from a study conducted in mice suggests that not all memory B cells contribute directly to the serological response to a pathogen[6]. Thus, a method capable of interrogating the composition and complexity of the circulating antibody repertoire elicited to specific vaccines or pathogens is necessary.

Here, we report the application of our strategy[1] to the identification and cloning of high-affinity, antigen-specific human monoclonal antibodies directly from plasma of a donor vaccinated against hepatitis B virus (HBV). We also clone potent neutralizing human monoclonal antibodies against human cytomegalovirus (HCMV) from a healthy, naturally infected individual.

A key first step of our approach is the affinity purification of polyclonal material from a serum or plasma sample to enrich for monoclonal antibodies with desired functional properties. Earlier we have enriched for antibodies with high specific activity in various immunoassays such as enzyme-linked immunosorbent assay (ELISA), western blotting, immunofluorescence, flow cytometry and immunohistochemistry[1]. But the affinity purification can be tailored to select

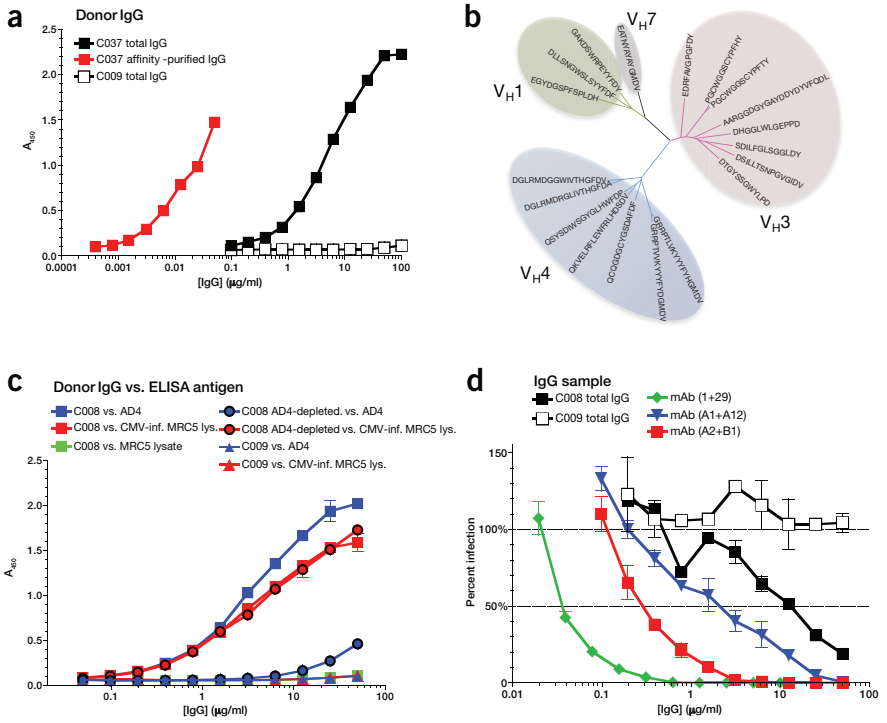

Figure 1 Affinity purification and identification of HBsAg-specific IgG from an HBV-vaccinated donor and of CMV-neutralizing IgG from a CMV-positive donor. (**a**) HBsAg ELISA titration of total IgG from HBV vaccinated (C037) and unvaccinated (C009) donors, and affinity-purified IgG from the vaccinated donor (red). (**b**) Phylogenetic tree (generated by aligning all 19 heavy chains using neighbor joining method for multiple sequence alignment by CLC Bio's genomic workbench software) of all 19 heavy chain variable sequences identified from affinity-purified IgG C037. Sequences that use V_H1, V_H3, V_H4 and V_H7 gene families are in green, pink, blue and gray, respectively. Scale represents number of substitutions per 100 residues (display was generated using Fig Tree; http:// tree.bio.ed.ac.uk/software/figtree/). (**c**) Antigen-specific ELISA binding of protein G-purified IgG from HCMV-positive (C008) and CMV-negative (C009) donors, compared with binding activity of protein G-purified C008 IgG after depletion of AD4 binding IgG following AD4 affinity purification. Antigens for which ELISA binding was tested were recombinant AD4 and lysate of HCMV-infected and uninfected MRC5 cells. (**d**) Neutralizing activity of total plasma IgG from HCMV-positive (C008) and HCMV-negative (C009) donors and three neutralizing monoclonal antibodies isolated from C008. In **a**, **c** and **d**, samples were tested in duplicate; error bars indicate the deviation of each value from the mean.

specifically for high-affinity interactions, potent neutralization or activity in other types of assays. Once the desired properties have been enriched, liquid chromatography tandem mass spectrometry (LC-MS/MS) is used to identify the monoclonal components of the purified fraction by matching to a reference database of antibody variable regions (V regions) produced by next-generation sequencing of the B cell immunoglobulin repertoire of the immunized animal.

In contrast to work in hyperimmunized laboratory animals, translating our approach to the isolation of circulating human antibodies faces several potential challenges. For instance, in our earlier work[1] we isolated B cells from the spleen of sacrificed animals, but in humans, accessing immunological organs that are highly enriched in B cells, such as the spleen, lymph nodes or even bone marrow, may not be feasible. Furthermore, in most cases humans are not hyperimmunized or maintained in a controlled environment, as would be the case for laboratory animals. In some cases, naturally exposed or convalescent donors might be the preferred or only source of polyclonal serum and B cell–derived genetic material, and therefore high-titer or adequate quantity of material could be limiting. In such cases, the immunogen is not known and thus the appropriate antigen for isolating the desired antibodies needs to be characterized before screening.

As a proof of principle of our methodology (**Supplementary Fig. 1**), we first sought to isolate immunogen-specific antibodies from HBV-vaccinated donors. Chronic HBV infection affects >350 million people, a considerable percentage of whom succumb to hepatic failure or have a high risk of developing hepatocellular carcinoma[7]. HBV small surface antigen (HBsAg) is the main clinical marker indicating acute or chronic HBV infection, and the recombinant version of this protein has been widely used in vaccine formulations since 1986 (ref. 8). To investigate the serological immune response in HBV vaccine recipients, we screened plasma samples from volunteers who were recently immunized during an in-house HBV vaccination program for reactivity to recombinant HBsAg by ELISA. Of the six volunteers who had received their second HBV immunization dose 7 d before blood collection, donor C037 showed the strongest plasma γ immunoglobulin (IgG) reactivity to the antigen (step A in **Supplementary Fig. 1**, **Supplementary Fig. 2**). The vaccine-specific response of this donor was confirmed by ELISA using total IgG purified with Sepharose-conjugated protein G, in which the half-maximal effective concentration (EC_{50}) was <10 μg ml^{-1} (**Fig. 1a**).

To enrich for vaccine-specific IgG from plasma for analysis by LC-MS/MS, we used a purification method based on magnetic microbead separation (step B in **Supplementary Fig. 1**) as described in **Supplementary Methods**. Elutions containing affinity-purified polyclonal antibody were neutralized and tested for antigen-specific binding activity by ELISA, showing an enrichment of ~150-fold when compared with the protein G–purified starting material (**Fig. 1a**).

To maximize identification of the V regions of the purified polyclonal mixture using LC-MS/MS, Fc regions were removed by digestion with IdeS (step C in **Supplementary Fig. 1**), an IgG-degrading cysteine proteinase from *Streptococcus pyogenes*[9], and the dimeric form of antigen-binding fragments ($F(ab')_2$) was gel purified (**Supplementary Fig. 3**), digested in-gel separately using chymotrypsin, LysC or trypsin (step D in **Supplementary Fig. 1**), and then extracted for analysis by LC-MS/MS as described in **Supplementary Methods**. Identification of V-region peptides was maximized by analysis of samples from each digest in duplicate using a 72-min gradient for six LC-MS/MS runs per purified antibody resulting in ~29,000 spectra per run.

To map MS/MS spectra to V-region sequences, a custom sequence reference database of antibody V regions from the donor is necessary (step E in **Supplementary Fig. 1**). Influenza vaccine–specific plasmablasts and memory B cells in circulation peak between days 7 and 14 (ref. 5). Thus, we reasoned that 7 d after HBV vaccine immunization would be an ideal time to collect memory B cells or plasmablasts for building a reference

Table 1 Monoclonal antibodies to HBsAg isolated from HBV-vaccinated donor C037

Clone	H-CDR3	L-CDR3	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_d (M)
C037(1+49)	DLLSNGWSLSYYFDF	QQYHTWPFT	1.69×10^5	4.27×10^{-4}	2.53×10^{-9}
C037(3+53)	AARGGDGYGAYDDYDYVFQDL	QQYSAYPIT	4.03×10^4	2.09×10^{-5}	5.17×10^{-10}
C037(5+55)	GRRPTVVKYYYFYDGMDV	QQSASSPRT	3.06×10^4	5.65×10^{-4}	1.85×10^{-8}
C037(7+63)	DGLRMDGGWIVTHGFDV	QHRSNWPPLFT	9.80×10^4	1.99×10^{-3}	2.03×10^{-8}
C037(9+61)	QCQGDGCYGSDAFDF	QQRSNWPMYT	ND	ND	ND
C037(11+63)	DGLRMDRGLIVTHGFDA	QHRSNWPPLFT	ND	ND	ND
C037(14+75)	DTGYSSGWYLPD	QQYYGAPIT	1.98×10^5	4.32×10^{-4}	2.18×10^{-9}
C037(15+91)	EDRFAVGPGFDY	GTWDTTLSAGV	8.95×10^4	4.10×10^{-4}	4.58×10^{-9}
C037(17+55)	GRRPTLVKYYYFYHGMDV	QQSASSPRT	4.20×10^4	4.99×10^{-4}	1.18×10^{-8}
C037(19+89)	QKVELRFLEWFRLHDSDV	ASWDDSLKGVL	7.80×10^4	4.71×10^{-4}	6.04×10^{-9}
C037(21+57)	GAKDSWRPEYYFDY	QQYNSFAFN	3.21×10^5	1.62×10^{-3}	5.04×10^{-9}
C037(25+105)	PGCWGGSCYPFTY	QSYDNSLSGHNYV	3.63×10^5	3.63×10^{-4}	2.08×10^{-9}
C037(27+103)	SDILFGLSGGLDY	QVWDGISNRVV	4.23×10^5	5.08×10^{-4}	1.20×10^{-9}
C037(29+107)	DHGGLWLGEPPD	ASFTPRYTWV	ND	ND	ND
C037(31+105)	PGCWGGSCYPFHY	QSYDNSLSGHNYV	5.63×10^5	7.45×10^{-4}	1.32×10^{-9}
C037(34+87)	QSYSDIWSGYGLHWFDP	HNYGTIPGT	1.77×10^4	4.50×10^{-4}	2.54×10^{-8}
C037(35+97)	DSILLTSNPGVGIDV	QVWDTTTDHMI	4.61×10^5	1.19×10^{-3}	2.57×10^{-9}
C037(37+95)	EATNYAYAYGMDV	QSYDGSLSRSTV	ND	ND	ND
C037(44+87)	EGYDGSPFSPLDH	HNYGTIPGT	ND	ND	ND

Each antibody clone is identified by heavy chain number plus light chain number in parentheses after donor identification. Heavy and light chain CDR3 sequences and binding kinetics constants measured by Biacore T200 are shown for HBsAg-specific monoclonal antibodies with unique heavy chains. Kinetics constants of monoclonal antibodies with lower relative binding as determined by ELISA titration (**Supplementary Fig. 3**) were not measured. H-CDR3, heavy chain complementarity determining region 3; L-CDR3, light chain complementarity determining region 3; k_a, association rate constant or on-rate; k_d, dissociation rate constant or off-rate; K_d, equilibrium dissociation constant or affinity; ND, not determined.

database. We isolated memory B cells from peripheral blood mononuclear cells of donor C037 using magnetic cell sorting to generate a cDNA library. In parallel, we also generated a cDNA library from antigen-enriched B cells, which were sorted using biotinylated HBsAg immobilized to magnetic streptavidin beads. The cDNA libraries from these cell pools were amplified by PCR using human γ-, κ- and λ-chain-specific fusion primers (**Supplementary Table 1**), followed by sequencing of heavy and light chain V regions using a Roche 454 FLX+ instrument.

After two independent sequencing runs of each library, we obtained 882,000 and 862,000 sequences that cover the entire V region from libraries of memory B cells and antigen-specific cells, respectively. Among these sequences, 132,373 unique γ-chain CDR3 sequences came from the memory B cell library but only 7,258 unique γ-chain CDR3 sequences came from the HBsAg-enriched B cell library. A smaller CDR3 sequence diversity was observed in the antigen-enriched B cell pool compared with the memory B cell pool, indicating as we expected that the antigen-enriched B cell repertoire is more restricted. We combined the pool of identified chains from both libraries into a single analysis to maximize the discovery of antibody V regions. Using this database, we confidently mapped 3,305 spectra to V-region sequences from donor C037 with a false discovery rate of 2%.

We next identified high-confidence V-region sequences containing unique CDR3s (see **Supplementary Fig. 4a,b** for example of a high-confidence heavy chain identified from HBV-vaccinated donor C037) using described methods[1] (**Supplementary Methods**). In this manner, we selected, synthesized and cloned 24 γ-, 20 κ- and 10 λ-chain V-region sequences (see **Supplementary Table 2** for percentage coverage of CDR3 and V region, and total number of mapped peptides for chains that yielded functional antibodies; see **Supplementary Tables 3** and **4** for identification of peptide spectral matches corresponding to V-region sequences of heavy chain 3 and light chain 53 from donor C037).

To produce recombinant monoclonal antibodies, heavy and light chains were combinatorially paired and transiently expressed in HEK293 cells in a 96-well format as described[1]. We screened the recombinant antibodies by ELISA and identified 37 unique heavy and light chain pairs (comprising 19 γ, 12 κ and 9 λ chains) that bound specifically to recombinant HBsAg. This result indicates that our criteria for selecting antibody chains were effective: 19 of 24 γ chains, 12 of 20 κ chains and 9 of 10 λ chains contributed to antigen-specific monoclonal antibodies when paired.

The V(D)J usage (**Supplementary Table 5**) and phylogenetic analysis of the positive heavy chain V regions (**Fig. 1b**) demonstrate that the identified antibody clones are of diverse B cell origin. Moreover, the high frequency of somatic hypermutation (**Supplementary Table 5**) in both the heavy and the light chains suggests that these antibodies have undergone substantial affinity maturation. By competition ELISA, we determined that these antibodies could be classified into at least four distinct epitope groups (data not shown). Furthermore, using overlapping linear peptides that span the entire length of the HBsAg *adw* subtype sequence, we mapped the epitope of one antibody (monoclonal antibody (15+91)) to within 10 residues and another (monoclonal antibody (17+69)) to 20 residues within the neutralization-sensitive, antigenic loop (**Supplementary Fig. 5**). The region of amino acids recognized by these monoclonal antibodies overlaps epitopes of described neutralizing antibodies[10]. Taken together, these results suggest that the monoclonal antibodies we obtained likely react against relatively diverse epitopes within HBsAg.

To further characterize these antigen-specific antibodies, we assessed the relative binding activity of all HBsAg-specific monoclonal antibodies by ELISA (**Supplementary Fig. 6**). Next, antibodies with higher relative binding activity were subjected to affinity measurements using a Biacore T200 instrument (GE Healthcare). We found 23 heavy and light chain pairs with

Table 2 Monoclonal antibodies to AD4 (HCMV gB domain) isolated from HCMV-positive donor C008

Clone	H-CDR3	L-CDR3	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_d (M)	CMV neutralization IC_{50}
C008(1+29)	EIIQWPRRWFDP	ATWDGTLTAGV	9.08×10^5	6.30×10^{-4}	6.95×10^{-10}	<0.04 μg ml^{-1}
C008(3+39)	DKDYGMVAATPDAFDI	GTWDSSLKTAIF	1.92×10^5	1.09×10^{-3}	5.67×10^{-9}	–
C008(5+31)	DPGDDSGGNSGLDY	GTWDSSLTAGV	1.13×10^6	3.69×10^{-3}	3.27×10^{-9}	–
C008(A1+A12)	ARRGESGAYGSSDY	QQYYSYPFT	1.96×10^5	4.71×10^{-4}	2.40×10^{-9}	<3 μg ml^{-1}
C008(A2+B1)	RYTEGDSVWYFDV	HQYKHWPRT	3.63×10^5	2.34×10^{-3}	6.44×10^{-9}	<0.3 μg ml^{-1}

Each antibody clone is identified by heavy chain number plus light chain number in parentheses after donor identification. CDR3 sequences (for heavy and light chains), binding kinetics constants measured by Biacore T200 and IC_{50} by *in vitro* HCMV microneutralization assay are presented. Dash indicates that neutralizing activity was not detected at highest IgG concentration tested (50 μg ml^{-1}).

affinities with K_d values ranging from 517 pM to 613 nM (**Table 1**, monoclonal antibodies with unique heavy chains; **Supplementary Table 6**, additional monoclonal antibodies). The highest-affinity monoclonal antibody was C037(3+53) (binding kinetics in **Supplementary Fig. 7a**). The antibodies we isolated using our stringent purification method could represent only a fraction of the total HBsAg-specific antibodies found in circulation owing to purification conditions that are too stringent for lower-affinity antibodies; limited B cell source for generation of the next-generation sequencing reference database for mass spectrometry; and/or the limit of detection of the mass spectrometry instruments. Nevertheless, the diversity of high-affinity antibodies identified by our proteomic approach from a single donor shows that HBV vaccination elicits a humoral response that triggers expansion of a diverse B cell response. This leads to circulating antibodies with high affinity for the HBV vaccine.

We next sought to identify neutralizing antibodies elicited against a pathogen from the plasma of a naturally infected or exposed individual. We focused on HCMV owing to its high prevalence in healthy adults and its clinical importance for immunocompromised individuals, solid organ transplant recipients and newborns, who may contract the virus congenitally[11].

In contrast to the HBV vaccine experiments in which we isolated a wide spectrum of antigen-specific antibodies, we designed our purification procedure for HCMV to enrich specifically for neutralizing antibodies. Antibody titer to envelope glycoprotein B (gB) in HCMV-infected individuals is correlated with neutralizing activity against the virus[12], and antibodies specific to a single discontinuous external region of gB known as antigenic domain 4 (AD4; residues 121–132 and 344–438, **Supplementary Fig. 8**) can effectively inhibit HCMV infection of a variety of cell types including fibroblasts and endothelial, epithelial and dendritic cells[13]. We therefore used the AD4 domain of gB to purify antibodies with the aim of identifying neutralizing human IgG against HCMV.

Healthy volunteers were screened for HCMV-neutralizing activity in their plasma using an *in vitro* microneutralization assay (**Supplementary Fig. 9a**)[14]. Plasma from one donor, C008, exhibited half-maximal inhibitory concentration (IC_{50}) titer of ~1:500 dilution (**Supplementary Fig. 9b**). Total plasma IgG from donor C008 was then purified using protein G, and showed very strong binding reactivity to AD4 and to total HCMV-infected cell lysate by ELISA (**Fig. 1c**). IgG purified from an HCMV-negative donor (with no detectable anti-HCMV IgG), C009, showed no reactivity in either assay (**Fig. 1c**).

To identify antibodies to HCMV from donor C008, we purified antibodies using recombinant AD4 and confirmed by ELISA that most of the AD4-specific activity was adsorbed but that binding activity to other HCMV proteins was unaffected (**Fig. 1c**). AD4-purified antibodies were analyzed by LC-MS/MS, and high-confidence heavy and light chain V regions (**Supplementary Table 7**) were identified (details in **Supplementary Methods**; see **Supplementary Fig. 4c,d** for example of high-confidence heavy chain identified from donor C008, and **Supplementary Tables 8** and **9** for identification of V-region peptide spectral matches corresponding to V regions of heavy chain 1 and light chain 29) and combinatorially expressed as described above. Fifteen heavy and light chain pairs (comprising five γ, three κ and ten λ chains) showed AD4-binding specificity by ELISA (data not shown). Five of ten γ chains, three of four κ chains and ten of ten λ chains selected led to productive antibodies, suggesting that the selection method based on coverage worked.

The 15 anti-AD4 antibodies were further characterized for binding affinity by Biacore T200 measurements and tested for *in vitro* neutralization activity against HCMV. These antibodies had affinities ranging from 278 pM to 7.76 nM (**Table 2**, binding kinetics constants of antibodies with unique heavy chains; **Supplementary Table 10**, V(D)J gene usage and frequency of somatic hypermutations; **Supplementary Table 11**, list of all other antibodies to AD4). Seven heavy and light chain pairs (with heavy chains 1, A1 and A2) showed neutralizing activity with IC_{50} values ranging from 0.04 μg ml^{-1} to 25 μg ml^{-1} (**Table 2** , **Supplementary Table 11** and **Fig. 1d**, neutralization curves of most potent neutralizers for each γ chain). The *in vitro* neutralizing activity of monoclonal antibodies C008(1+29), C008(1+27) and C008(A2+B1), with IC_{50} values of <0.04 μg ml^{-1}, <0.07 μg ml^{-1} and <0.3 μg ml^{-1}, respectively, is among the most potent reported so far for antibodies that inhibit HCMV infection of fibroblasts[13,15], with K_d values of 695 pM, 549 pM and 6.44 nM, respectively (**Table 2** and **Supplementary Table 11**).

In summary, we present two separate proof-of-concept cases to demonstrate our proteomics-based antibody discovery approach to isolate and clone antiviral monoclonal antibodies directly from human circulation. In vaccine research, antibody titers measured by ELISA are often used to identify and/or map neutralizing epitopes and to establish correlates of immune protection[16]. More recently, high-throughput single-cell cloning strategies have been applied to investigate the memory and plasma cell pool in vaccinated or naturally exposed donors to reflect a specific humoral response[5,17]. However, none of these techniques allow direct interrogation of the circulating antibody population. In the first proof of concept, we took advantage of the serological response of an HBV vaccine recipient to rapidly obtain monoclonal antibodies to the recombinant HBsAg vaccine. Some of these vaccine-specific antibodies have high affinities that are comparable to those of recently identified HBV-neutralizing antibodies[2], and two of them bind to described neutralizing epitopes[10]. Our method could be used to track changes in the vaccine-specific circulating antibody repertoire over time and examine their correlation with the evolving functional characteristics of individual antibody components.

Characterization of biochemical properties and biological activity of isolated antibodies during purification is a critical step that helps focus the proteomics process on the identification of monoclonal antibodies with desired functional properties. We demonstrate this principle in the second proof of concept, namely the isolation of HCMV-neutralizing human monoclonal antibodies from a naturally infected donor. To accomplish this task, we first screened for donors' plasma with potent neutralizing activity *in vitro* (**Supplementary Fig. 9a**). Using plasma from one such donor and a purification strategy restricted to the use of AD4 domain of gB from HCMV[13], we isolated monoclonal antibodies to HCMV with high affinities (up to 278 pM) and potent neutralization activity (IC_{50} values as low as 0.04 μg ml^{-1}). In future experiments, one could envision discovering additional neutralizing monoclonal antibodies from the same donor by using additional components of gB[13] or other HCMV glycoproteins[18] for the affinity purification step. In this way, one could reconstitute a combined pool of potent neutralizing, fully human monoclonal antibodies. Considering that pooled HCMV hyperimmune globulin preparations are still the only available antibody-based HCMV-specific therapy, a neutralizing mixture made by recombinant human monoclonal antibodies would provide an improved clinical tool for passive immunotherapy against HCMV.

Manipulating purification conditions upfront facilitates the isolation and identification of antigen-specific human monoclonal antibodies with various biophysical or biochemical characteristics such as acid resistance, high heat tolerance, specific association and/or dissociation rates, ability to compete with a specific ligand, binding to a protein domain or a combination of any of these properties. We tracked antigen-specific binding activity to monitor enrichment of the desired polyclonal fraction (**Fig. 1a,c**). Such enrichment before mass spectrometry analysis enhances the probability of reconstituting functional heavy and light chain matches by combinatorial pairing, and we speculate that some of the matches may correspond to cognate pairs.

Finally, functionally validated antibody sequences identified using this approach can be used as a guide for further mining of additional clonally related antibody chains from the next-generation sequencing database generated from the same donor[19] (**Supplementary Fig. 10**). These additional antibody chains could have been missed owing to their very low affinity or very low abundance in serum, or because they were encoded by memory B cells, that did not contribute to the serological response at the time the blood sample was drawn. In addition, we cannot rule out that other specific antibodies enriched through purification could not be identified because they were expressed only by plasma cells in the bone marrow or other lymphoid organs and thus their sequences may be absent in the cDNA sequence databases from circulating B cells.

Whether the goal is to identify a broad antibody pool against a whole protein antigen or a more restricted set of neutralizing antibodies against a smaller domain, these results demonstrate that our proteomics approach is applicable in humans and thus may be useful to address questions in humoral immunity and facilitate the development of human antibody therapeutics.

Note: Supplementary information is available at http://www.nature.com/doifinder/10.1038/nbt.2406.

ACKNOWLEDGMENTS

We thank J. Fisher for biotinylation of antigen; D. Moore-Lai, T. Manganaro, T. Palazzola and K. Riley for antibody expression and purification; J. Knott and J. MacNeill for peptide synthesis; C. Manning and M. Nelson for high-content analysis; A. Funicella for plasmid purification, C. Reeves for DNA sequencing of expression constructs; K. Lee and A. Moritz for insightful discussion on mass spectrometry; and S. Martin and E. Savinelli for coordinating donor blood collection. We thank M. Mach for allowing us to adopt the HCMV gB structural image in our manuscript. We thank R. Matthews, T. Singleton and D. Comb for designing graphics. We thank M. Comb, T. Sulahian, K. Huynh, P. Hornbeck, C. Hoffman and L. Morrison for insightful comments and discussion on the manuscript. Finally, we are very grateful to all the volunteers who donated blood, without whom this project would not have been feasible.

AUTHOR CONTRIBUTIONS

S.S., S.A.B., W.C.C. and R.D.P developed the methodology, designed experiments, analyzed data and wrote the manuscript. S.A.B. did bioinformatic analyses. S.S., S.A.B., L.P., J.G.B., R.K.R., X.Z., J.S.W. and S.M.S. did experiments.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available at http://www.nature.com/doifinder/10.1038/nbt.2406.

Shuji Sato[1,2], Sean A Beausoleil[1,2], Lana Popova[1], Jason G Beaudet[1], Ravi K Ramenani[1], Xiaowu Zhang[1], James S Wieler[1], Sandra M Schieferl[1], Wan Cheung Cheung[1] & Roberto D Polakiewicz[1]

[1]Cell Signaling Technology, Danvers, Massachusetts, USA. [2]These authors contributed equally to this work. Correspondence should be addressed to W.C.C. (gcheung@cellsignal.com) or R.D.P. (rpolakiewicz@cellsignal.com).

1. Cheung, W.C. *et al. Nat. Biotechnol.* **30**, 447–452 (2012).
2. Jin, A. *et al. Nat. Med.* **15**, 1088–1092 (2009).
3. Meijer, P.J. *et al. J. Mol. Biol.* **358**, 764–772 (2006).
4. Schmaljohn, C., Cui, Y., Kerby, S., Pennock, D. & Spik, K. *Virology* **258**, 189–200 (1999).
5. Wrammert, J. *et al. Nature* **453**, 667–671 (2008).
6. Purtha, W.E., Tedder, T.F., Johnson, S., Bhattacharya, D. & Diamond, M.S. *J. Exp. Med.* **208**, 2599–2606 (2011).
7. Zuckerman, J.N. & Zuckerman, A.J. *J. Infect.* **41**, 130–136 (2000).
8. Hilleman, M.R. *Infection* **15**, 3–7 (1987).
9. von Pawel-Rammingen, U., Johansson, B.P. & Bjorck, L. *EMBO J.* **21**, 1607–1615 (2002).
10. Tajiri, K. *et al. Antiviral Res.* **87**, 40–49 (2010).
11. Staras, S.A. *et al. Clin. Infect. Dis.* **43**, 1143–1151 (2006).
12. Marshall, G.S., Rabalais, G.P., Stout, G.G. & Waldeyer, S.L. *J. Infect. Dis.* **165**, 381–384 (1992).
13. Potzsch, S. *et al. PLoS Pathog.* **7**, e1002172 (2011).
14. Abai, A.M., Smith, L.R. & Wloch, M.K. *J. Immunol. Methods* **322**, 82–93 (2007).
15. Macagno, A. *et al. J. Virol.* **84**, 1005–1013 (2010).
16. Barrette, R.W., Urbonas, J. & Silbart, L.K. *Clin. Vaccine Immunol.* **13**, 802–805 (2006).
17. Scheid, J.F. *et al. Nature* **458**, 636–640 (2009).
18. Ryckman, B.J. *et al. J. Virol.* **82**, 60–70 (2008).
19. Wu, X. *et al. Science* **333**, 1593–1602 (2011).

Analyzing the association of SCNA boundaries with replication timing

To the Editor:

A paper by De and Michor[1] published in last December's issue claimed that DNA replication timing and long-range DNA interactions can predict mutational landscapes of cancer genomes. We would like to draw readers' attention to a statistical weakness in their analysis, which in our opinion negates one of the main claims of the article.

The paper, which presents integrative analysis of a database of somatic copy-number alterations (SCNAs)[2], DNA replication timing[3] and three-dimensional conformation data[4] (Hi-C), argues that the data provide evidence that the formation of SCNAs is governed by a DNA replication–driven model, in which alterations occur opportunistically when two active replication forks at the boundaries of the nascent SCNA are in close three-dimensional proximity. The empirical evidence presented has two essential components: first, an increase in the density