



Functional interrogation and mining of natively paired human $V_H:V_L$ antibody repertoires

Bo Wang^{1,12}, Brandon J DeKosky^{2-4,12}, Morgan R Timm², Jiwon Lee¹, Erica Normandin², John Misasi², Rui Kong², Jonathan R McDaniel¹, George Delidakis¹ , Kendra E Leigh² , Thomas Niezold², Chang W Choi², Elise G Viox², Ahmed Fahad⁴, Alberto Cagigi², Aurélie Ploquin², Kwanyee Leung², Eun Sung Yang², Wing-Pui Kong², William N Voss¹, Aaron G Schmidt⁵, M Anthony Moody^{6,7}, David R Ambrozak², Amy R Henry², Farida Laboune², Julie E Ledgerwood², Barney S Graham², Mark Connors⁸, Daniel C Douek², Nancy J Sullivan², Andrew D Ellington^{9,10}, John R Mascola² & George Georgioui^{1,9-11}

We present a technology to screen millions of B cells for natively paired human antibody repertoires. Libraries of natively paired, variable region heavy and light ($V_H:V_L$) amplicons are expressed in a yeast display platform that is optimized for human Fab surface expression. Using our method we identify HIV-1 broadly neutralizing antibodies (bNAbs) from an HIV-1 slow progressor and high-affinity neutralizing antibodies against Ebola virus glycoprotein and influenza hemagglutinin.

The human B-cell receptor repertoire constitutes an invaluable resource for discovery of therapeutic antibodies^{1,2}. Cloning from individual B cells obtained via immortalization and expansion *in vitro*, or from single B cells obtained by limiting dilution or fluorescence-activated cell sorting (FACS), has been used extensively to discover anti-infective antibodies, including broadly neutralizing antibodies (bNAbs) to HIV-1 and influenza^{3,4}. In parallel, over the last 25 years the screening of combinatorial libraries generated by random pairing of amplified V_H and V_L genes from human B cells has yielded numerous antibodies, leading to dozens of experimental or approved drug products^{5,6}. However, single-cell cloning is time- and resource-intensive, and is therefore limited to analysis of a small fraction of the human antibody repertoire^{7,8}, whereas combinatorial library screening has the capacity to interrogate

antibody function from millions of B cells. However, the non-cognate pairing of V_H and V_L sequences in these libraries frequently gives rise to antibodies with lower selectivity and inferior biophysical properties compared to authentic human immunoglobulins^{9,10}.

We report a technology for large-scale functional interrogation of the natively paired $V_H:V_L$ antibody repertoire (Fig. 1). Because V_H and V_L genes are encoded by separate mRNA transcripts, they are first physically linked into a single amplicon for subsequent cloning into an expression vector. V_H and V_L linkage is accomplished by a two-step single-cell emulsion lysis and oligo-dT capture of V_H and V_L mRNAs from the same B cell, followed by a second reverse transcription (RT) and overlap-extension (OE)-PCR step to create contiguous $V_H:V_L$ amplicons^{11,12}. In these amplicons the V_H and V_L genes are joined through a linker designed to enable one-step sub-cloning into a yeast Fab surface-expression vector, whereby the V_H and V_L genes are transcribed from a galactose-inducible bidirectional promoter with C_{H1} (human IgG1 isotype) and C_L (human κ and $\lambda 2$ isotypes) at the C terminus of the V_H and V_L , respectively (Fig. 1a,b, Supplementary Fig. 1 and Supplementary Table 1).

Human antibodies often express poorly in microbial hosts¹³, and while expression efficiency in yeast is substantially higher relative to *Escherichia coli* or phage¹³⁻¹⁵, examination of a panel of 13 previously reported human influenza hemagglutinin (HA)-specific antibodies revealed that only 7/13 antibodies (53%) bound antigen when displayed on yeast (Fig. 1c and Supplementary Fig. 2)¹⁶. Consistent with earlier reports¹⁷, co-expression of protein disulfide isomerase (PDI) increased display efficiency, as monitored by two-color flow cytometry, to 10/13 antibodies (Fig. 1c). In addition to PDI expression, the enhanced dimerization of heavy and light chains via fusion to C-terminal leucine-zipper domains resulted in the display of the full set of 13/13 human anti-HA antibodies (Fig. 1c and Supplementary Fig. 2)¹⁸. We then tested the display efficiency of three other human antibodies (two anti-Ebola virus (EBOV) antibodies: c13c6 and KZ52, and the anti-HIV-1 bNAb N123-VRC34.01 that targets the HIV-1 fusion peptide¹⁹). All three antibodies displayed efficiently and were shown to bind selectively to their respective antigens in the optimized system (Supplementary Fig. 2). Extensive earlier studies demonstrated that yeast display enables interrogation of the human antibody repertoire based on affinities or off-rates, for epitope specificity, and for other properties including stability¹⁵ (Fig. 1d). Clones of interest can then be expressed either as Fab or as IgG for detailed functional and biochemical assays (Fig. 1d).

We used this approach to analyze the antibody repertoire of an individual 6 d after immunization with an experimental EBOV vaccine²⁰.

¹Department of Chemical Engineering, The University of Texas at Austin, Austin, Texas, USA. ²Vaccine Research Center, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland, USA. ³Department of Chemical & Petroleum Engineering, The University of Kansas, Lawrence, Kansas, USA. ⁴Department of Pharmaceutical Chemistry, The University of Kansas, Lawrence, Kansas, USA. ⁵Laboratory of Molecular Medicine, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA. ⁶Duke Human Vaccine Institute, Duke University Medical School, Durham, North Carolina, USA. ⁷Department of Pediatrics, Duke University Medical School, Durham, North Carolina, USA. ⁸National Institute of Allergy and Infectious Diseases, Bethesda, Maryland, USA. ⁹Center for Systems and Synthetic Biology, The University of Texas at Austin, Austin, Texas, USA. ¹⁰Department of Molecular Biosciences, The University of Texas at Austin, Austin, Texas, USA. ¹¹Department of Bioengineering, The University of Texas at Austin, Austin, Texas, USA. ¹²These authors contributed equally to this work. Correspondence should be addressed to J.R.Mascola (jmascola@mail.nih.gov) or G.G. (gg@che.utexas.edu).

Received 14 June 2017; accepted 6 December 2017; published online 8 January 2018; doi:10.1038/nbt.4052

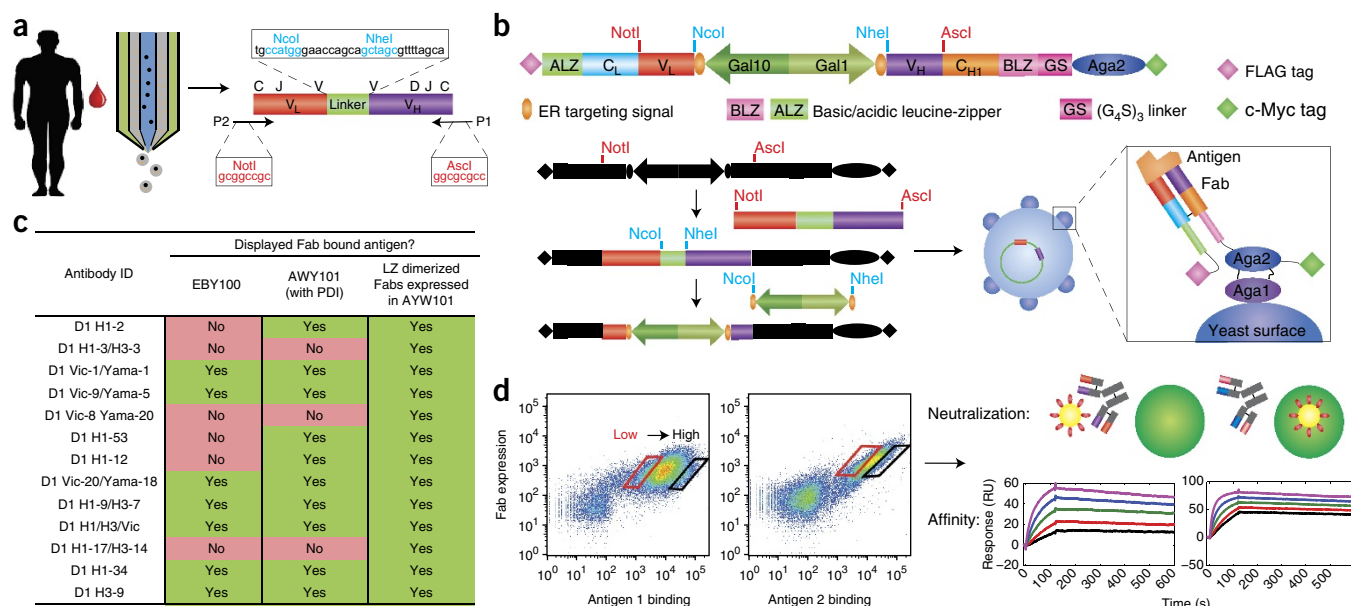


Figure 1 High-throughput cloning, yeast display, and functional analysis of the human naturally paired $V_H:V_L$ antibody repertoire. **(a)** Using an axisymmetric flow focusing device, peripheral blood B cells are compartmentalized inside microdroplets and lysed, and single-cell mRNA is captured as overlap extension PCR template to produce V_H and V_L cDNAs joined by a 32-nucleotide linker sequence containing *NcoI* and *NheI* restriction sites. **(b)** Naturally paired $V_H:V_L$ amplicon libraries are subcloned *en masse* into a Fab expression vector with a galactose-inducible bidirectional promoter *Gal1/Gal10* for transcription of C_{H1} - V_H and C_{L1} - V_L , along with expression tags (c-Myc, FLAG, respectively) and leucine-zipper (LZ) dimerization domains. **(c)** Display characteristics for a panel of 13 anti-HA antibodies in yeast strains EBY100, AWY101 that overexpresses PDI, and AWY101 with LZ-forced dimerization. **(d)** Sequential rounds of FACS under increasingly stringent conditions (e.g., lower antigen concentrations, co-incubation with competitor antibodies) are used to bin libraries within various windows of affinity. Antibodies are recovered from sorted yeast, and expressed and characterized.

This peak plasmablast $V_H:V_L$ repertoire was displayed in yeast, and cells were analyzed and sorted for binding to EBOV mucin-like domain deleted glycoprotein ($GP_{\Delta\text{muc}}$). We used high-throughput sequencing (HTS) to track antibody lineages throughout the screening process. Of 1,189 unique CDRH3:CDRL3 nucleotide clusters obtained from 5,002 plasmablasts after highly stringent sequence quality filtering, 828 were verified as cloned and displayed in the system using HTS (70% overall efficiency for library construction and display). As expected for the peak post-vaccination plasmablast response, an appreciable (6%) fraction of repertoire-expressing yeast cells in the pre-sort library bound to antigen, and antigen-specific clones were highly enriched after the third round of sorting (Fig. 2a and Supplementary Fig. 3). Single-colony analysis of yeast yielded seven antibody lineages that bound to $GP_{\Delta\text{muc}}$ (EBOV.YD.01–EBOV.YD.04, EBOV.YD.09–EBOV.YD.11; Supplementary Table 2 and Supplementary Fig. 4). Comparison of HTS data sets for the pre-sort library and the sorted library after three rounds of screening revealed that all seven clones isolated above had been enriched by ≥ 120 -fold. Four of these antibodies were randomly selected (EBOV.YD.01–EBOV.YD.04) and expressed as IgG1s in HEK293 cells, then digested to generate Fabs, which were shown to bind $GP_{\Delta\text{muc}}$ with nM affinities by biolayer interferometry (BLI) (Supplementary Fig. 5 and Supplementary Table 2). All four antibodies blocked infection by EBOV GP-pseudotyped lentiviral particles, with neutralization ranging from 55% to 99% at 10 $\mu\text{g}/\text{ml}$ (Fig. 2b and Supplementary Fig. 6). Competition assays revealed that these antibodies targeted distinct non-overlapping epitopes (Fig. 2c). EBOV.YD.03 competed with the well-characterized neutralizing antibody KZ52, indicating that it binds an epitope similar to antibodies generated during natural infection²¹ (Fig. 2c).

HTS surveillance of antibody clonal prevalence during screening enabled us to retrieve other antibodies that were enriched across

rounds, but not identified by single-colony picking. We synthesized four additional antibody lineages that had been enriched more than eightfold in HTS data sets after multiple rounds of FACS. Three out of four antibodies identified by HTS bound to $GP_{\Delta\text{muc}}$ with single-digit nM K_D (EBOV.YD.06–EBOV.YD.08; Supplementary Fig. 5 and Supplementary Table 2) while another clone, EBOV.YD.05, bound weakly ($\sim 3 \mu\text{M}$ K_D as a Fab).

We then applied this yeast display technology to assess an antibody lineage in an HIV-1-infected donor. Kong and co-workers recently identified N123-VRC34, an HIV-1 bNAb lineage that binds to the fusion peptide (FP)¹⁹. We interrogated the antibody repertoire of 1.42 million peripheral B cells from this donor (N123) and amplified the $V_H:V_L$ repertoire with a unique human FR1 primer set that was also supplemented with lineage-specific primers. Only the inclusion of lineage-specific primers enabled us to successfully recover the N123-VRC34 lineage, which contains several reported FR1 mutations (Supplementary Tables 3 and 4) and is extremely rare at this time point within the donor (roughly 0.003% of all B cells¹⁹). Yeast libraries were sorted using an epitope protein scaffold containing the eight terminal AA of the fusion peptide (VRC34-epitope scaffold-FP-APC) and a version of the scaffold alone without the fusion peptide (VRC34-epitope scaffold-KO-PE)¹⁹. As with EBOV antibody libraries, the HIV-1 antibody libraries were highly enriched for FP-specific clones by the third round of sorting (Fig. 2d and Supplementary Fig. 7).

HTS revealed that after three rounds of screening, VRC34-lineage antibodies far outcompeted other antibody lineages, constituting 98.7% of high-quality sequences and suggesting that the VRC34 lineage dominated the FP-specific repertoire in this donor. Three prevalent $V_H:V_L$ clones were expressed and the respective Fabs were shown to bind to the HIV-1 fusion peptide probe with high affinity (Fig. 2f and Supplementary Fig. 8). To further “bin” FP-binding clones based

BRIEF COMMUNICATIONS

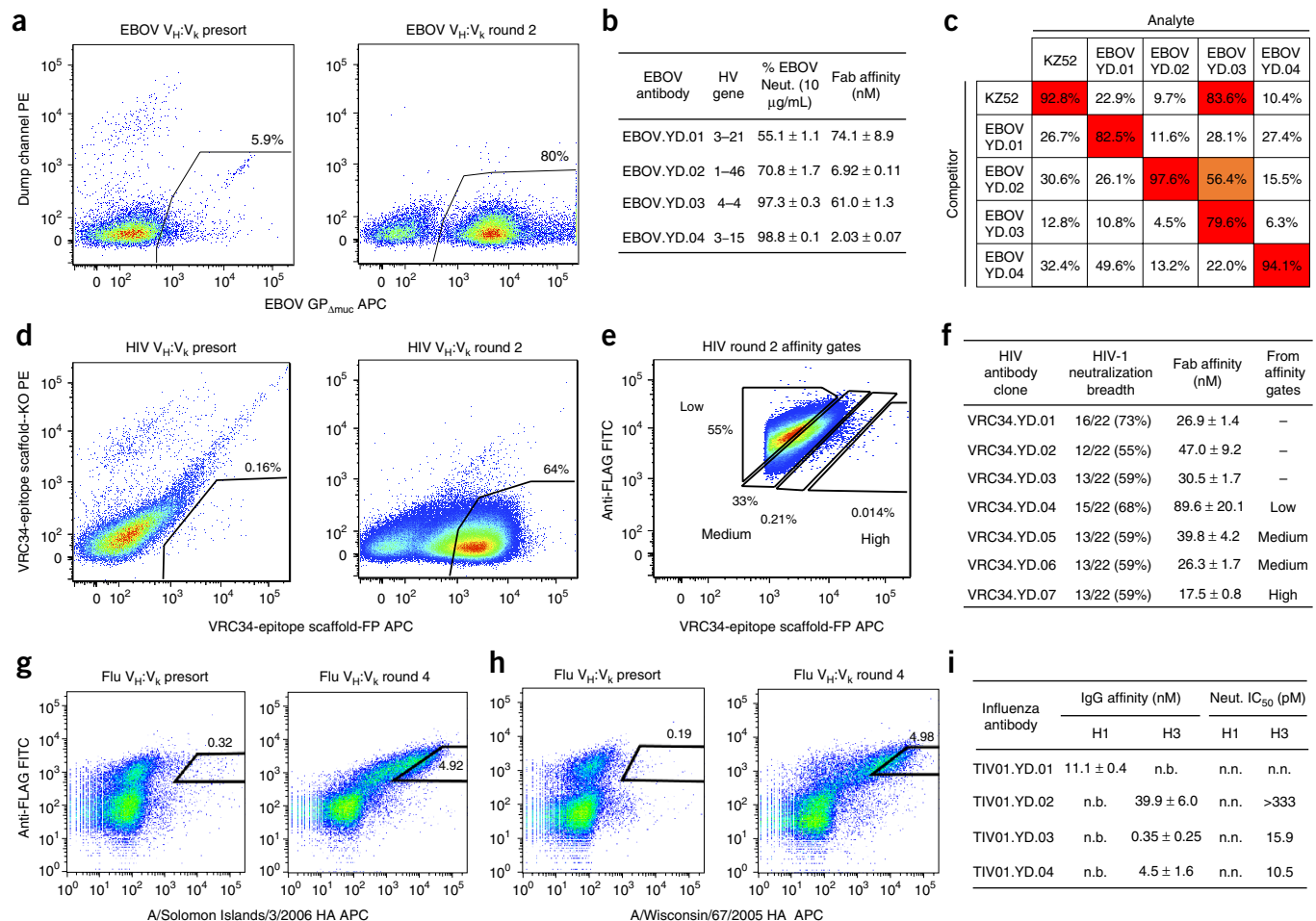


Figure 2 Examples of natively paired antibody repertoire analysis and functional characterization. **(a)** FACS analysis of the natively paired heavy:light antibody repertoire from 5,002 peripheral plasmablasts isolated from an EBOV vaccinee and screened for binding to EBOV GP $_{\Delta muc}$. **(b)** Neutralization and affinity of GP $_{\Delta muc}$ antibodies randomly selected after the third round of sorting. **(c)** Competition analysis of anti-GP $_{\Delta muc}$ antibodies from **b** and the EBOV-neutralizing antibody KZ52. **(d)** FACS analysis of the yeast display repertoire from 1.42×10^6 peripheral B cells for binding to HIV-1 FP probe. **(e)** Diagonal FACS gates were used to bin the sorted (round 2) repertoire based on affinity. **(f)** K_D values and HIV-1 neutralization breadth of purified antibodies following screening. **(g)** The steady-state peripheral B-cell repertoire was interrogated with A/Solomon Islands/3/2006 H1 HA. **(h)** The same B-cell repertoire was mined with A/Wisconsin/67/2005 H3 HA. **(i)** Affinity and neutralization of anti-HA antibodies isolated after the fourth round of antigen screening. n.b., no binding; n.n., no neutralization. Affinity values and EBOV neutralization are reported as average \pm s.d. from three technical replicates. Influenza neutralization is reported as average from three technical replicates.

on affinity, we gated the yeast population during the third round of sorting by increasing fluorescence intensity to FP; Fabs of four clones restricted to a high-, medium-, or low-affinity gated population had K_D values consistent with their respective FACS profile (Fig. 2e,f and Supplementary Fig. 8). In total, seven unique VRC34 lineage antibodies were identified and all were broadly neutralizing (Supplementary Fig. 9). Notably, three double-nucleotide changes within a codon were observed that resulted in nonsynonymous amino acid substitutions, which are highly unlikely to have resulted from PCR or other artifacts and thus likely arose from somatic hypermutation in the donor, suggesting site-specific selection *in vivo* (Supplementary Fig. 10). These results suggest that genetic-lineage targeting coupled with yeast display can be useful for antibody discovery against HIV-1 or other difficult pathogens for which bNAbs are reported to have specific genetic requirements^{22,23}.

Finally, we constructed a paired $V_H:V_L$ library from 12 million peripheral B cells harvested 270 d after immunization with seasonal, trivalent inactivated influenza vaccine (IIV3) (ref. 24) when HA-specific B cells occur at a frequency of $\sim 0.01\%$ ²⁵. We isolated single yeast

colonies after one round of sorting for antibody-expressing cells and four rounds of screening with Group 1 HA (18 clones) and separately, Group 2 HA (16 clones) included in IIV3 (Group 1: H1 from A/Solomon Islands/3/2006; Group 2: H3 from A/Wisconsin/67/2005); decreasing concentrations of antigen were used across rounds to increase selection stringency (Fig. 2g,h and Supplementary Fig. 11). Of these, 15/34 (44%) colonies encoded four unique antibody lineages to HA (one targeting H1 and three targeting H3) that bound to recombinant HAs with affinities ranging from 0.35 to 39.9 nM when expressed as IgGs (Fig. 2i, Supplementary Fig. 12 and Supplementary Table 5); an additional 7/34 colonies (21%) recognized HAs but with lower affinities. Two of the four antibodies reported in Figure 2i neutralized influenza with picomolar inhibitory concentrations (IC₅₀) (Supplementary Fig. 13).

The display of a properly folded, functional antibody repertoire in yeast constitutes a renewable resource for the isolation of human antibodies and also for repeated analyses of the antibody response based on properties such as affinity, epitope coverage (e.g., by sorting in the presence of competitor antibodies), and stability¹⁵. Yeast

surface display has been reported to have a lower expression bias relative to other microbial display technologies^{13–15}, and the yeast display optimization reported here further ensured bona fide expression of human antibody repertoires. We expect native $V_H:V_L$ antibodies to show superior selectivity and biophysical properties compared to randomly paired V_H and V_L antibodies isolated using other display platforms^{9,10}. Native antibody libraries displayed on yeast can be screened for antigens that bind to B-cell surface ligands (e.g., sialic acid²⁶ or CR2 (ref. 27)) and are therefore not suitable for single-B-cell sorting, and also can be used to discover antibodies targeting insoluble antigens, including membrane proteins^{28–30}.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the [online version of the paper](#).

ACKNOWLEDGMENTS

We gratefully acknowledge B. Hartman for assistance with figures, N. Doria-Rose, G. Ippolito, and M. Kanekiyo for advice and guidance, S. Lucas, and K. Zhou for help with experiments, S. Darko for assistance with data processing, and E. Shusta and S. Harrison for kindly sharing reagents and research tools. This work was funded in part by the intramural research program of the Vaccine Research Center, NIAID, NIH, NIH grant DP5OD023118-01 to B.J.D., NIH grant 5R21CA191239-01 to A.D.E., Leidos Biomedical Research Inc. contract 15X219, DTRA contract HDTRA1-12-C-0105, and NIH grant 1R56AI106006 to G.G.

AUTHOR CONTRIBUTIONS

B.W., B.J.D., J.R. Mascola, and G.G. conceived the study and designed the experiments. B.W. and B.J.D. conducted the experiments with help from M.R.T., J.L., E.N., J.M., R.K., J.R. McDaniel, G.D., K.E.L., T.N., C.W.C., E.G.V., A.F., A.C., A.P., K.L., E.S.Y., W.-P.K., W.N.V., A.G.S., M.A.M., D.R.A., A.R.H., F.L., J.E.L., B.S.G., M.C., and D.C.D. N.J.S., A.D.E., J.R. Mascola, and G.G. supervised the study. B.W., B.J.D., J.R. Mascola, and G.G. wrote the manuscript with input from all authors.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the [online version of the paper](#).

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Chan, A.C. & Carter, P.J. *Nat. Rev. Immunol.* **10**, 301–316 (2010).
2. Brekke, O.H. & Sandlie, I. *Nat. Rev. Drug Discov.* **2**, 52–62 (2003).
3. Corti, D. & Lanzavecchia, A. *Annu. Rev. Immunol.* **31**, 705–742 (2013).
4. Burton, D.R. & Hangartner, L. *Annu. Rev. Immunol.* **34**, 635–659 (2016).
5. Bradbury, A.R.M., Sidhu, S., Dübel, S. & McCafferty, J. *Nat. Biotechnol.* **29**, 245–254 (2011).
6. Eckler, D.M., Jones, S.D. & Levine, H.L. *MAbs* **7**, 9–14 (2015).
7. Wilson, P.C. & Andrews, S.F. *Nat. Rev. Immunol.* **12**, 709–719 (2012).
8. Georgiou, G. *et al. Nat. Biotechnol.* **32**, 158–168 (2014).
9. Jayaram, N., Bhowmick, P. & Martin, A.C.R. *Protein Eng. Des. Sel.* **25**, 523–529 (2012).
10. Ponsel, D., Neugebauer, J., Ladetzki-Baehs, K. & Tissot, K. *Molecules* **16**, 3675–3700 (2011).
11. DeKosky, B.J. *et al. Nat. Med.* **21**, 86–91 (2015).
12. McDaniel, J.R., DeKosky, B.J., Tanno, H., Ellington, A.D. & Georgiou, G. *Nat. Protoc.* **11**, 429–442 (2016).
13. Spadiut, O., Capone, S., Krainer, F., Glieder, A. & Herwig, C. *Trends Biotechnol.* **32**, 54–60 (2014).
14. Bowley, D.R., Labrijn, A.F., Zwick, M.B. & Burton, D.R. *Protein Eng. Des. Sel.* **20**, 81–90 (2007).
15. Feldhaus, M.J. *et al. Nat. Biotechnol.* **21**, 163–170 (2003).
16. Lee, J. *et al. Nat. Med.* **22**, 1456–1464 (2016).
17. Wentz, A.E. & Shusta, E.V. *Appl. Environ. Microbiol.* **73**, 1189–1198 (2007).
18. Ojima-Kato, T. *et al. Protein Eng. Des. Sel.* **29**, 149–157 (2016).
19. Kong, R. *et al. Science* **352**, 828–833 (2016).
20. Stanley, D.A. *et al. Nat. Med.* **20**, 1126–1129 (2014).
21. Maruyama, T. *et al. J. Virol.* **73**, 6024–6030 (1999).
22. Tian, M. *et al. Cell* **166**, 1471–1484. e18 (2016).
23. Joyce, M.G. *et al. Cell* **166**, 609–623 (2016).
24. Moody, M.A. *et al. PLoS One* **6**, e25797 (2011).
25. Pinna, D., Corti, D., Jarrossay, D., Sallusto, F. & Lanzavecchia, A. *Eur. J. Immunol.* **39**, 1260–1270 (2009).
26. Whittle, J.R. *et al. J. Virol.* **88**, 4047–4057 (2014).
27. Kanekiyo, M. *et al. Cell* **162**, 1090–1100 (2015).
28. Tillotson, B.J., Cho, Y.K. & Shusta, E.V. *Methods* **60**, 27–37 (2013).
29. Wang, X.X., Cho, Y.K. & Shusta, E.V. *Nat. Methods* **4**, 143–145 (2007).
30. Fang, Y., Chu, T.H., Ackerman, M.E. & Griswold, K.E. *MAbs* **9**, 1253–1261 (2017).

ONLINE METHODS

Strain and media. The yeast strain AWY101 (MAT α AGA1::GAL1-AGA1::URA3 PDI1::GAPDH-PDI1::LEU2 *ura3-52 trp1 leu2 Δ 1 his3 Δ 200 pep4::HIS3 prb1 Δ 1.6R can1 GAL*) (kind gift from Eric Shusta, University of Wisconsin-Madison) was used for library construction and screening. EBY100 (MAT α AGA1::GAL1-AGA1::URA3 *ura3-52 trp1 leu2 Δ 1 his3 Δ 200 pep4::HIS3 prb1 Δ 1.6R can1 GAL*) was used for initial native human antibody display. Yeast cells were maintained in YPD medium (20 g/l dextrose, 20 g/l peptone, and 10 g/l yeast extract); after library transformation, yeast cells were maintained in SDCAA medium (20 g/l dextrose, 6.7 g/l yeast nitrogen base, 5 g/l casamino acids, 8.56 g/l NaH₂PO₄·H₂O, and 10.2 g/l Na₂HPO₄·7H₂O). SGDCAA medium (SDCAA with 20 g/l galactose, 2 g/l dextrose) was used for library induction.

Antigens and antibodies. Recombinant Ebola virus glycoprotein with the mucin-like domain deleted (GP Δ muc), and HIV-1 fusion peptide probe (VRC34-epitope scaffold-FP) and knockout scaffold probe (VRC34-epitope scaffold-KO) were produced as described previously^{19,31,32}. Proteins were biotinylated and conjugated with streptavidin-APC (GP Δ muc and VRC34-epitope scaffold-FP) or streptavidin-PE (VRC34-epitope scaffold-KO) (Thermo Fisher Scientific), respectively. Recombinant hemagglutinins (A/California/07/2009, A/Solomon Islands/3/2006, and A/Wisconsin/67/2005) were produced as before³³ or acquired from BEI Resources (A/Victoria/210/2009, and B/Brisbane/60/2008) and were biotinylated using an EZ-Link Sulfo-NHS-LC-Biotin kit (Thermo Fisher Scientific). Anti-FLAG fluorescein isothiocyanate (FITC) antibody was purchased from Sigma-Aldrich (clone M2).

Optimization of native human antibody display. Vectors encoding previously reported anti-influenza virus HA monoclonal antibodies with or without leucine zipper domains were transformed into EBY100 or AWY101 using a Frozen-EZ Yeast Transformation II kit (Zymo Research)¹⁶. After culturing in SDCAA to an OD₆₀₀ of 2 at 30 °C, Fab surface expression was induced by transferring cells to SGDCAA medium at OD₆₀₀ of 0.5. After 2 d of induction at 20 °C, 10⁶ cells were collected and washed twice with PBS + 0.5% BSA + 2 mM EDTA, and incubated with 100 nM biotinylated hemagglutinin at room temperature for 30 min, followed by staining with 2 μ g/ml anti-FLAG FITC and 2 μ g/ml streptavidin-APC at 4 °C for 15 min. Cells were washed twice with ice-cold PBS + 0.5% BSA + 2 mM EDTA and analyzed on a FACS Aria II (BD Biosciences).

Analysis of anti-EBOV GP Δ muc (c13c6, KZ52) and anti-HIV-1 FP (VRC34.01) antibodies was performed as above^{19,34,35}, except 23 nM GP Δ muc-APC or 50 nM VRC34-epitope scaffold-FP-APC and 50 nM VRC34-epitope scaffold-KO-PE, respectively, were used for antigen staining.

Generation of natively paired V_H:V_L from peripheral B cells, library construction, yeast display and FACS screening. Human subject protocols were approved by the NIAID Institutional Review Board. For anti-EBOV GP Δ muc antibody isolation, peripheral blood mononuclear cells (PBMCs) were isolated from a healthy human volunteer after immunization with a phase 1 Ebola GP vaccine (NCT02408913)²⁰. The volunteer was first immunized with chimpanzee-derived replication-defective adenovirus encoding EBOV GP, then boosted 30 weeks and 5 d later with modified vaccinia Ankara encoding EBOV GP. Previous studies showed that plasmablasts in peripheral blood peak around 6 d post-boost immunization³⁶. 10 ml of blood were collected and PBMCs isolated using Ficoll-Paque PLUS (GE Healthcare) 6 d post-boost. PBMCs were stained with a multi-color flow cytometry panel consisting of fluorophore-labeled antibodies against CD3 (Brilliant Violet 421, clone SP34-2, BD Biosciences), CD19 (PE-Cy7, clone HIB19, BD Biosciences), CD4 (Brilliant Violet 421, clone OKT4, BioLegend), CD8a (Brilliant Violet 421, clone RPA-T8, BioLegend), CD14 (Brilliant Violet 421, clone M5E2, BioLegend), CD20 (Brilliant Violet 605, clone 2H7, BioLegend), CD27 (Brilliant Violet 711, clone O323, BioLegend), and CD38 (Alexa Fluor 680, clone OKT10, custom-conjugated at the Vaccine Research Center, NIAID), and 7-aminoactinomycin D (7-AAD, Thermo Fisher Scientific) to exclude dead cells. 5,002 CD3⁺CD4⁺CD8⁺CD14⁺CD19⁺CD20⁺CD27⁺CD38⁺ plasmablasts were isolated using a FACS Aria sorter (BD Biosciences) and subsequently used for emulsion V_H:V_L overlap extension RT-PCR.

For HIV-1 FP antibody isolation, PBMCs were collected from donor N123 on June 22, 2009. This donor is a chronically HIV-1-infected individual enrolled in NIAID under a clinical protocol approved by the NIAID Institutional Review Board^{37,38}. This donor was diagnosed with HIV-1 in 2000. After more than nine years of infection, this donor showed a CD4 T cell count of 463 cells/ml and a plasma HIV-1 viral load of 4,920 RNA copies/ml. This donor was not on antiretroviral treatment. 1.42 \times 10⁶ peripheral B cells were isolated from 25 million PBMCs using a human B-cell selection kit (Stemcell Technologies).

For the isolation of influenza HA-specific antibodies, a healthy donor was vaccinated with a trivalent inactivated influenza vaccine (IIV3: A/Solomon Islands/3/2006, A/Wisconsin/67/2005, B/Malaysia/2508/2004)²⁴. Subsequently, 270 d after vaccination, 1.2 \times 10⁷ B cells were isolated from blood leukapheresis using a human pan B-cell isolation kit (Miltenyi Biotec).

We used a flow-focusing nozzle to rapidly compartmentalize B cells in single-cell emulsion droplets, followed by single-B-cell lysis inside droplets and single-cell mRNA capture with oligo(dT)-coated magnetic beads^{11,12}. Overlap extension RT-PCR was then performed to link heavy and light chains using a Superscript III RT-PCR kit (Thermo Fisher Scientific)^{11,12}. We included NcoI and NheI restriction sites in the linker region of the overlap-extension RT-PCR primers that link V_H and V_L into an ~850 bp amplicon (**Supplementary Table 3**). For HIV-1 experiments, we also included additional primers specific to the VRC34 lineage (**Supplementary Table 4**). For library construction, 100 ng of V_H:V_L cDNA was amplified under the following conditions with AccuPrime Pfx DNA polymerase (Thermo Fisher Scientific) to introduce NotI and AscI sites, respectively (**Supplementary Table 1**): 2 min initial denaturation at 95 °C, denaturation at 95 °C for 20 s for 20 cycles, annealing at 60 °C for 20 s and extension at 68 °C for 60 s, final extension at 68 °C for 5 min. The DNA product was digested and ligated into pCT-VHVL-K1 (for V_H:V_K libraries) and pCT-VHVL-L1 (for V_H:V_L libraries), and transformed into electrocompetent *E. coli* for library cloning *en masse*. Plasmid DNA encoding V_H:V_L libraries was miniprep, digested with NcoI and NheI, ligated with the bidirectional promoter, and transformed into *E. coli* again. The final library DNA was miniprep, then amplified using the transformation primers (**Supplementary Table 1**) to generate library inserts with homologous ends to NotI and AscI double-digested vectors, and then inserts were co-transformed into electrocompetent AWY101 together with the digested vectors to generate libraries via yeast homologous recombination³⁹. Library sizes for EBOV, HIV-1 and flu repertoires were EBOV_K: 2 \times 10⁷, EBOV_L: 10⁷, HIV-1_K: 10⁷, HIV-1_L: 9 \times 10⁶, flu_K: 7 \times 10⁷, and flu_L: 3 \times 10⁷, respectively, as determined by colony counting.

For library screening, natively paired human V_H:V_L libraries were displayed on yeast by growing cells resuspended in SGDCAA medium at 20 °C for 2 d to induce Fab expression. The fraction of Fab-expressing cells 2 d post induction were consistent with previous reports for yeast displaying naïve human scFv¹⁵. For EBOV vaccinee and HIV-1 donor libraries, three rounds of sorting were performed against GP Δ muc, or VRC34-epitope scaffold-FP-APC and VRC34-epitope scaffold-KO-PE, respectively. The VRC34-epitope scaffold-FP was designed to present the FP in an optimal conformation and provide a glycan in a similar context as that presented by the native HIV-1 trimer¹⁹. In the first round of screening, at least tenfold coverage in yeast clones relative to library size were labeled with 2 μ g/ml anti-FLAG-FITC and either (i) 23 nM GP Δ muc-APC for isolating of EBOV GP Δ muc-specific antibodies, or (ii) 50 nM VRC34-epitope scaffold-FP-APC and 50 nM VRC34-epitope scaffold-KO-PE for the isolation of HIV-1 FP-specific antibodies. For EBOV antibody libraries, the PE channel was also included to correct for yeast autofluorescence. Cells were stained at room temperature for 30 min and washed twice with ice-cold PBS + 0.5% BSA + 2 mM EDTA, then analyzed by FACS. FITC⁺APC⁺PE⁻ cells were selected and recovered in SDCAA medium at 30 °C. Subsequent screening rounds were performed similarly, except that for the EBOV GP antibody library, at least 5 \times 10⁵ cells were screened in rounds 2 and 3, and for HIV-1 FP antibody library, at least 10⁷ cells were screened in rounds 2 and 3. Affinity binning of anti-EBOV GP Δ muc and anti-HIV-1 FP antibody repertoires was performed similarly as described⁴⁰.

Influenza HA-specific antibodies were isolated following five total rounds of sorting (two rounds of MACS (magnetic-activated cell sorting) and three rounds of FACS enrichment for binding to fluorescent HAs) as follows. For the

first round, at least tenfold coverage in yeast clones relative to library size were labeled with 2 µg/ml anti-FLAG-FITC at room temperature for 30 min. After washing, cells were labeled with anti-FITC microbeads (Miltenyi Biotec) at 4 °C for 15 min, and Fab-expressing cells were selected by MACS. For the second round of sorting, cells were labeled with 1 µM biotinylated recombinant HA (H1 A/Solomon Islands/3/2006, H3 A/Wisconsin/67/2005), and then selected with streptavidin microbeads (Miltenyi Biotec) using MACS as previously described⁴¹. Subsequently the library was screened using three rounds of FACS by labeling 5×10^6 cells with 2 µg/ml anti-FLAG-FITC together with 1 µM HA (first FACS round), 200 nM HA (second FACS round), or 40 nM HA (third FACS round), at room temperature for 30 min, followed by incubation with 2 µg/ml streptavidin-APC at 4 °C for 15 min before sorting.

Sequencing of the natively paired antibody repertoire from B cells was performed as previously described⁴². Bioinformatic analysis was performed as previously described⁴².

Recovery and expression of antibody clones from enriched libraries. Yeast cells from the final round of sorting were plated on SDCAA plates. A minimum of ten colonies were selected following the last round of each screening campaign. Colony PCR was performed on heavy and light chain variable regions of each clone. Clones were sequenced, and the unique antibodies were named as *project_name.YD.unique_clone_number*. For antibody expression, restriction sites were incorporated for insertion into the VRC8400 IgG1, and Igκ or Igλ expression vectors (for anti-EBOV GP_{Δmuc} and HIV-1 FP antibodies), or Gibson assembly was used to clone the variable regions into modified pcDNA3.4 IgG1, and Igκ or Igλ vectors (for anti-HA antibodies). Expi293 cells were co-transfected with heavy- and light-chain plasmids for each antibody, and secreted antibodies were purified on a Protein A column⁴³. Fabs were produced by digestion of IgG1 with Lys-C Protease (Thermo Fisher Scientific) and separated from Fc using Protein A or Protein G columns.

For EBOV vaccinee libraries, the population of sorted FITC⁺ Fab-expressing yeast cells in the first round of FACS were recovered in SDCAA medium at 30 °C. Plasmid DNA was extracted using high-efficiency yeast plasmid recovery protocols as reported previously⁴⁴ and V_H genes in sorted libraries were PCR-amplified using primers that targeted the yeast expression plasmid vector backbone (2YDrec_heavy_Vfor_MSrev1: TCTCGTGGGCTCGGAG ATGTGTATAAGACAGACAGNNNNCTGTTATTGCTAGCGTTTATAGCA, and 2YDrec_huIgH_Crev_MSfor1: TCGTCGGCAGCGTCAGATGTGTATAAGA GACAGNNNNAAGGCGCGCTGTACTTGC). Libraries were prepped by an additional round of PCR-based primer extension and Illumina adaptor addition to incorporate unique DNA barcodes for each sample, and sequenced using the Illumina 2 × 300 bp MiSeq platform. Similarly, FACS-sorted yeast from EBOV vaccinee libraries analyzed for binding to GP_{Δmuc} and FACS-sorted yeast from HIV-1 donor libraries analyzed for binding to VRC34-epitope scaffold-FP-APC were recovered, mini-prepped, and sequenced in the same way after each round of sorting to quantify V_H gene clonal enrichment across library sorting rounds.

For determining the EBOV vaccinee antibody library display efficiency, the original plasmablast V_H:V_L repertoire underwent highly stringent quality filtering (≥15 CDRH3:CDRL3 reads, 96% CDRH3 nt clustering), and CDR-H3 nucleotide junctions were mapped to CDR-H3 junctions recovered in the FITC⁺ Fab-expressing yeast library. Mapping was performed using usearch v5.2.32 with an exact nucleotide length match requirement and a 96% cutoff threshold for CDR-H3 junction nucleotide sequence match. For EBOV GP_{Δmuc} library antibody discovery via NGS (next-generation sequencing) clonal lineage tracking, CDR-H3 amino acid sequences from NGS data sets that were enriched more than eightfold across multiple rounds of screening were synthesized, expressed in HEK293 cells, and tested for soluble binding to GP_{Δmuc} as Fabs (EBOV.YD.05–EBOV.YD.08). Consensus sequences for NGS-discovered antibodies were generated based on exact CDR-H3 and CDR-L3 nucleotide junction matches between the originally paired V_H:V_L, separate V_H, and separate V_L sequencing libraries (before yeast display screening), as previously described for antibody discovery from paired heavy:light sequence data sets^{42,45,46}. Briefly, consensus sequences were generated using usearch v5.2.32 from exact match reads to the CDR-H3 nucleotide or CDR-L3 nucleotide junctions for heavy or light chains, respectively, and plasmids containing antibody heavy and light chain sequences

were expressed via transient transfection in HEK293 cells for soluble antibody generation as previously described³².

Affinity characterization. Binding kinetics of anti-EBOV GP_{Δmuc} and anti-HIV-1 FP Fabs were determined using biolayer interferometry on a FortéBio Octet HTX instrument³². For EBOV GP_{Δmuc}-targeting antibodies, AR2G biosensors were coupled with GP_{Δmuc} (10 µg/ml in 10 mM acetate, pH 4.5) for 600 s. Typical capture levels after quenching with 1 M ethanolamine (pH 8.0) for 300 s were between 2 and 2.5 nm, and variability within the same protein did not exceed 0.25 nm. Biosensors were then equilibrated for 420 s in PBST-BSA (PBS + 1% BSA + 0.01% Tween + 0.02% sodium azide) before binding assessment of the Fab. Association of Fab was measured for 300–600 s and dissociation was measured for 300–600 s, both in PBST-BSA. Correction to subtract non-specific baseline drift was carried out by subtracting the measurements recorded for a sensor loaded with unrelated antigen (HIV-1 gp120).

For HIV-1 FP-targeting antibodies, streptavidin biosensors were used to capture VRC34-epitope scaffold-FP at 0.5 µg/mL in PBST-BSA. Typical capture levels for FP probe were between 0.4 and 0.7 nm. Biosensors were then equilibrated for 60 s in PBST-BSA before binding assessment of the Fab. Association of Fab was measured for 150 s and dissociation was measured for 150 s, both in PBST-BSA. Correction to subtract non-specific baseline drift was carried out by subtracting the measurements recorded for a sensor loaded without Fab. All assays were performed with agitation set to 1,000 r.p.m. at 30 °C. Data analysis and curve fitting were carried out using the Octet analysis software, version 9.0. Experimental data were fitted using a 1:1 binding model for all experiments. Global analyses of the complete data sets, assuming binding was reversible (full dissociation), were carried out using nonlinear least-squares fitting allowing a single set of binding parameters to be obtained simultaneously for all concentrations used in each experiment.

For anti-HA IgGs, recombinant HAs (H1 A/Solomon Islands/3/2006, H3 A/Wisconsin/67/2005)³³ were immobilized in separate channels by amine-coupling at pH 6.0. BSA was immobilized in the reference channel, to correct for buffer effects and non-specific binding signal. All SPR measurements were performed in HBS-EP running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% vol/vol surfactant P20; GE Healthcare). Serially diluted antibodies were injected in triplicates at 30 µl/min for 2 min and allowed a dissociation time of 10 min. The chip was regenerated after each binding event with 50 mM Tris, pH 11.5, with a contact time of 15 s. The resulting sensorgrams were fitted with a two-state model (with conformation change) using Biaevaluation 3.0 software. The K_D values reported are the average of the three technical replicates ± s.d.

For EBOV GP_{Δmuc} antibody cross-competition assays, 10 µg/mL of GP_{Δmuc} was loaded onto Octet HTX biosensors using amine coupling (AR2G, FortéBio) for 600 s. The newly identified antibodies, the KZ52 antibody as a positive control²¹, and the isotype negative control (VRC01 (ref. 47)) were diluted to 50 µg/mL in PBST-BSA. Binding of competitor and analyte monoclonal antibodies (mAbs) were each assessed for 1,800 s. The assay was performed in duplicate with agitation at 1,000 r.p.m. at 30 °C. The percent inhibition (PI) was calculated by the equation: PI = 100 – [(probing mAb binding in the presence of competitor mAb)/(probing mAb binding in the absence of competitor mAb)] × 100.

Neutralization assays. For EBOV neutralization, GP-pseudotyped lentiviral particles expressing a luciferase reporter gene were produced as described previously⁴⁸, and were incubated 1 h at 37 °C with serially diluted purified mAbs. HEK293T cells were infected with the lentivirus:mAb mixture for 72 h in presence of polybrene (5 µg/ml, Sigma-Aldrich). Luciferase expression was assessed with Bright Glo (Promega) using a Victor X3 Plate Reader (PerkinElmer). Cell infection was calculated relative to the negative control antibody VRC01.

HIV-1 neutralization was assessed in TZM-bl cells as described previously¹⁹. Briefly, 293T cells were co-transfected by a pSG3ΔEnv backbone and an HIV-1 Env expression plasmid to produce Env-pseudotyped virus stocks. Viruses were mixed with fivefold serially diluted mAbs starting at 50 µg/ml, and incubated at 37 °C for 1 h before being added to the cells. After incubation at 37 °C for 48 h, the supernatants were removed and the cells were lysed. Luciferase activity was measured. 50% inhibitory concentrations (IC₅₀) were determined as described¹⁹.

For flu neutralization, influenza pseudotyped lentiviral vectors expressing a luciferase reporter gene were produced as described⁴⁹. Briefly, the following plasmids: 17.5 µg pCMVΔR8.2, 17.5 µg pHRCMV-Luc, 0.3 µg pCMV Sport/h TMPRSS2, and 1 µg CMV/R-HA and 0.125 µg corresponding CMV/R-NA of a given strain of influenza virus were transiently co-transfected into 15-cm tissue culture plate of 293T cells using Fugene6 (Promega). Cells were transfected overnight and replenished with fresh medium. Forty-eight hours later, supernatants were harvested, filtered through a 0.45-µm PES (polyethersulfone) membrane filter, aliquoted, and frozen at -80 °C. For neutralization assays, monoclonal antibodies at various dilutions were mixed with pseudoviruses and incubated at 37 °C for 1 h before adding to 293A cells in 96-well plates (10,000 cells per well). Seventy-two hours later, cells were lysed in cell culture lysis buffer (Promega, Madison, WI) before mixing with luciferase assay reagent (Promega). Light intensity was quantitated with a PerkinElmer microplate reader and antibody neutralization results, in lower light intensity.

Statistical analysis. Refer to the **Life Sciences Reporting Summary**.

Data availability. Antibody sequences for VRC34.YD.01-VRC34.YD.07 have been deposited in GenBank ([MF990262](#)–[MF990275](#)). NGS data sets for EBOV

and HIV-1 library screening have been deposited in the NCBI Short Read Archive (SRA) under accession number [SRP125922](#).

31. Côté, M. *et al. Nature* **477**, 344–348 (2011).
32. Misasi, J. *et al. Science* **351**, 1343–1346 (2016).
33. Whittle, J.R.R. *et al. Proc. Natl. Acad. Sci. USA* **108**, 14216–14221 (2011).
34. Lee, J.E. *et al. Nature* **454**, 177–182 (2008).
35. Olinger, G.G. Jr. *et al. Proc. Natl. Acad. Sci. USA* **109**, 18030–18035 (2012).
36. Lavinder, J.J. *et al. Proc. Natl. Acad. Sci. USA* **111**, 2259–2264 (2014).
37. Doria-Rose, N.A. *et al. J. Virol.* **83**, 188–199 (2009).
38. Doria-Rose, N.A. *et al. J. Virol.* **84**, 1631–1636 (2010).
39. Benatuil, L., Perez, J.M., Belk, J. & Hsieh, C.M. *Protein Eng. Des. Sel.* **23**, 155–159 (2010).
40. Reich, L.L., Dutta, S. & Keating, A.E. *J. Mol. Biol.* **427**, 2135–2150 (2015).
41. Wang, B. *et al. MAbs* **8**, 1035–1044 (2016).
42. DeKosky, B.J. *et al. Proc. Natl. Acad. Sci. USA* **113**, E2636–E2645 (2016).
43. Cale, E.M. *et al. Immunity* **46**, 777–791. e10 (2017).
44. Whitehead, T.A. *et al. Nat. Biotechnol.* **30**, 543–548 (2012).
45. DeKosky, B.J. *et al. Nat. Biotechnol.* **31**, 166–169 (2013).
46. Wang, B. *et al. Sci. Rep.* **5**, 13926 (2015).
47. Wu, X. *et al. Science* **329**, 856–861 (2010).
48. Sullivan, N.J. *et al. PLoS Med.* **3**, 0865–0873 (2006).
49. Yang, Z.Y. *et al. J. Virol.* **78**, 4029–4036 (2004).

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

► Experimental design

1. Sample size

Describe how sample size was determined.

No sample-size calculation was performed

2. Data exclusions

Describe any data exclusions.

No data were excluded

3. Replication

Describe whether the experimental findings were reliably reproduced.

All attempts at replication were successful

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

This was not relevant to this study

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

This was not relevant to this study

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- ☒ ☐ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ A statement indicating how many times each experiment was replicated
- ☒ ☐ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- ☒ ☐ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- ☒ ☐ The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- ☒ ☐ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- ☐ ☒ Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

FlowJo 10.0.7, Octet analysis software, version 9.0, Biaevaluation 3.0, usearch v5.2.32

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Anti-flag FITC (clone M2, F4049, Sigma-Aldrich), anti-human CD3 (Brilliant Violet 421, clone SP34-2, BD Biosciences), anti-human CD19 (PE-Cy7, clone HIB19, BD Biosciences), anti-human CD4 (Brilliant Violet 421, clone OKT4, BioLegend), anti-human CD8a (Brilliant Violet 421, clone RPA-T8, BioLegend), anti-human CD14 (Brilliant Violet 421, clone M5E2, BioLegend), anti-human CD20 (Brilliant Violet 605, clone 2H7, BioLegend), anti-human CD27 (Brilliant Violet 711, clone O323, BioLegend), and anti-human CD38 (Alexa Fluor® 680, clone OKT10, custom-conjugated at the Vaccine Research Center, NIAID)

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

HEK293T and TZM-bl were from NIH

b. Describe the method of cell line authentication used.

They have been authenticated by morphology checking, and through use in other experiments

c. Report whether the cell lines were tested for mycoplasma contamination.

They were tested negative for mycoplasma contamination

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No commonly misidentified cell lines were used

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

No animals were used

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

EBOV vaccine donor is a healthy human volunteer enrolled in a clinical trial with a Phase I Ebola GP vaccine (NCT02408913). This volunteer was first immunized with chimpanzee-derived replication-defective adenovirus encoding EBOV GP, then boosted 30 weeks and 5 days later with modified vaccinia Ankara encoding EBOV GP.

Donor N123 is a chronically HIV-1 infected individual enrolled in NIAID under a clinical protocol approved by the NIAID Institutional Review Board. This donor was diagnosed with HIV-1 in 2000. After more than nine years of infection, this donor showed a CD4 T cell count of 463 cells/ml and a plasma HIV-1 viral load of 4,920 RNA copies/ml. This donor was not on antiretroviral treatment.

Influenza vaccine donor is a healthy human volunteer vaccinated with a trivalent inactivated influenza vaccine (IIV3: A/Solomon Islands/3/2006, A/Wisconsin/67/2005, B/Malaysia/2508/2004).

Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

► Data presentation

For all flow cytometry data, confirm that:

- ☒ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ 3. All plots are contour plots with outliers or pseudocolor plots.
- ☒ 4. A numerical value for number of cells or percentage (with statistics) is provided.

► Methodological details

- | | |
|--|--|
| 5. Describe the sample preparation. | Yeast cells with surface expressed Fabs were collected, washed with PBS +0.5% BSA+2mM EDTA, then stained with different fluorescent antigens at room temperature for 30min, followed by staining with 2µg/ml anti-flag FITC at 4°C for 15min. Cells were washed twice with ice-cold PBS+0.5% BSA +2mM EDTA and analyzed on a FACS Aria II. |
| 6. Identify the instrument used for data collection. | BD FACS Aria II |
| 7. Describe the software used to collect and analyze the flow cytometry data. | BD FACS Diva was used to collect data and control the sorting process, FlowJo 10.0.7 was used to analyze data and prepare figs |
| 8. Describe the abundance of the relevant cell populations within post-sort fractions. | The abundance of relevant yeast cells after each round of sorting for EBOV and HIV-1 libraries was assessed by high-throughput sequencing as described in the manuscript |
| 9. Describe the gating strategy used. | FSC/SSC gates were first drawn to isolate single yeast cells, and gates for Fab expression positive and antigen binding positive cells were drawn by comparing the library with both positive and negative control cells. |

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☒