

1 **Shared HIV envelope-specific B cell clonotypes induced by a pox-protein vaccine**  
2 **regimen**

3 Kristen W. Cohen<sup>1</sup>, Lamar Ballweber-Fleming<sup>1</sup>, Michael Duff<sup>1</sup>, Rachael E. Whaley<sup>1</sup>, Aaron  
4 Seese<sup>1</sup>, Daniel Imhoff<sup>1,3</sup>, Zoe Moodie<sup>1</sup>, Ollivier Hyrien<sup>1,\*</sup>, M. Juliana McElrath<sup>1,2,\*</sup>

5 <sup>1</sup>Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle,  
6 WA, U.S.A.

7 <sup>2</sup>Department of Medicine, University of Washington, Seattle, WA, U.S.A.

8 <sup>3</sup>Currently at: Bristol Myers Squibb, Seattle, WA, U.S.A.

9 \*Address correspondence to:

10 Ollivier Hyrien and M. Juliana McElrath, 1100 Fairview Ave. N., Seattle, WA 98109. Email:

11 [jmcelrat@fredhutch.org](mailto:jmcelrat@fredhutch.org) and [ohyrien@fredhutch.org](mailto:ohyrien@fredhutch.org)

12

13

## 14 **Summary**

15           An effective HIV-1 vaccine will likely induce potent, broad neutralizing antibodies.  
16 No candidate vaccines have elicited these responses presumably because they fail to  
17 activate human B cell precursors that can affinity mature to generate broad neutralizing  
18 antibodies. To identify the B cell clonotypes that are elicited, we conducted in-depth  
19 analyses of the envelope-specific B cell repertoire in recipients of ALVAC-HIV vector  
20 (vCP2438) and bivalent subtype C gp120 protein (HVTN100). We observed high  
21 frequencies of envelope-specific IgG<sup>+</sup> memory B cells with restricted immunogenetic  
22 diversity, relative to non-vaccine induced memory B cells, with preferential expansions of  
23 distinct variable genes but limited accumulation of mutations. Many envelope-specific  
24 clonotypes were shared across vaccinees, but did not overlap with the envelope-negative  
25 memory repertoire, within and across subjects. Single-cell sequencing of envelope-  
26 specific IgG<sup>+</sup> memory B cells often revealed *VH1-2\*02* and *VK3-20* sequence co-  
27 expression and in one case, contained a 5 amino acid CDRL3, the canonical signature of  
28 VRC01-class antibodies, confirming that these B cells are extremely rare but detectable.  
29 Our study provides evidence that immunogens play a critical role in selecting and  
30 restricting the responding B cell repertoire and supports the rational design of HIV  
31 vaccines targeting specific B cell lineages for induction of broadly-reactive neutralizing  
32 antibodies.

33

34

35

## 36 Introduction

37 The HIV-1 epidemic is now in its fourth decade, and developing a highly effective  
38 preventive vaccine remains critical to end this global health crisis. The induction of  
39 broadly neutralizing antibodies (bnAbs) is widely thought to be a necessary component  
40 of a successful vaccine to prevent HIV-1 infection. This assumption is based on  
41 numerous preclinical studies indicating that passive administration of bnAbs can prevent  
42 SIV or SHIV infection following challenge (Balazs et al., 2014; Shingai et al., 2014), and  
43 it is further substantiated in humans, based on more recent findings that repeated IV  
44 infusions of the broadly neutralizing VRC01 mAb was associated with protection in  
45 humans against circulating VRC01-sensitive HIV-1 strains (Corey et al., 2021). However,  
46 no clinical vaccine strategies have thus far succeeded in reliably eliciting tier 2, including  
47 autologous strain-specific, neutralizing antibody responses. By contrast, approximately  
48 30% of people living with HIV generate bnAbs within 2 years of infection, indicating that  
49 viral diversity pushes antibody maturation toward broad neutralizing activities (Doria-  
50 Rose et al., 2010; Hraber et al., 2014; Mikell et al., 2011; Piantadosi et al., 2009; Sather  
51 et al., 2012; Simek et al., 2009; van Gils et al., 2009). Thus in some cases, human B  
52 cells have the capacity to generate anti-HIV-1 bnAbs, and understanding the B cell events  
53 and molecular profiles required to induce bnAbs following immunization is a major priority  
54 in HIV development (Kwong et al., 2013).

55 When B cells differentiate from naïve to memory, they also undergo rounds of  
56 mutation and selection for higher affinity B cell receptors (BCR) and produce antibodies  
57 with increasing affinity to HIV-1 envelope (Env). B cells predicted to have the potential of  
58 evolving into mature B cells able to secrete bnAbs are found at extremely low frequencies

59 in the naïve repertoire (Havenar-Daughton et al., 2018; Jardine et al., 2016a).  
60 Nevertheless, germline-targeting HIV vaccine designs have advanced into clinical trials  
61 to first prime germline bnAb precursors that can then be directed to affinity mature with  
62 boosts of sequentially more native-like HIV-1 envelopes (Briney et al., 2016; Jardine et  
63 al., 2013; Jardine et al., 2015; Jardine et al., 2016b; McGuire, 2019; McGuire et al., 2016;  
64 McGuire et al., 2013; Sok et al., 2016). Clinical trials of germline-targeting HIV  
65 immunogens aim to induce memory B cells expressing BCRs that share similarities to the  
66 unmutated common ancestors of known bnAbs. These trials rely on deep sequencing of  
67 the vaccine-induced BCR repertoire to evaluate the immunogenicity of the vaccine  
68 regimens. Few studies have investigated the molecular signatures of Env-specific BCR  
69 repertoires induced by HIV-1 vaccines and whether repeated immunizations with Env  
70 vaccines induce somatic hypermutation and features frequently associated with bnAbs  
71 (Basu et al., 2020; Easterhoff et al., 2017).

72 To address this issue, we examined the Env-specific memory B cell responses  
73 elicited in HVTN 100 (NCT02404311), a phase 1/2 randomized, double-blind, placebo-  
74 controlled clinical trial that evaluated the safety and immunogenicity of a clade B/C  
75 ALVAC-HIV (vCP2438) (months 0, 1) with bivalent subtype C gp120/MF59 (months 0, 1,  
76 3, 6, 12) in HIV-seronegative low-risk populations South Africans (Bekker et al., 2018).  
77 Notably, the subtype C RV144-like vaccine regimen was subsequently tested in HVTN  
78 702, a phase 2b-3 efficacy trial to evaluate HIV-1 prevention in populations at-risk in sub-  
79 Saharan Africa (Gray et al., 2021). The vaccine regimen did not induce nAbs (Bekker et  
80 al., 2018), and failed to protect against HIV-1 acquisition (Gray et al., 2021).

81 Here, we interrogated the BCR repertoire, the diversity of the Env-specific memory  
82 B cell repertoire, the extent of somatic hypermutation after repeated vaccination, the  
83 between-subject similarities and overlap between clonotypes and similarities to known  
84 bnAbs. Our findings of Env-specific BCR repertoires produced by conventional and  
85 ineffective Env gp120 immunogens in non-neutralizing HIV-1 prime-boost vaccine  
86 regimens offers a benchmark against which novel HIV-1 immunogens may be compared  
87 and provides additional context for the evaluation of germline-targeting immunogens.

88

## 89 **Results**

90 *HVTN 100 evaluated a subtype C based modified RV144-like vaccine regimen*

91 HVTN 100 was a phase 1 HIV-1 vaccine clinical trial of ALVAC-HIV vCP2438  
92 including subtypes B and C Env immunogens plus MF59-adjuvanted bivalent subtype C  
93 1086 and TV1 gp120 proteins. The study enrolled 250 participants in South Africa.  
94 (Bekker et al., 2018). Similar to the RV144 vaccine, the HVTN 100 regimen induced non-  
95 neutralizing Env-specific binding antibody responses that had the ability to mediate  
96 antibody-dependent cellular cytotoxicity (ADCC) along with polyfunctional Env-specific T  
97 cell responses (Bekker et al., 2018; Laher et al., 2020).

98 We performed in-depth analysis of the vaccine-induced peripheral Env-specific  
99 memory B cells in a subset of 14 out of 185 per protocol vaccine recipients enrolled in the  
100 HVTN 100 trial (Figure S1A). The subset was randomly selected among those with a  
101 positive serum binding antibody response to the vaccine-matched HIV envelope, 1086  
102 gp120, in the bivalent protein immunogen. The demographics and vaccine-matched Env-

103 specific binding antibody, antibody dependent cellular cytotoxicity (ADCC) and CD4+ T-  
104 cell responses of the subset were representative of those from the entire trial cohort  
105 (Figure S1B-S1E). No participants in the subset became HIV-infected during the study.

106

107 *Durability of Env-specific memory B cells 6 months after vaccination correlates with their*  
108 *expansion after boost immunization*

109 To characterize the induction and maintenance of vaccine-elicited memory B cells,  
110 we evaluated, we evaluated the phenotype and magnitude of vaccine-matched subtype  
111 C 1086 gp120-specific memory B cells from vaccine recipients 6 months after the 4<sup>th</sup>  
112 immunization, the day of 5<sup>th</sup> immunization (month 12), and 2 weeks later (month 12.5)  
113 (Figure 1A and S1). Six months after the 4<sup>th</sup> immunization, 1086 gp120-specific IgD<sup>-</sup>  
114 memory B cells were low but detectable in all subjects (median 0.15%) and then  
115 increased significantly in response to the boost immunization (median 0.6%, Figure 1B  
116  $p=0.0001$ ). A minority of the total Env-specific IgD<sup>-</sup> B cells were the IgM isotype (median  
117 5.6%), which decreased to 1% after the 5<sup>th</sup> vaccination (Figure 1C,  $p<0.0001$ ). The  
118 majority of responding Env-specific IgD<sup>-</sup> memory B cells were IgG<sup>+</sup> at month 12 (median  
119 64%). Two weeks after the 5<sup>th</sup> vaccination (month 12.5), the proportion of Env-specific  
120 memory B cells that were IgG isotype increased significantly to a median of 79%  
121 ( $p=0.0047$ ).

122 The median frequency of gp120-specific IgG<sup>+</sup> memory B cells was 0.28% at month  
123 12 and increased 5-fold to 1.33% of IgG<sup>+</sup> B cells at month 12.5 (Figure 1D,  $p=0.0001$ ).  
124 Following the 5<sup>th</sup> vaccination, the Env-specific IgG<sup>+</sup> B cells also increased in CD27

125 expression (Figure 1E,  $p=0.0019$ ) suggesting that the resting vaccine-specific memory B  
126 cells had lower CD27 expression 6 months post-vaccination, whereas the responding  
127 vaccine-specific B cells upregulated expression of CD27 and class-switched to IgG  
128 quickly after the boost.

129 Interestingly, the frequencies of  $gp120^+$   $IgG^+$  B cells six months post 4<sup>th</sup>  
130 vaccination/pre-boost (month 12) correlated significantly with the magnitude of  $gp120^+$   
131  $IgG^+$  B cells post-5<sup>th</sup> vaccination (month 12.5) (Figure 1F;  $r=0.87$ ,  $p=0.0027$ ) and the  
132 magnitude of vaccine-matched serum binding antibody responses at peak (2 weeks post-  
133 4<sup>th</sup> vaccination; Figure 1G;  $r=0.77$ ,  $p=0.038$ ). Thus, the frequencies of circulating Env-  
134 specific memory B cells were positively associated with the vaccine-elicited titers of Env-  
135 specific serum antibodies and predicted the magnitude of circulating B cells after boost.

136

137 *The vaccine-induced Env-specific BCR repertoire is narrower than the polyclonal non-*  
138 *vaccine selected memory repertoire*

139 We explored the molecular immunogenetics of the vaccine-elicited memory B cell  
140 repertoire in order to identify characteristics shared by vaccine recipients. We isolated  
141 vaccine-specific B cells using fluorescent probes of the vaccine-matched 1086  $gp120$ ,  
142 which was the vaccine-matched Env with the highest serum antibody response. To  
143 characterize the B cell receptor (BCR) repertoire of Env-specific memory B cells, we  
144 sorted  $1086\ gp120^+$  and  $gp120^-$  (negative)  $IgG^+$  B cells two weeks post-5<sup>th</sup> vaccination  
145 (month 12.5), which was the timepoint with the highest frequency of Env-specific memory  
146 B cells, and directly sequenced their  $V_H$ ,  $V_K$  and  $V_L$  immunoglobulin genes. We restricted

147 cell sorting to IgG<sup>+</sup> B cells because most vaccine-specific memory B cells were of IgG  
148 isotype. The majority of the gp120<sup>-</sup> IgG<sup>+</sup> cells are expected to be representative of the  
149 pre-vaccination polyclonal repertoire expanded in response to other infections or  
150 immunizations but not in response to the HIV vaccination. Thus, we sequenced the BCR  
151 of sorted 1086 gp120<sup>+</sup> and gp120<sup>-</sup> IgG<sup>+</sup> B cells in parallel to compare the two repertoires.

152 We hypothesized that five doses of the Env-expressing vaccine regimen over one  
153 year drove the vaccine-specific B cells to undergo multiple rounds of expansion and  
154 selection, resulting in a narrower repertoire compared to the polyclonal non-vaccine  
155 specific memory B cell repertoire. Thus, we sought to compare diversity between the BCR  
156 repertoires of the 1086 gp120<sup>+</sup> and gp120<sup>-</sup> memory B cells (Figure S3). Since the  
157 diversity of a BCR repertoire is partly determined by the number of germline genes utilized  
158 to encode its heavy and light chains, we first assessed the diversity of the V genes  
159 encoding the gp120<sup>+</sup> and gp120<sup>-</sup> IgG<sup>+</sup> repertoires using three indices of alpha diversity.  
160 The richness index is defined as the number of V genes encoding a given repertoire,  
161 indicated that the V regions of the heavy, kappa, and lambda chains of the vaccine-  
162 induced repertoire were encoded by fewer V genes than the non-vaccine induced  
163 repertoire ( $p < 0.001$  in all cases; LMM adjusting for cell count and read count; Figure 2A-  
164 2C). The gp120<sup>+</sup> IgG<sup>+</sup> B cell repertoire had lower richness in 13 out 14 participants for the  
165 heavy chain (median of 55 versus 70 V<sub>H</sub> genes), in 13 out 14 participants for the kappa  
166 chain (median of 32 versus 43 V<sub>K</sub> genes), and in 14 out 14 participants for the lambda  
167 chain (median of 18 versus 30 V<sub>L</sub> genes). Since the V gene richness may be sensitive to  
168 cell and read counts, we considered Shannon's and Simpson's indices as alternative  
169 metrics to assess diversity, both of which take evenness into consideration. Shannon's



170 index also suggested that the heavy, kappa, and lambda chains of the vaccine-induced  
171 repertoire were encoded by fewer V genes than in the non-vaccine induced repertoire ( $p$   
172  $< 0.001$  in all cases; LMM adjusting for cell count and read count; Figure S2A). The  
173 Simpson index also indicated reduced diversity in the gp120<sup>+</sup> repertoire of the kappa and  
174 lambda chains ( $p < 0.001$ ; LMM adjusting for read count), but it did not reach significance  
175 for the heavy chain ( $p = 0.73$ ; LMM adjusting for cell count and read count; Figure S2B).  
176 Taken together, these results suggest that, compared to the gp120<sup>-</sup> IgG<sup>+</sup> BCR repertoire,  
177 the heavy and light chains of gp120<sup>+</sup> IgG<sup>+</sup> BCR repertoire are encoded by a smaller set  
178 of V genes, and that within these V genes, some are preferentially selected and  
179 expanded. The observation of a focused vaccine-induced repertoire may be attributed to  
180 the preferential expansion of B cells with specific genetic signatures.

181

182 *In the gp120<sup>+</sup> BCR repertoire, specific V genes were preferentially expanded across*  
183 *individuals*

184 Ideally, a germline-targeting vaccine would induce an on-target B cell response  
185 consistent across vaccine recipients. To assess the similarity of the gp120<sup>+</sup> BCR  
186 repertoire across vaccine-recipients, we measured the overlap between the sets of V  
187 genes that encoded each of the heavy, kappa, and lambda chains of gp120<sup>+</sup> BCR from  
188 different individuals using the Jaccard index. The average Jaccard indices for the heavy,  
189 kappa and lambda chains of the gp120<sup>+</sup> BCR repertoire were 0.64, 0.64 and 0.59,  
190 respectively. Thus, for all three chain types, on average about two thirds (66%) of the  
191 BCR in the gp120<sup>+</sup> repertoire of two distinct vaccinees were encoded by the same pool  
192 of V genes, whereas the remaining one third (33%) of gp120<sup>+</sup> BCR were encoded by

193 distinct (i.e., non-overlapping) V genes. Although the value of two-third may under-  
194 estimate the overlap between vaccine-induced repertoires, it provides a benchmark for  
195 trials that target specific B cell lineages.

196         Since the Jaccard index indicated significant overlap in the use of V genes among  
197 the gp120<sup>+</sup> B cells of distinct individuals, we next sought to determine which V genes  
198 were over-utilized by the gp120<sup>+</sup> repertoire across participants. In evaluating the  
199 distribution of V gene families used by the gp120<sup>+</sup> and gp120<sup>-</sup> repertoires, we observed  
200 an immunodominance of the *V<sub>H</sub>1* family, in particular *V<sub>H</sub>1-69* and *V<sub>H</sub>1-2*, and increased  
201 usage of the *V<sub>H</sub>4* gene family and decreased usage of the *V<sub>H</sub>6* and *V<sub>H</sub>2* gene families in  
202 the gp120<sup>+</sup> compared to the gp120<sup>-</sup> repertoires (Figure S3). In contrast, the distributions  
203 of kappa chain V gene families were similar between the repertoires and were dominated  
204 by *V<sub>K</sub>1*, *V<sub>K</sub>2* and *V<sub>K</sub>3* families (Figure S3). For lambda chains, the *V<sub>L</sub>4* family was enriched  
205 in the gp120<sup>-</sup> repertoire and underrepresented in the gp120<sup>+</sup> repertoire (Figure S3). We  
206 performed differential expression analysis to identify specific V genes unequally  
207 represented in the two repertoires (Figure 2D-2F; full list of comparisons can be found in  
208 the Supplement). For the *V<sub>H</sub>* BCR repertoire, we found that the *V<sub>H</sub>1-69-2*, *V<sub>H</sub>1-2*, *V<sub>H</sub>4-31*  
209 and *V<sub>H</sub>4-34* genes were significantly expanded in the gp120<sup>+</sup> repertoire compared to the  
210 gp120<sup>-</sup> repertoire (unadjusted p values < 0.01 and FDR < 10%; Figure 2D). For the *V<sub>K</sub>*  
211 BCR repertoire, the *V<sub>K</sub>3-20* and *V<sub>K</sub>1D-12* genes were the most significantly enriched in  
212 the gp120<sup>+</sup> relative to gp120<sup>-</sup> BCR repertoire (unadjusted p=0.009 and 0.013,  
213 respectively; Figure 2E). For the *V<sub>L</sub>* BCR repertoire, the *V<sub>L</sub>2-14* gene was the most  
214 significantly enriched in the gp120<sup>+</sup> compared to the gp120<sup>-</sup> repertoire (unadjusted  
215 p=0.0052; Figure 2F). However, in contrast to the heavy chain, none of the light chain V

216 genes were significant after controlling FDR for multiple comparisons. Additionally,  
217 multiple V genes were significantly under-represented in the gp120<sup>+</sup> BCR repertoire and  
218 some V genes did not encode any of the gp120<sup>+</sup> BCR isolated from the 14 vaccine-  
219 recipients (e.g., *V<sub>H</sub>6-1*).

220

### 221 *Evidence of limited somatic hypermutation in the gp120<sup>+</sup> BCR repertoire*

222 Since we observed selection and expansion of specific genes among the vaccine-specific  
223 B cells, we next asked whether the gp120<sup>+</sup> IgG<sup>+</sup> B cells had undergone substantial affinity  
224 maturation by month 12.5, after 5 vaccinations. When evaluating the frequency of  
225 nucleotide mutations of individual BCR sequences compared to their assigned germline,  
226 we found that the rates of mutation in the V gene segments of the heavy, kappa and  
227 lambda sequences were lower among the gp120<sup>+</sup> compared to the gp120<sup>-</sup> repertoires  
228 (Figure 3A-3C;  $p < 0.001$ ). For the *V<sub>H</sub>* BCR repertoire, the average percent mutation was  
229 7.43% in the gp120<sup>+</sup> repertoire and 8.4% in the gp120<sup>-</sup> repertoire (Figure 3A). However,  
230 the rate of mutation observed in the gp120<sup>+</sup> repertoire was very similar to the mutation  
231 rate observed after a 5<sup>th</sup> vaccination of the RV144 regimen in the trial RV305 (avg 6.7%)  
232 (Easterhoff et al., 2017). The lower mutation rates among the vaccine-specific memory  
233 B cells after five vaccinations is indicative of limited affinity maturation, possibly due to  
234 limited or short-lived germinal center activity.

235

236 *Vaccination did not induce antibodies with bnAb characteristics such as longer CDRH3*  
237 *or shorter CDRL3 regions*

238 Several anti-HIV-1 bnAbs have exceptionally long CDRH3 regions (e.g., PGDM1400,  
239 CDRH3 length=36; PGT121, CDRH3 length= 26). Their unusual lengths allow the bnAbs  
240 to reach into and between glycans on the HIV Env protein to make contacts with the  
241 amino acids (Sok et al., 2014). However, antibodies with long CDRH3 regions are  
242 uncommon in the naïve B cell repertoire and believed to be the result of rare VDJ  
243 recombination events. The CDR3 region is highly variable due to the random  
244 rearrangement of V(D)J gene segments and the addition and subtraction of nucleotides  
245 at the ends of these genes during somatic recombination. The average CDR3 lengths for  
246 the  $V_H$ ,  $V_K$  and  $V_L$  were not significantly different between the gp120<sup>+</sup> and gp120<sup>-</sup>  
247 repertoires (Figure 3D-3F). However, there was a trend toward longer CDRH3 lengths  
248 among the gp120<sup>+</sup> compared to gp120<sup>-</sup> BCRs (median=16.1 versus 14.9;  $p=0.068$ ; Fig.  
249 3D), but no significant differences in the frequency of BCRs with CDRH3 regions greater  
250 than 24 a.a. in gp120<sup>+</sup> compared to gp120<sup>-</sup> IgG<sup>+</sup> B cells ( $p=0.30$ ).

251 The average CDRL3 of both kappa and lambda chains is 9 a.a. long, but some  
252 HIV bnAbs, including VRC01-class bnAbs, are characterized by a 5 a.a. CDRL3 that is  
253 required to maintain a critical angle of approach to the CD4bs and prevent clashing with  
254 particular gp120 residues. These short CDRL3 are known to be rare in the naïve and  
255 memory B cell repertoires (Jardine et al., 2016a). We evaluated the CDRL3 length  
256 distribution of kappa and lambda BCR sequences from the gp120<sup>+</sup> B cells compared to  
257 gp120<sup>-</sup> B cells to assess expansion of light chains with short CDRL3 sequences. The  
258 median CDRL3 lengths among the kappa and lambda gp120<sup>+</sup> BCR sequences, 9 and 10  
259 a.a., respectively, were not significantly different from those of gp120<sup>-</sup> B cells (Figure 3E

260 and 3F). In addition, the frequency of light chains with a 5 a.a. CDRL3 sequence were  
261 rare and not different between the gp120<sup>+</sup> and negative repertoires.

262

263 *Non-synonymous mutations concentrate in the complementary determining regions of V*  
264 *genes but don't correlate with vaccine-specific V gene expansions*

265 The rate of non-synonymous mutations in the gp120<sup>+</sup> repertoire was explored by  
266 comparing the a.a. sequence of the V-gene encoded region of the light and heavy chains  
267 of each BCR sequence to that of their assigned germline gene. The average mutation  
268 rate per a.a. position in the FR1-3 and CDR1-2, by chain type and gene family (Figure 3G-  
269 3I). When stratifying the analysis by framework (FR) and complementary determining  
270 regions (CDR; Figure 3G-3I), as expected, we observed higher accumulation of mutations  
271 in the CDRs compared to the FRs. The mutation rates in the  $V_H$  genes continued to be  
272 lower among the gp120<sup>+</sup> IgG<sup>+</sup> B cells compared to gp120<sup>-</sup> IgG<sup>+</sup> B cells, even among the  
273  $V_H$  BCR sequences encoded by V genes that were specifically expanded in the gp120<sup>+</sup>  
274 repertoire (i.e.  $V_{H1-69-2}$ ,  $V_{H1-2}$ ,  $V_{H4-31}$  and  $V_{H4-34}$ ). Possible exceptions included the  
275 FR1 of BCR sequences encoded by the  $V_{H4-61}$  family as well as the CDR2 of BCR  
276 sequences encoded by the  $V_{H2-26}$  and  $V_{H3-66}$  gene families which have accumulated  
277 more mutations in the gp120<sup>+</sup> than in the gp120<sup>-</sup> BCR repertoire (Figure 3I). However,  
278 these  $V_H$  genes were underrepresented in the gp120<sup>+</sup> repertoire. Similarly, the rate of  
279 mutation in the  $V_L$  genes was lower in gp120<sup>+</sup> IgG<sup>+</sup> B cells compared to gp120<sup>-</sup> IgG<sup>+</sup> B  
280 cells even when subdivided by V gene and region (Figure S4A and S4B). Overall, the  
281 CDR1 and CDR2 were the most mutated regions, followed by the FR1 and FR3 regions.  
282 The FR2 region was virtually unmutated. The amount of mutation varied by gene.

283 Interestingly BCRs encoded by any of the  $V_H$  genes found significantly over-expressed in  
284 the gp120+ repertoire were not among those with the highest mutation rates. In particular,  
285 the amount of mutations accumulated by the CDR1 and CDR2 regions of  $V_{H1-2*02}$ -  
286 encoded BCR were relatively low compared to the CDR1 and CDR2 regions of BCR  
287 encoded by other gene families. Similar observations can be made for the kappa and  
288 lambda chains (Figure S4).

289

290 *Accumulation of mutations in  $V_{H1-2}$  gp120+ repertoire actually diverges the repertoire*  
291 *further away from representative  $V_{H1-2}$  bnAbs*

292 Although vaccination induced BCRs with limited somatic hypermutations, we assessed  
293 whether it selected for BCR with mutations identical to those found in any  $V_{H1-2*02}$ -  
294 encoded bnAbs. Although encoded by a same V gene, the heavy chain of these bnAbs  
295 have accumulated distinct mutations, and we considered a representative subset of five  
296  $V_{H1-2*02}$ -containing bnAbs: VRC01 (Gristick et al., 2016), IOMA (Gristick et al., 2016),  
297 N6 (Huang et al., 2016), 3BNC117 (Scheid et al., 2011), and DH270.1 (Bonsignori et al.,  
298 2017). We calculated the Hamming distance between each  $V_{H1-2*02}$ -encoded read  
299 sequence (n=553 gp120+ BCR; n=35 gp120- BCR) and the heavy chain of the five  
300 selected bnAbs, by region (FR1-3, and CDR1-2) and repertoire (gp120- vs. gp120+)  
301 (Figure 4). The results indicate that distances to the bnAbs were region-, bnAb-, and  
302 repertoire-specific. For instance, the distance to the CDR2 region of VRC01 was virtually  
303 identical between the gp120- and gp120+ repertoires; however, the distances to the CDR2  
304 region of the other four bnAbs were overall smaller in the gp120- repertoire. Despite  
305 evidence for shared mutations with some bnAbs, the gp120+ BCRs encoded by the  $V_{H1}$ -

306 2\*02 gene remained overall distant to the selected ones. All FR3 read sequences were  
307 at least 9 a.a. mutations away from that of IOMA, 16 a.a. mutations from that of N6, 4 a.a.  
308 mutations from that of DH270.1, and 27 a.a. mutations from that of VRC-PG04. Overall,  
309 the mutation-selection profiles in each repertoire, by region and bnAb and suggests that  
310 mutations that accumulated in the gp120<sup>+</sup> B cells moved the repertoire further away from  
311 the V<sub>H</sub>1-2\*02 bnAbs (Figure 4).

312

313 *Shared vaccine-induced heavy chain clonotypes were detected across multiple subjects*

314 We identified heavy chain clonotypes by clustering BCR sequences based on the  
315 similarity of the a.a. sequence of their CDRH3. The clustering process was blinded to the  
316 V and J gene calls of the sequences, their antigenic specificity (gp120<sup>+</sup> vs. gp120<sup>-</sup>), and  
317 the participants from which they were sampled. Out of 13,381 unique sequences, we  
318 identified a total of 6,697 clonotypes. By inspecting the antigen-specificity of the BCR  
319 sequences assigned to identical clusters, we found that none of these clusters included  
320 sequences from both the gp120<sup>+</sup> and gp120<sup>-</sup> BCR repertoires (Figure 5A and 5D). In  
321 contrast, we observed that multiple clonotypes of unrelated origin were shared by several  
322 participants (Figure 5B-5D). Shared clonotypes are less likely to include BCR sequences  
323 from multiple participants if their clusters include few BCR sequences, and our analyses  
324 adjusted for cluster size (Figure 5C). Amongst the gp120<sup>+</sup> clonotypes that consisted of at  
325 least 20 BCR sequences, at least 60% of them were shared by several participants  
326 (Figure 5C and 5D). In contrast, amongst the gp120<sup>-</sup> clonotypes including at least 20  
327 BCR sequences, we found that only about 15% of clonotypes were shared by at least two

328 participants (Figure 5C and 5D). We conclude that the vaccine-induced repertoire was  
329 selected for in independent subjects indicative of convergent selection and evolution.

330

331 *High-throughput multiplexed next generation single-cell paired  $V_H/V_L$  sequencing*  
332 *demonstrate detection of rare  $gp120^+$   $IgG^+$  BCR clones*

333 We sought to characterize the molecular characteristics of vaccine-induced BCRs with  
334 paired  $V_H$  and  $V_L$  BCR sequences, and selected the four vaccine-recipients with the  
335 largest expansions of  $V_{H1-2*02}$  in their  $gp120$ -specific BCR repertoire analysis to  
336 evaluate the paired  $V_H$  and  $V_L$  BCRs. We performed single-cell sorting of  $1086 gp120^+$   
337  $IgG^+$  memory B cells followed by next generation sequencing of their BCRs.

338 We obtained heavy and/or light chain BCR sequences from 1,912  $gp120^+$   $IgG^+$  B  
339 cells (1,364 heavy chain, 379 kappa chain and 568 lambda chain BCR sequences). A  
340 total of 339  $gp120^+$   $IgG^+$  B cells had paired heavy and light chain BCR sequences (159  
341 kappa chains and 203 lambda chains, 23 cells registered both kappa and lambda chains).  
342 Even though the single-cell BCR sequencing analysis was limited to a subset of the  
343 participants analyzed in the bulk BCR sequencing analysis ( $n=4$ ), it independently  
344 confirmed many of the patterns of gene usage observed in the bulk data set. In particular,  
345 we observed expansions of  $V_{H1-2*02}$ ,  $V_{K3-20}$ , and  $V_{L2-14}$  (Figure 6A and 6B).  
346 Furthermore, the analysis of B cells expressing paired heavy and light chains confirmed  
347 that the  $V_{H1-2*02}$  heavy chain sequences were often paired with  $V_{K3-20}$  light chains,  
348 which is noteworthy since the combination is a common motif among many VRC01-class  
349 BCRs. Specifically of the 1,364 cells with heavy chain sequences, 181 contained a  $V_{H1-}$



350 2\*02 and of these, 38 of these were paired with a light chain. The average length of the  
351 amino acid sequence of the CDR3 of the light chain of the 38 sequences with a paired  
352  $V_H1-2*02$  heavy chain V gene was 8.6 for kappa and 10.3 for lambda. Two light chains  
353 paired with  $V_H1-2*02$  had five amino acid-long CDRL3s ( $V_K3-20$ , QHMYT and  $V_K1-5$ ,  
354 LPFRV), with one being a  $V_K3-20$ . These data confirm that Env-specific B cells with  
355 molecular signatures associated with VRC01-class of bnAbs can be induced by a non-  
356 neutralizing gp120-based prime-boost vaccine but these B cells are extremely rare (~  
357 2.6% of  $V_H1-2*02$  gp120<sup>+</sup> B cells or 0.29% of gp120<sup>+</sup> B cells) and their relevance is  
358 unknown at this point, except that it provides a benchmark to which VRC01 germline-  
359 targeting vaccines can be prepared.

360

## 361 Discussion

362 Most effective vaccines may protect against pathogens through multiple  
363 mechanisms, and antibodies may contribute through both neutralizing and non-  
364 neutralizing effector activities. Despite tremendous effort spanning many decades, an  
365 efficacious HIV vaccine continues to elude the field. However, recent technical advances,  
366 allowing for improved Env immunogen design and engineering, and tracking of the  
367 resultant humoral responses have provided renewed enthusiasm. Here, we have detailed  
368 the phenotypic and molecular characteristics of Env-specific B cells generated by a  
369 vaccine regimen in a pivotal trial (HVTN 100) that supported the decision to advance to a  
370 phase 2b/3 vaccine efficacy trial (HVTN 702). Although the vaccine regimen tested in  
371 HVTN 100 was not found to prevent HIV infection, we observed that the pox/protein

372 prime-boost regimen induced high frequencies of class-switched Env-specific memory B  
373 cells that they were durable up to 6 month post 4<sup>th</sup> vaccination. Of note, the original 4  
374 vaccination RV144 vaccine regimen waned in efficacy after 6 months post 4<sup>th</sup> vaccination  
375 (Robb et al., 2012). Consequently, a 5<sup>th</sup> vaccination was added to the HVTN 100 regimen  
376 at 6 months post-4<sup>th</sup> vaccination in order to boost and extend vaccine-induced immunity.  
377 Our findings indicate that the 5<sup>th</sup> vaccination rapidly increased the Env-specific memory  
378 B cells frequency (~5-fold), CD27 expression and IgG isotype, which together support the  
379 finding that the vaccine regimen, including the 5<sup>th</sup> vaccination/boost, effectively induced  
380 and expanded HIV gp120-specific memory B cells. Moreover, the frequencies of gp120<sup>+</sup>  
381 IgG<sup>+</sup> B cells prior to the 5<sup>th</sup> vaccination correlated with the frequency of gp120<sup>+</sup> IgG<sup>+</sup> B  
382 cells 2 weeks later and peak serum antibody responses, highlighting the direct  
383 relationship between memory B cells in circulation and bone marrow plasma cells, which  
384 secrete the serum antibodies.

385 In order to characterize the immunogenetics and clonotypes of vaccine-induced  
386 Env-specific memory B cells, we conducted next generation sequencing of the BCR  
387 repertoires of gp120-specific (gp120<sup>+</sup>) and non-vaccine specific (gp120<sup>-</sup>) IgG<sup>+</sup> memory B  
388 cells from the 14 subjects 2 weeks after the 5<sup>th</sup> vaccination. We found that the gp120<sup>-</sup>  
389 specific memory BCR repertoire was largely of the V<sub>H</sub>1 gene family. This is similar to a  
390 recent study of early circulating plasmablast and long-lived (>8 months) bone marrow  
391 plasma cell repertoires from a RV144-like DNA/gp120 vaccine regimen (Basu et al.,  
392 2020). In particular, in the gp120<sup>+</sup> repertoire, V<sub>H</sub>1-69 and V<sub>H</sub>1-2 were immunodominant,  
393 which have been previously observed among Env-specific repertoires induced both by  
394 HIV vaccination and infection (Basu et al., 2020; Huang et al., 2004; Smith et al., 2018).

395 The gp120-specific repertoire was less diverse than the gp120<sup>-</sup> memory repertoire, which  
396 is consistent with the hypothesis that the gp120<sup>-</sup> repertoire comprised of non-vaccine  
397 induced B cells with diverse specificities in contrast to the gp120<sup>+</sup> vaccine-induced  
398 repertoire. Interestingly, the estimates of BCR beta diversity further suggested that about  
399 two-thirds of the heavy and light V genes in the gp120<sup>+</sup> repertoires were shared between  
400 any two subjects. In fact, multiple V<sub>H</sub> genes were significantly expanded specifically in the  
401 gp120<sup>+</sup> repertoire compared to the gp120<sup>-</sup> memory repertoire consistently across  
402 subjects, including *V<sub>H</sub>1-69-2*, *V<sub>H</sub>1-2*, *V<sub>H</sub>4-31* and *V<sub>H</sub>4-34*. The *V<sub>H</sub>1-2* expansion is  
403 particularly interesting as it is a defining feature of the VRC01-class of bnAbs and its  
404 germline sequence encodes for many of the residues that are critical for binding envelope  
405 and neutralizing activity (Kwong and Mascola, 2012; West et al., 2012). Indeed, novel  
406 immunogens have been engineered to engage and activate VRC01-class bnAb  
407 precursors and are now being tested in clinical trials for their ability to expand VRC01-like  
408 B cells (CD4bs-specific with *V<sub>H</sub>1-2\*02,03* or *04* paired with a light chain with a 5 a.a.  
409 CDRL3) (Jardine et al., 2013; Jardine et al., 2016a; McGuire et al., 2013). Thus, the  
410 consistent expansion of gp120<sup>+</sup> *V<sub>H</sub>1-2\*02* memory B cells by vaccination with a gp120-  
411 based HIV vaccine suggests these B cells can be readily expanded by vaccination and  
412 that VRC01 germline-targeting vaccines will need to further focus the response to avoid  
413 off-target responses.

414 We also identified *V<sub>H</sub>4-34* as one of the V gene segments statistically significantly  
415 over-represented in the vaccine-induced repertoire. Previous studies have reported that  
416 *V<sub>H</sub>4-34* BCRs may be associated with autoimmune disorders and autoreactivity (Bhat et  
417 al., 2002) and enriched among highly mutated BCRs of HIV-1-infected individuals (Roskin

418 et al., 2020). Although over-represented in the vaccine-induced repertoire, the frequency  
419 of  $V_{H4-34}$  BCR remained low in our cohort (mean=0.39% vs 0.027% in the  $gp120^+$  vs  
420  $gp120^-$  repertoire). In addition, we observed that the FR1 of  $V_{H4-34}$  BCR tended to be  
421 more mutated than the FR1 of those encoded by other V gene segments. Somatic  
422 hypermutations also accumulated in their CDR1, but mutations were overall rare in the  
423 FR2, CDR2 and FR3. This mutational profile is consistent with previous reports that  
424 implicated FR1 in the antigen binding by  $V_{H4-34}$  (Li et al., 1996). Taken together, these  
425 results suggest that vaccination might have partly recapitulated HIV exposure in inducing  
426  $V_{H4-34}$  clones; however, while the FR1 and CDR1 were the most prone to somatic  
427 hypermutation, we did not find evidence of higher mutation rates of codon 26 in the  
428  $gp120^+$  than in the  $gp120^-$  repertoire, unlike previously observed in a cohort of HIV-1  
429 infected patients (Roskin et al., 2020). The enrichment of certain  $V_H$  genes among  
430 antigen-specific B cells has also been observed in the humoral response to COVID-19  
431 and SARS-CoV-2 vaccination (Chen et al., 2021b; Kim et al., 2021). It is not unexpected  
432 that the vaccine-induced repertoire would be focused and is likely attributable to the  
433 preferential expansion of B cells with specific V gene characteristics in response to  
434 vaccination.

435         Since we observed clear selection and expansions of specific gene families among  
436 the vaccine-specific B cells, we were surprised to discover that the  $gp120^+$  IgG<sup>+</sup> B cells  
437 had undergone limited somatic hypermutation after 5 vaccinations. The rate of mutation  
438 in the  $gp120^+$  memory  $V_H$  repertoire was similar to previous reports of RV305 vaccine  
439 recipients who had received 5 doses of the RV144 vaccine regimen (Easterhoff et al.,  
440 2017). However, the mutation rate remained significantly lower than the  $gp120^-$  memory

441 repertoire. Our data suggest that despite receiving five immunizations, vaccine-  
442 responding B cells display a reduced level of affinity maturation as compared to B cells  
443 from natural infection with other pathogens, as indicated by the gp120<sup>-</sup> repertoire. The  
444 rate of non-synonymous mutations in the gp120<sup>+</sup> repertoire was further explored by  
445 comparing the amino acid sequence of the V-gene encoded region of the light and heavy  
446 chains of each BCR sequence to that of their assigned germline gene. We found that the  
447 non-synonymous mutations accumulated in the complementary determining regions of V  
448 genes but did not correlate with the V genes that had preferentially expanded. This  
449 suggests that although some V genes were preferentially expanded, they did not undergo  
450 sustained germinal center-associated maturation. Although the vaccination-induced  
451 BCRs had limited somatic hypermutations, we assessed the tendency of *V<sub>H</sub>1-2\*02*-  
452 encoded gp120<sup>+</sup> BCRs to share mutations with those found in *V<sub>H</sub>1-2\*02*-encoded bnAbs  
453 and found that the limited somatic hypermutation in the gp120<sup>+</sup> B cells appeared to  
454 diverge further from the *V<sub>H</sub>1-2\*02* bnAbs.

455 To further understand the similarity of B cell responses across subjects, we  
456 evaluated CDRH3 clonotypes, identifying 6,697 unique clonotypes, none of which  
457 overlapped between the gp120<sup>+</sup> and gp120<sup>-</sup> repertoires even within individuals. The lack  
458 of overlap between the gp120<sup>+</sup> and gp120<sup>-</sup> clonotypes provides further support that the  
459 gp120<sup>-</sup> pool of memory B cells is representative of the non-vaccine induced repertoire.  
460 Among the vaccine-specific clusters (gp120<sup>+</sup>), we identified multiple publicly shared  
461 vaccine-induced clonotypes across multiple subjects. Public clonotypes have been  
462 described in response to HIV-1 infection, including Env-specific *V<sub>H</sub>1-69* and *V<sub>H</sub>1-2*  
463 clonotypes (Setliff et al., 2018; Smith et al., 2018), as well as to other pathogens such as

464 SARS-CoV-2 vaccination and infection, including neutralizing and non-neutralizing  
465 clonotypes (Chen et al., 2021a). These observations further confirm the role of the  
466 immunogen, including recombinant gp120, in selecting the repertoire of the responding B  
467 cells with shared molecular characteristics across subjects.

468         Lastly, independent high-throughput multiplexed next generation single-cell  
469 paired  $V_H/V_L$  sequencing confirmed the enrichment of vaccine-induced gp120<sup>+</sup> B cells co-  
470 expressing the canonical VRC01-class V genes  $V_{H1-2*02}$  paired with  $V_K 3-20$  light chain.  
471 Because inferred germline versions of VRC01-class bnAbs do not bind recombinant  
472 gp120 *in vitro*, it is hypothesized that current HIV-1 Env-based immunogens would not  
473 activate these BCRs. gIVRC01-class targeting immunogens aim to expand B cells with  
474 BCRs containing  $V_{H1-2*02}$  (\*03 or \*04) alleles paired with a short (5 a.a.) CDRL3.  
475 However, only 1 of these  $V_H 1-2*02/V_K 3-20$  expressing gp120<sup>+</sup> B cells had a light chain  
476 with a 5 a.a. CDRL3, and the resulting monoclonal antibody, FH1, was found to not bind  
477 the CD4bs. Therefore, rare vaccine-induced gp120<sup>+</sup> clones with VRC01-class molecular  
478 characteristics can be detected, but it is critical to know the epitope specificity to ascertain  
479 whether the without the corresponding epitope specificity their presence could be  
480 misinterpreted.

481         A major hurdle in driving development of HIV-specific bnAbs by vaccination is the  
482 ability to activate specific B cell lineages and induce affinity maturation through multiple  
483 rounds of mutation and selection. We have shown that the pox-protein vaccine strategy  
484 induced limited somatic hypermutation but did engage a restricted repertoire of vaccine-  
485 reactive B cells. Understanding the lineages of B cells recruited and expanded by non-

486 neutralizing gp120-based HIV vaccine strategies will help inform interpretation of  
487 upcoming germline-targeting HIV vaccine trials and guide vaccine strategies.

488

## 489 **Materials and Methods**

### 490 *Study design, population, products, and procedures*

491 HVTN 100 was a randomized, controlled, double-blind study phase 1-2 clinical trial, which  
492 enrolled 252 healthy HIV-uninfected 18- to 40-year-old participants at 6 sites in South  
493 Africa. The trial was registered with the South African National Clinical Trials Registry  
494 (DOH-27-0215-4796) and ClinicalTrials.gov ([NCT02404311](#)). The investigational  
495 products were ALVAC-HIV (vCP2438) and MF59-adjuvanted bivalent subtype C gp120.  
496 ALVAC-HIV (vCP2438) (viral titer nominal dose of  $10^7$  50% cell culture infectious dose)  
497 expressed the envelope gp120 of the subtype C ZM96.C strain with the gp41  
498 transmembrane sequence of the subtype B LAI strain, as well as *gag* and *protease* from  
499 the subtype B LAI strain. Bivalent subtype C gp120 was a combination of 100 µg each  
500 mixed with the MF59 adjuvant, a squalene oil-in-water emulsion. Placebo for ALVAC-HIV  
501 was a mixture of virus stabilizer and freeze-drying medium reconstituted with 0.4% NaCl,  
502 and placebo for the bivalent subtype C gp120 was 0.9% NaCl for injection. Participants  
503 were randomized in a 5:1 vaccine:placebo ratio to receive ALVAC-HIV at months 0 and  
504 1 followed by 3 administrations of ALVAC-HIV along with bivalent subtype C gp120/MF59  
505 at months 3, 6, and 12, or to receive placebo throughout, all by intramuscular injection.  
506 The research ethics committees of the University of the Witwatersrand, the University of  
507 Cape Town, the University of KwaZulu-Natal, and the Medical Research Council

508 approved the study. All participants gave written informed consent in English or their local  
509 language (Setswana, Sotho, Xhosa, or Zulu). Additional detail about the HVTN 100 study  
510 design, eligibility criteria, participants and their baseline characteristics, randomization,  
511 blinding, study products is available in previous reports (Bekker et al., 2018).

512

### 513 *Binding antibody multiplex assay*

514 HIV-1-specific IgG binding antibody responses were measured with 1:200 dilutions of  
515 serum for presence of IgG that was able to bind to the gp120 immunogens by an HIV-1  
516 binding antibody multiplex assay as described previously (Tomaras et al., 2008; Yates et  
517 al., 2018).

518

### 519 *Intracellular cytokine staining assay*

520 CD4<sup>+</sup> T-cell responses to HIV vaccine insert-matched peptides were measured by  
521 intracellular cytokine staining. The assay detects the production and accumulation of  
522 cytokines on inhibition of intracellular transport after brief cell stimulation. Cryopreserved  
523 peripheral blood mononuclear cells were thawed, rested overnight, and stimulated with a  
524 peptide pools representing the 108.c gp120 ALVAC insert ZM96 gp120 and dimethyl  
525 sulfoxide (DMSO; negative control) or *staphylococcal enterotoxin B* (SEB; positive  
526 control) in the presence of costimulatory antibodies (CD28 and CD49d), and intracellular  
527 transport inhibitors brefeldin A and monensin for 6 h at 37°C. Next, cells were washed  
528 and incubated with EDTA (edetic acid) overnight at 4°C, then stained with a 16-colour  
529 panel (Moncunill et al., 2015), acquired on a BD LSRII flow cytometer (BD Biosciences,



530 San Jose, CA, USA), and analyzed using FlowJo version 9.9.4 (BD, Franklin Lakes, NJ,  
531 USA).

532

### 533 *B cell phenotyping and sorting*

534 Cryopreserved peripheral blood mononuclear cells (PBMCs) were thawed and  
535 transferred to warm thaw media supplemented with benzonase [RPMI, 10% fetal bovine  
536 serum (FBS), 1% L-Glutamine, 1% penicillin-streptomycin, benzonase 50 U/mL] followed  
537 by two additional washes with thaw media. Samples were rested overnight at 37°C 5%  
538 CO<sub>2</sub> and centrifuged at 250 x g for 10 minutes before transferring to a 96-well U-bottom  
539 plate for staining. Cells were pelleted for 3 minutes at 750 x g and resuspended in  
540 Live/Dead marker (AViD, Invitrogen) diluted in 1x PBS. After a 15 to 20-minute incubation  
541 at 4°C, Live/Dead stain was washed off with PBS and PBMCs were incubated for 8-15  
542 minutes at 4°C in anti-human CD4 (BD Biosciences) diluted in stain buffer (PBS, 10%  
543 FBS) to block non-B cell receptor binding of antigens to cells via the Envelope CD4-  
544 binding site. Two probes were prepared in advance at a molar ratio of 4:1, protein to  
545 streptavidin (SA) using HIV-1<sub>1086</sub> Env gp120 conjugated to phycoerythrin (PE)-SA (BD  
546 Biosciences) and AlexaFluor647-SA (Invitrogen). The two tetramers were diluted in stain  
547 buffer and incubated with cells for 30 minutes at 4°C. Samples were washed with stain  
548 buffer, pelleted and resuspended with the following antibody panel diluted in stain buffer  
549 for 15-20 minutes at 4°C: anti-CD3 BV510 (BD Biosciences, UCHT1), anti-CD14 BV510  
550 (BD Biosciences, MfP9), anti-CD56 BV510 (BD Biosciences, NCAM16.2), anti-CD19 PE-  
551 Cy7 (BD Biosciences, SJ25C1), anti-CD27 BV605 (BD Biosciences), anti-IgD FITC (BD  
552 Biosciences), anti-IgM PE-Dazzle594 (BioLegend, MHM-88), anti-IgG BV450 (BD

553 Biosciences), and anti-CD38 PerCP-Cy5.5 (BD Biosciences, HIT2), anti-CD21 BV711  
554 (BD Biosciences), anti-CD16 (BD Biosciences), anti-CD32 (BD Biosciences), anti-CD64  
555 (BD Biosciences), and Mouse IgG. The samples were washed and resuspended in stain  
556 buffer for collection on a BD FACS Aria II (BD Biosciences). Two populations (HIV-1<sub>1086</sub>  
557 Env gp120 positive and HIV-1<sub>1086</sub> Env gp120 negative) were bulk sorted into separate  
558 FACS tubes containing lysis buffer [DNA suspension buffer (Teknova), 10% BSA, 2.5%  
559 RNaseOut] using the following gating strategy: singlets, lymphocytes, Live/Dead-, CD3-,  
560 CD56-, CD14-, CD19+ IgD-, IgG+, 1086 gp120+ or 1086 gp120-.

561 Four of the original subset were selected for further single-cell analysis and were  
562 thawed and stained as described above with the following modifications: IgD BV650 (BD  
563 Biosciences, IA6-2) and IgG BV786 (BD Biosciences, G18-145). Cells that were 1086  
564 gp120+ were single cell sorted using the gating scheme described above into PCR plates  
565 containing lysis buffer. Plates with sorted cells were sealed and frozen on dry ice prior to  
566 transferring to -80°C for storage.

567

#### 568 *Next generation BCR repertoire sequencing*

569 Total RNA from HIV-1<sub>1086</sub> Env gp120 positive and HIV-1<sub>1086</sub> Env gp120 negative bulk-  
570 sorted cells were extracted using the RNeasy Micro Kit (Qiagen) according to the  
571 manufacturer's protocol. Random primed cDNA synthesis was carried out with 10 µL of  
572 total RNA using SuperScript III Reverse Transcriptase Kit (Invitrogen/Thermo Fisher) per  
573 the manufacturer's protocol.

574 For the heavy chain immunoglobulin transcripts were amplified using three  
575 separate primary previously published PCR primer pools (G1-G3, see Supplemental for  
576 list of primers) (Scheid et al., 2011), and used the same 3' Cg CH1 reverse primer in each  
577 PCR reaction consisting of 3µl of cDNA, 1X Amplitaq Gold 360 Master Mix (Applied  
578 Biosystems), 125 nM of the reverse primer and 125 nM of the primer pool. The three IgH  
579 gamma pool amplicons were pooled and 2<sup>nd</sup> round PCR reactions used the pooled  
580 primary PCR IgH amplicons as input. The second round PCR reactions mixes consisted  
581 of 1x AmpliTaQ Gold 360 Master Mix, 500 nM 3' IgGint Reverse Primer and 500 nM of  
582 respective nested Heavy Chain Forward Primer pool. Each round of IGH PCR was  
583 performed for 50 cycles at 94°C for 30 sec, annealing temp specified in NGS primer table  
584 for 30 sec, 72°C for 55 sec (1<sup>st</sup> PCR) or 45 sec (2<sup>nd</sup> PCR) and a final 7 min extension  
585 cycle.

586 Light chain specific (Igλ or Igκ) primer pools were used to amplify light chain specific  
587 immunoglobulin transcripts using previously published primers by two rounds of nested  
588 PCR for each light chain (see Supplemental for list of primers).

589 The primary PCR used 5 ul of cDNA as input and consisted of 500 nM of either  
590 CK-ext Reverse Primer (kappa) or CL-new-ext Reverse Primer (lambda) and either 2.5  
591 µM Kappa Forward Primer Mix or 5 µM Lambda Forward Primer Mix and 1x AmpliTaQ  
592 Gold 360 Master Mix. Nested PCR used 3 µL of the respective primary light chain  
593 amplicon with 1x AmpliTaQ Gold 360 Master Mix, 500 nM of either CK-int Reverse Primer  
594 (kappa) or CL-int Reverse Primer (lambda) and 500 nM of respective Nested  
595 Kappa/Lambda Chain Forward Primer. Each round of light chain PCR used the following  
596 amplification parameters 95°C for 5 min, 50 cycles of 95°C for 30 sec, Annealing

597 temperature specified in primer table, for 30 sec and 72°C for 60 sec, followed by a final  
598 extension of 72°C for 7 min.

599         Nested PCR products for all chains were cleaned and normalized using the  
600 AxyPrep MAG PCR Normalizer Kit (Axygen). Concentrations post-normalization were  
601 confirmed with Qubit dsDNA High Sensitivity Assay Kit (Invitrogen) and diluted where  
602 necessary. Sample pools were constructed by taking 5 µL of each normalized PCR  
603 product from matching samples and combining together into one 1.5 mL tube. Each pool  
604 was mixed by pipetting, and then 5 µL of each pool were added into a 96-well plate, with  
605 each sample pool added to a different well. Sequencing libraries were constructed using  
606 the Nextera XT DNA Library Preparation Kit (Illumina) according to the manufacturer's  
607 protocol. Indexed libraries were normalized, pooled and sequenced using (300 paired  
608 end, version 3 600 cycle kit, Illumina) on a MiSeq Illumina sequencer.

609

#### 610 *Processing and BCR repertoire analysis*

611 Paired-end sequencing reads were assembled by invoking IMMCANTATION/pRESTO  
612 AssemblePairs.py align. Details of BCR assembly have been published previously  
613 (Vander Heiden et al., 2014). Post-assembly filtering of the assembled reads was  
614 accomplished using IMMCANTATION/pRESTO FilterSeq.py, first by removing  
615 assembled reads with length less than 250 nt, and then by filtering out assembled reads  
616 with quality scores less than 30. The remaining assembled reads were written to FASTA  
617 files and submitted to IMG/HighV-QUEST for analysis, including further filtering-out of  
618 reads determined to be unproductive (Lefranc et al., 2009).

619

620 *Single-cell BCR sequencing*

621 Samples were single-cell sorted into 96-well PCR plates containing 10  $\mu$ L of DNA  
622 Suspension Buffer (Teknova) with 1 mg/mL BSA (Sigma-Aldrich) per well. Plates were  
623 sealed and briefly centrifuged before being placed on dry ice. Plates were stored at  $-80^{\circ}\text{C}$   
624 at least overnight before proceeding. Plates were thawed on ice and spun at 800 x g for  
625 1 minute to collect all liquid. Random primed cDNA synthesis was performed on each  
626 well using the SuperScript III Reverse Transcriptase kit with the following changes to the  
627 reaction mix: 5 mM DTT, 200 U SuperScript III and 0.8 mM each dNTP mix.  
628 Amplification proceeded with 2 rounds of nested PCR targeting either the heavy chain,  
629 kappa chain and lambda chain (Supplemental) (Liao et al., 2009; Murugan et al., 2015).  
630 In each reaction, 2  $\mu$ l of the cDNA product were used with 1x AmpliTaq Gold 360 Master  
631 Mix, 300 nM of chain-matched forward primer pool and 200 nM of chain-matched reverse  
632 primer. For heavy chain only, the reverse primer was in the form of two pooled primers  
633 that target either IgG or IgM sequences. Amplification parameters were  $95^{\circ}\text{C}$  for 5 min,  
634 50 cycles of  $95^{\circ}\text{C}$  for 30 sec,  $57^{\circ}\text{C}$  (heavy and kappa) or  $60^{\circ}\text{C}$  (lambda) for 30 sec and  
635  $72^{\circ}\text{C}$  for 55 sec, followed by a final extension of  $72^{\circ}\text{C}$  for 10 min.

636 Nested PCRs were run for each chain using 4  $\mu$ l of unpurified first round PCR  
637 product. Three master mixes consisting of 1x AmpliTaq Gold 360 Master Mix and 200  
638 nM chain-matched reverse primer were created and aliquoted to 96-well PCR plates. The  
639 first round PCR product was added to its chain-matched plate containing the nested PCR  
640 mix. Forward primers for each chain contained 12 nucleotide well IDs (heavy and kappa),  
641 the lambda chain forward primers contained the 8 nucleotide Chain specific plate IDs and

642 reverse primers contained 8 nucleotide plate IDs (heavy and kappa) at their respective 5'  
643 ends, and the lambda chain reverse primer contained 12 nucleotide well IDs. These  
644 barcoded forward primers were then spiked into the nested PCR plate of the matching  
645 chain at 200 nM. Amplification parameters were 95°C for 5 min, 50 cycles of 95°C for 30  
646 sec, 57°C (heavy and kappa) or 60°C (lambda) for 30 sec and 72°C for 50 sec, followed  
647 by a final extension of 72°C for 10 min.

648 The wells from each corresponding heavy, kappa and lambda plate were the  
649 pooled by plate and bead cleaned using 0.6X HighPrep™ PCR Clean-up System SPRI  
650 magnetic beads (MagBio Genomics). Cleaned amplicons were indexed using the Qiaseq  
651 1 step Amplicon DNA library kit (Qiagen) per manufacturer's instructions. Each pooled  
652 and indexed sample was quantified and normalized to 4 nM using Kapa Biosystems  
653 Illumina library quantification kit per the manufacturer's instructions. Indexed and  
654 normalized samples were then pooled and sequenced using (300 paired end, version 3,  
655 600 cycle kit, Illumina) on a MiSeq Illumina sequencer.

656

### 657 *Processing and single-cell BCR sequencing analysis*

658 Single-cell read sequences have a number of attributes, including the originating plate  
659 and well IDs, as well as information about the particular primer associated with a given  
660 reaction within a well (for heavy chains this is the "VREG" identifier while for light chains  
661 it is the "CPRIMER" attribute). The (plate, well, primer) triples serve to partition the read  
662 sequences into groups sharing the same attributes, and each of these read groups is then  
663 independently subjected to the following filtering steps:

664 (1) Retain a read sequence only if its duplicate count is  $\geq 2$ ,

665 (2) Sort the reads within a given group into descending order of read-sequence duplicate

666 count. Note the V gene associated with the most frequently-occurring read sequence in

667 the group and call its duplicate-count COUNT\_MAX and its v-gene VGENE\_MAX (the v-

668 gene, including allele, is computed via IMGT/HighV-QUEST).

669 (3) Consider the other read sequences in the sorted list whose duplicate counts are

670 greater than or equal to  $\text{COUNT\_MAX} / 2$ . If any of these sequences' V genes differ from

671 VGENE\_MAX then remove the entire group of read-sequences from further

672 consideration. Otherwise, retain the read-sequences that have duplicate counts greater

673 than or equal to  $\text{COUNT\_MAX} / 2$  and have common V gene VGENE\_MAX.

674 Step 1 of filtering resulted in 171,010 read sequences with duplicate counts  $\geq 2$  and Steps

675 2 and 3 further reduced this set to 11,758 read sequences.

676 In an effort to eliminate the effects of possible cross contamination, after the above

677 filtering steps had been applied to all read sequence groups, we constructed a

678 mapping of the filtered set's read sequences to the subjects that generated them. If a

679 given read sequence was associated with more than one subject then all read

680 sequence groups sharing that read sequence were filtered out (applying this shared-

681 subject filtering step resulted in 6099 read sequences), and we proceeded to the next

682 step of constructing a single consensus sequence for each group from these filtered

683 retained sequences. Read sequence clustering and consensus resulted in 2449

684 sequences associated with 1911 wells, 339 of which had light chains paired with heavy

685 chains.

686

687 *Statistical analyses*

688 Paired comparisons (e.g., frequencies of V genes between the gp120<sup>+</sup> and gp120<sup>-</sup>  
689 repertoires) were carried out using the Wilcoxon signed-rank (WSR) test. Associations  
690 between paired variables were measured using Spearman's rank correlation coefficients.  
691 P-values adjustment for multiple comparisons was done using Benjamini and Yekutieli's  
692 procedure (Benjamini and Yekutieli, 2001). Adjusted p-values smaller than 0.10 were  
693 considered statistically significant. The diversity of V genes within individuals that  
694 encoded the heavy, kappa, and lambda chains of gp120<sup>+</sup> and gp120<sup>-</sup> B cells were  
695 compared using three diversity indices: the richness index, defined as the number of  
696 distinct V gene families observed in gp120<sup>+</sup> and gp120<sup>-</sup> samples; Shannon diversity,  
697 which is based on the geometric mean of the proportional abundances of observed V  
698 gene frequencies; and Simpson diversity index, which estimates the probability that two  
699 read sequences sampled at random without replacement belong to the same V gene  
700 family. Unlike the richness and Shannon indices, a *lower* Simpson diversity indicates a  
701 *higher* diversity. These indices were compared between classes (gp120<sup>+</sup> and gp120<sup>-</sup> B  
702 cells) using linear mixed models (LMM) adjusting for log-read count and class, and  
703 including subject-specific random intercepts. We used Jaccard's beta diversity index to  
704 evaluate the overlap between the pool of V genes used by the repertoire of different  
705 vaccine-recipients to encode their gp120<sup>+</sup> and gp120<sup>-</sup> heavy and light chains. The Jaccard  
706 index between two repertoires was calculated as  $J_{12} = C_{12}/(C_{12} + U_1 + U_2)$  where  $C_{12}$   
707 denotes the number of genes shared by the two repertoires and  $U_1$  and  $U_2$  are the  
708 numbers of V genes unique to the first and second repertoire, respectively.



709 To assess evolution of BCR due to somatic hypermutation and antigen-driven  
710 selection during germinal center reactions, we calculated the Hamming amino acid  
711 distance between the V region of read sequences and their germline V gene. To examine  
712 convergence of the vaccine-induced repertoire toward  $V_H1-2^*02$ -restricted bnAbs, we  
713 computed distances between read sequences and 5 such HIV-1 bnAbs: VRC01, IOMA,  
714 N6, DH270.1, and VRC-PG04. Distances were separately computed by segment of the  
715 V region: FR1-3, and CDR1-2. Calculations of Hamming distances were restricted to read  
716 sequences with complete coverage of the FR1-FR3 regions.

717 To identify CDRH3 clonotypes, we first computed pairwise distances between AA  
718 CDRH3 read sequences using the Levenshtein distance. We next performed single-  
719 linkage clustering to identify BCR with similar CDRH3 a.a. sequences. The threshold used  
720 to identify the minimum distance between clusters was determined by inspecting the  
721 distribution of the pairwise distances. A sensitivity analysis was carried out to evaluate  
722 the impact of the value of the threshold on the results.

723

## 724 **Acknowledgements**

725 We like to thank the HVTN 100 protocol chairs, protocol teams and participants for making  
726 this study possible.

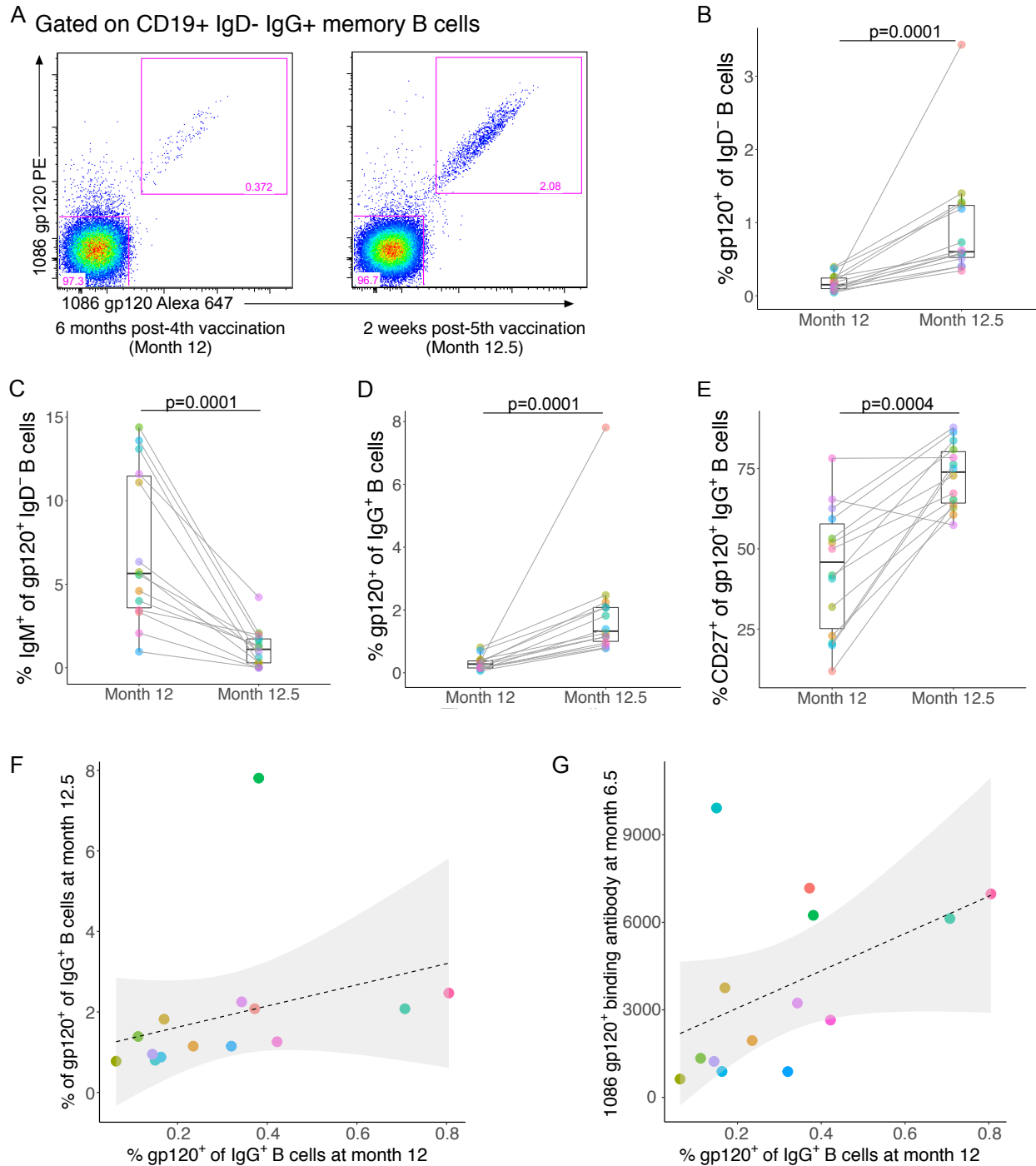
727

## 728 **Funding**

729 Research reported in this publication was supported by the National Institute of Allergy  
730 and Infectious Disease of the National Institutes of Health under award numbers  
731 UM1AI068618, UM1AI068635, and 5R01AI129518, and award number OPP1151646  
732 from the Bill and Melinda Gates Foundation. The content is solely the responsibility of  
733 the authors and does not necessarily represent the official views of the National  
734 Institutes of Health.

735

736 **Figures**

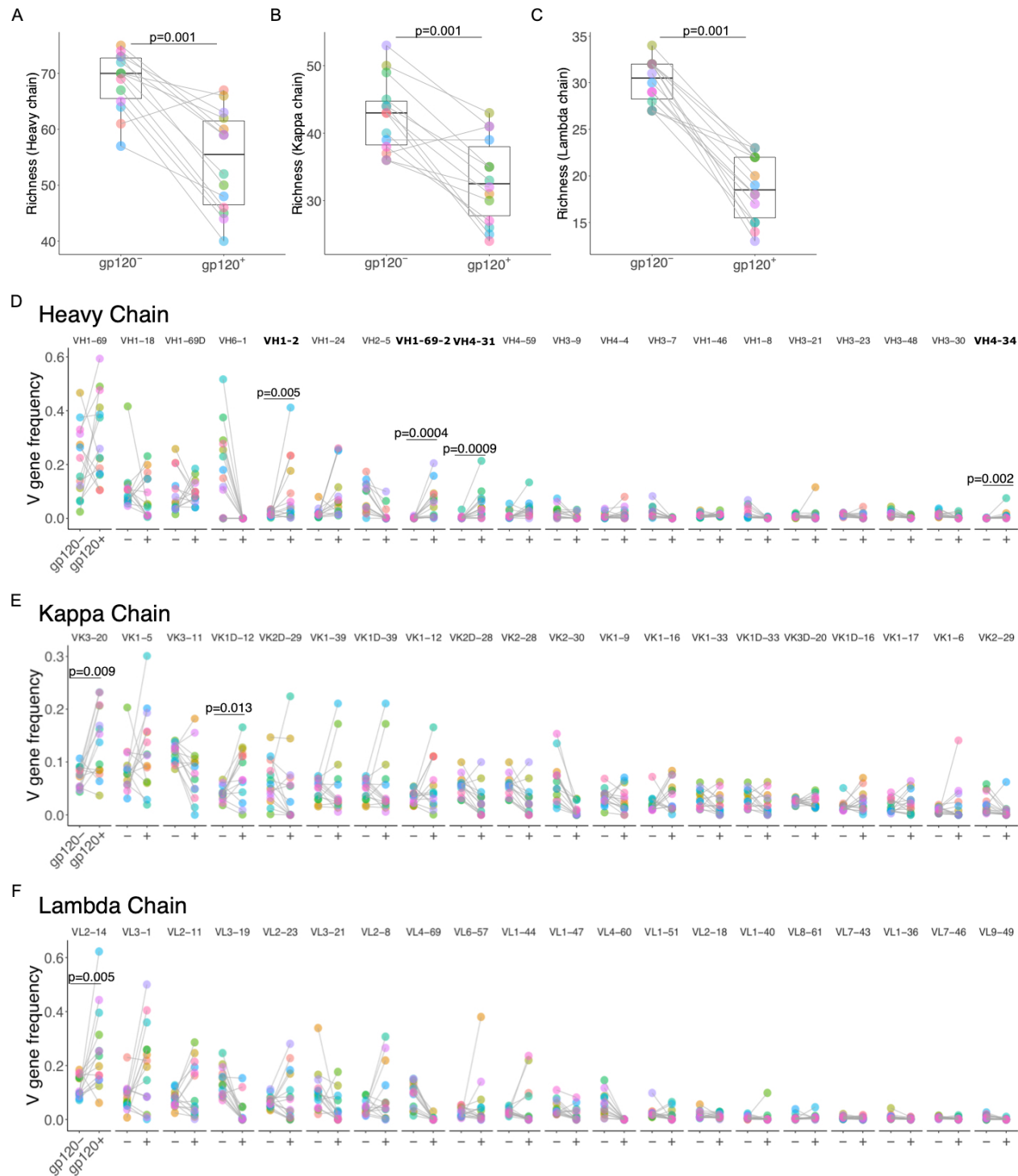


737

738 **Figure 1. Env-specific memory B cells are detected 6 months after vaccination and expand**  
 739 **significantly after boost immunization (n=14).** Example of peripheral blood Env-specific  
 740 memory B cells from one HVTN 100 participant gated on CD3<sup>-</sup>, CD14<sup>-</sup>, CD56<sup>-</sup>, CD19<sup>+</sup>, IgD<sup>-</sup>, and

741 IgG<sup>+</sup> live lymphocytes at A) month 12 (6 months post-4<sup>th</sup> vaccination and day of boost) and month  
742 12.5 (2 weeks after 5<sup>th</sup> vaccination). Env-specificity is determined by double-staining with  
743 fluorescently-labeled 1086 gp120 probes. B) The frequencies of 1086 gp120<sup>+</sup> IgD<sup>-</sup> B cells at  
744 month 12 and 12.5. C) The proportion of gp120<sup>+</sup> IgD<sup>-</sup> B cells that were IgM<sup>+</sup>. D) The frequencies  
745 of 1086 gp120<sup>+</sup> IgG<sup>+</sup> B cells at month 12 and 12.5. E) The proportion of gp120<sup>+</sup> IgG<sup>+</sup> B cells that  
746 were CD27<sup>+</sup>. Boxes indicate the median (at line) and interquartile range. Statistical significance  
747 was determined using Wilcoxon signed rank test. P-values less than 0.05 were considered  
748 significant. The frequencies of 1086 gp120<sup>+</sup> IgG<sup>+</sup> B cells measured in PBMCs at month 12  
749 correlated with the F) with the frequencies of 1086 gp120<sup>+</sup> IgG<sup>+</sup> B cells at visit month 12.5 and G)  
750 corresponding vaccine-matched serum binding antibody responses (net MFI by binding antibody  
751 multiplex assay) at month 6.5 (2 weeks post-4<sup>th</sup>).

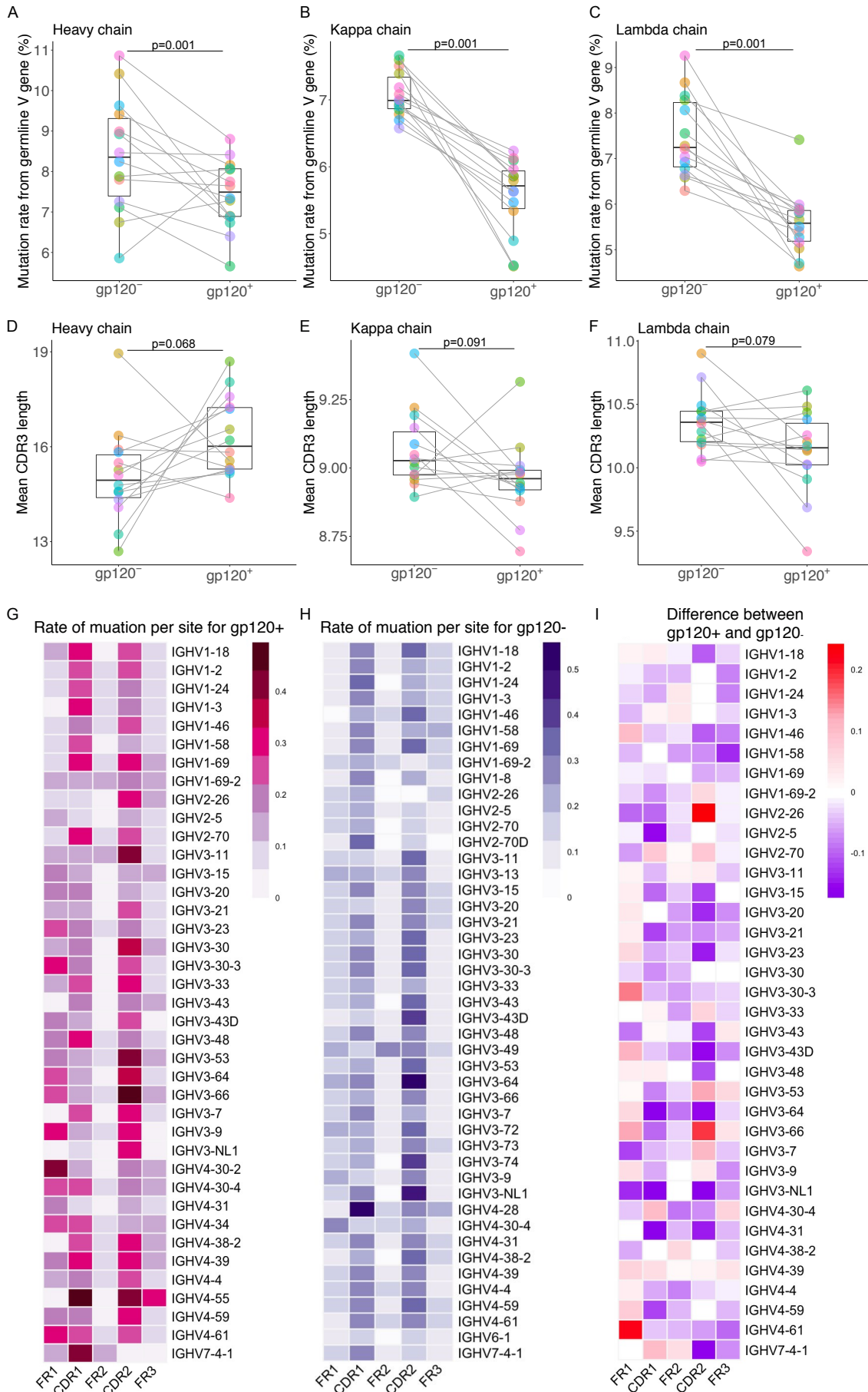
752



753

754 **Figure 2. Diversity and V-gene distribution of paired vaccine gp120-specific and non-**  
 755 **specific IgG<sup>+</sup> B cell receptor repertoires (n=14).** Diversity of V genes of gp120<sup>+</sup> and gp120<sup>-</sup>  
 756 IgG<sup>+</sup> B cell for V<sub>H</sub>, V<sub>K</sub>, and V<sub>L</sub> BCR sequences at month 12.5. A) heavy B) kappa and C) lambda  
 757 chain diversity indices as calculated by Richness score among gp120<sup>-/+</sup> IgG<sup>+</sup> B cells. The paired  
 758 frequencies per participant of 1086 gp120<sup>+</sup> and gp120<sup>-</sup> IgG BCR sequences expressing the top

759 twenty: D)  $V_H$ , E)  $V_K$  or F)  $V_L$  genes ranked by average frequency. D) Only the  $V_H$  genes that  
760 were significantly enriched in the gp120<sup>+</sup> compared to gp120<sup>-</sup> after adjusting for a FDR < 10% are  
761 indicated (in bold). All E)  $V_K$  and F)  $V_L$  genes that were enriched in the gp120<sup>+</sup> compared to gp120<sup>-</sup>  
762 before adjusting for FDR are indicated; no  $V_K$  and  $V_L$  genes were significant after controlling for  
763 FDR. Comparisons were conducted by Wilcoxon signed rank tests and show the unadjusted p  
764 values.



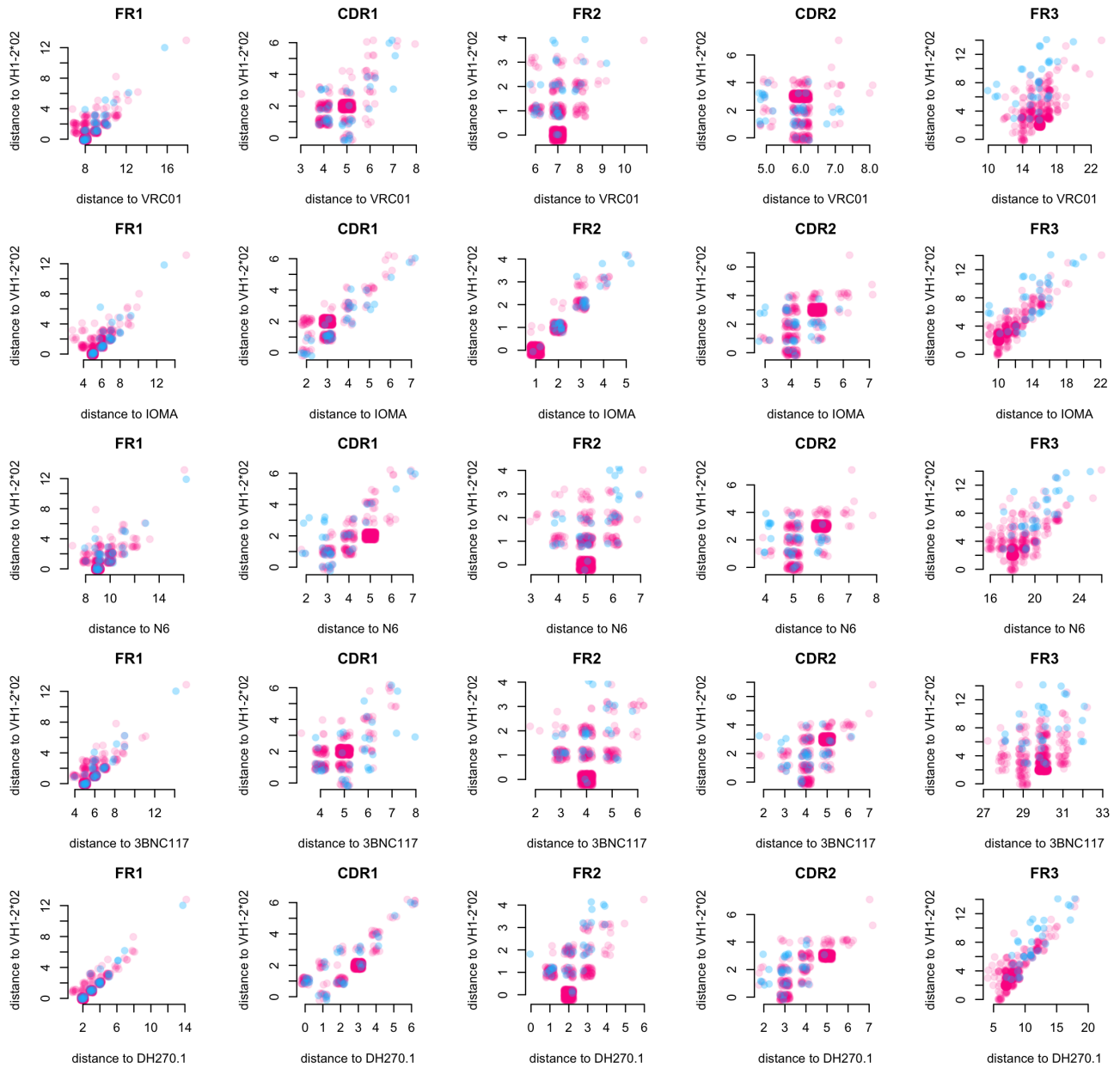
766 **Figure 3. Mutation rate and CDR3 length among gp120<sup>+</sup> and gp120<sup>-</sup> V<sub>H</sub>, V<sub>K</sub> and V<sub>L</sub> BCR**  
767 **sequences at month 12.5 (n=14).** Mean per subject percent mutation of the nucleic acid V gene  
768 sequences for the A) V<sub>H</sub>, B) V<sub>K</sub> and C) V<sub>L</sub>. Mean CDR3 lengths per subject for the D) heavy, E)  
769 kappa and F) lambda chains. Paired differences and corresponding P values were assessed  
770 using Wilcoxon signed rank tests. Mean amino acid (a.a.) mutations per site in the respective  
771 framework (FR) or complementary determining (CDR) regions of the V<sub>H</sub> sequence split out by V<sub>H</sub>  
772 gene among G) gp120<sup>+</sup> or H) gp120<sup>-</sup> IgG<sup>+</sup> B cells. I) the difference in a.a. mutation rate per V<sub>H</sub>  
773 region between the gp120<sup>+</sup> and gp120<sup>-</sup> V<sub>H</sub> sequences; cells in red indicate higher mutation rates  
774 in the gp120<sup>+</sup> than in the gp120<sup>-</sup> repertoires, and vice versa for those in blue.

775



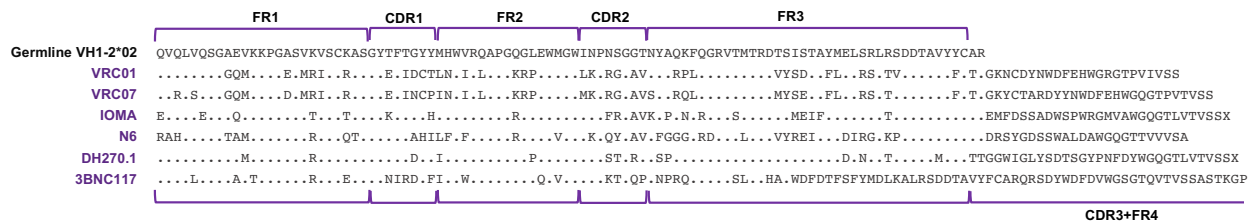
776

A



77

B



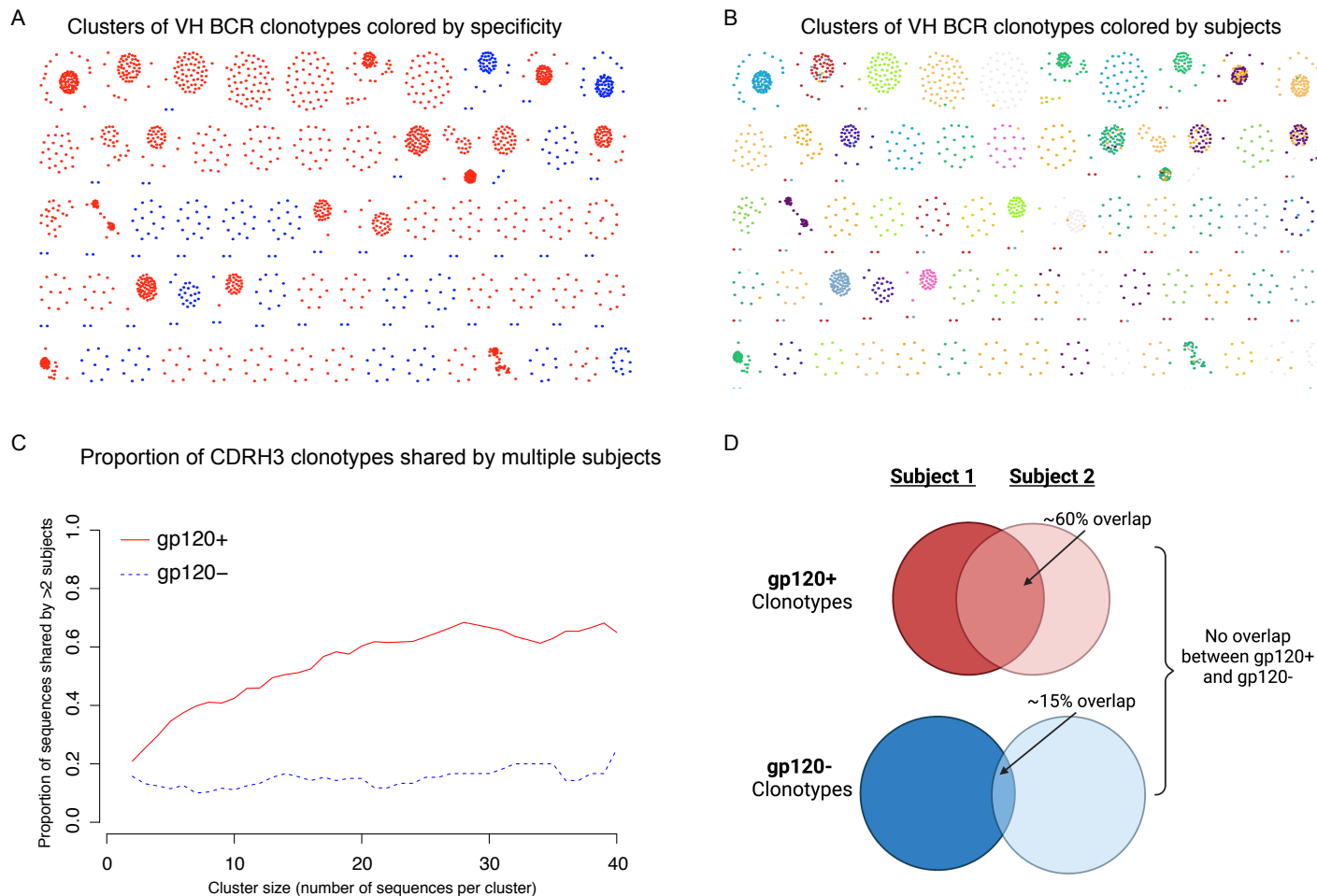
778

779 **Figure 4. Mutations in the VH1-2\*02-expressing Env-positive B cells do not lead to**  
 780 **increased similarity to VH1-2\*02 bnAbs. A) Among  $V_{H1-2*02}$ -encoded  $gp120^+$  (pink) and**

781 gp120<sup>-</sup> (blue) BCR sequences, the distance from the germline V<sub>H</sub>1-2\*02 (y-axes) plotted against  
782 the distance from a subset of mature V<sub>H</sub>1-2-expressing bnAbs (from top to bottom rows): VRC01,  
783 IOMA, N6, 3BNC117, and DH270.1 per V gene framework (FR) or complementary determining  
784 (CDR) region. Each dot represents one BCR sequence; their color indicates the repertoire they  
785 were sampled from (red: gp120<sup>+</sup>; blue: gp120<sup>-</sup>). B) Comparison of the amino acid sequence of 6  
786 bnAbs encoded by the V<sub>H</sub>1-2\*02 germline V gene allele. Positions with a mutation relative to the  
787 germline V<sub>H</sub>1-2\*02 sequence (top) are indicated.

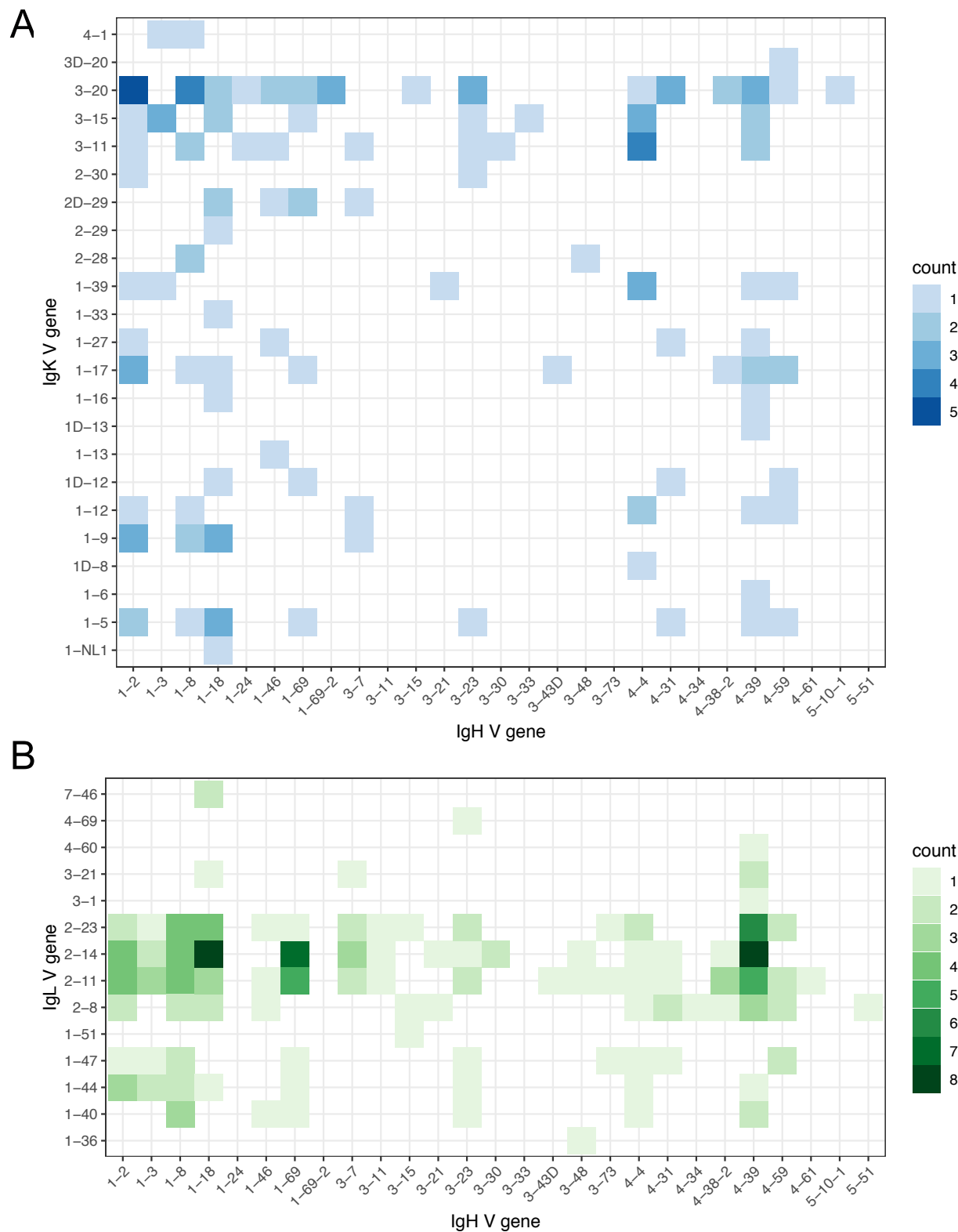
788

789



790

791 **Figure 5. Shared B cell clonotypes were detected across subjects among the gp120+, but**  
792 **not gp120- BCR repertoires (n=14).** Clonotypes were determined by clustering all CDRH3  
793 sequences, regardless of class (gp120<sup>+</sup> and gp120<sup>-</sup>) and subjects. A) Representative subset of  
794 the clusters of gp120<sup>+</sup> and gp120<sup>-</sup> BCR V<sub>H</sub> clonotypes, colored by the specificity of the BCR:  
795 gp120<sup>+</sup> (red) and gp120<sup>-</sup> (blue). B) Same representative subset of the clusters of gp120<sup>+</sup> and  
796 gp120<sup>-</sup> BCR V<sub>H</sub> clonotypes, instead colored by subject that the sequence came from. C) The  
797 proportion of clonotypes shared by multiple subjects as a function of the size of the cluster of  
798 clonotypes for gp120<sup>+</sup> and gp120<sup>-</sup> B cells, respectively. D) A schematic of the average overlap in  
799 clonotypes between any two subjects for the gp120<sup>+</sup> and gp120<sup>-</sup> repertoires.



800

801 **Figure 6. Single-cell BCR sequencing identifies immunodominant heavy and light chain**

802 **pairs.** Counts of paired heavy and light chain sequences from single-cell gp120<sup>+</sup> IgG<sup>+</sup> B cells

803 from 4 vaccinated HVTN 100 subjects with expanded  $V_H1-2^{*02}$  gp120<sup>+</sup> B cells for A)  $V_H/V_K$  and  
804 B)  $V_H/V_L$ , respectively.

## 805 References

- 806 Balazs, A.B., Ouyang, Y., Hong, C.M., Chen, J., Nguyen, S.M., Rao, D.S., An, D.S., and  
807 Baltimore, D. (2014). Vectored immunoprophylaxis protects humanized mice from  
808 mucosal HIV transmission. *Nature medicine* 20, 296-300.
- 809 Basu, M., Piepenbrink, M.S., Francois, C., Roche, F., Zheng, B., Spencer, D.A.,  
810 Hessel, A.J., Fucile, C.F., Rosenberg, A.F., Bunce, C.A., *et al.* (2020). Persistence of  
811 HIV-1 Env-Specific Plasmablast Lineages in Plasma Cells after Vaccination in Humans.  
812 *Cell Rep Med* 1.
- 813 Bekker, L.G., Moodie, Z., Grunenberg, N., Laher, F., Tomaras, G.D., Cohen, K.W.,  
814 Allen, M., Malahleha, M., Mngadi, K., Daniels, B., *et al.* (2018). Subtype C ALVAC-HIV  
815 and bivalent subtype C gp120/MF59 HIV-1 vaccine in low-risk, HIV-uninfected, South  
816 African adults: a phase 1/2 trial. *Lancet HIV* 5, e366-e378.
- 817 Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple  
818 testing under dependency. *The Annals of Statistics* 29, 1165-1188, 1124.
- 819 Bhat, N.M., Lee, L.M., van Vollenhoven, R.F., Teng, N.N., and Bieber, M.M. (2002).  
820 VH4-34 encoded antibody in systemic lupus erythematosus: effect of isotype. *J*  
821 *Rheumatol* 29, 2114-2121.
- 822 Bonsignori, M., Kreider, E.F., Fera, D., Meyerhoff, R.R., Bradley, T., Wiehe, K., Alam,  
823 S.M., Aussedat, B., Walkowicz, W.E., Hwang, K.K., *et al.* (2017). Staged induction of  
824 HIV-1 glycan-dependent broadly neutralizing antibodies. *Sci Transl Med* 9.
- 825 Briney, B., Sok, D., Jardine, J.G., Kulp, D.W., Skog, P., Menis, S., Jacak, R.,  
826 Kalyuzhniy, O., de Val, N., Sesterhenn, F., *et al.* (2016). Tailored Immunogens Direct  
827 Affinity Maturation toward HIV Neutralizing Antibodies. *Cell* 166, 1459-1470 e1411.
- 828 Chen, E.C., Gilchuk, P., Zost, S.J., Suryadevara, N., Winkler, E.S., Cabel, C.R.,  
829 Binshtein, E., Chen, R., Sutton, R.E., Rodriguez, J., *et al.* (2021a). Convergent antibody  
830 responses to the SARS-CoV-2 spike protein in convalescent and vaccinated individuals.  
831 *Cell Reports*, 109604.
- 832 Chen, E.C., Gilchuk, P., Zost, S.J., Suryadevara, N., Winkler, E.S., Cabel, C.R.,  
833 Binshtein, E., Chen, R.E., Sutton, R.E., Rodriguez, J., *et al.* (2021b). Convergent  
834 antibody responses to the SARS-CoV-2 spike protein in convalescent and vaccinated  
835 individuals. *Cell Rep*, 109604.
- 836 Corey, L., Gilbert, P.B., Juraska, M., Montefiori, D.C., Morris, L., Karuna, S.T.,  
837 Edupuganti, S., Mgodhi, N.M., deCamp, A.C., Rudnicki, E., *et al.* (2021). Two  
838 Randomized Trials of Neutralizing Antibodies to Prevent HIV-1 Acquisition. *N Engl J*  
839 *Med* 384, 1003-1014.
- 840 Doria-Rose, N.A., Klein, R.M., Daniels, M.G., O'Dell, S., Nason, M., Lapedes, A.,  
841 Bhattacharya, T., Migueles, S.A., Wyatt, R.T., Korber, B.T., *et al.* (2010). Breadth of  
842 human immunodeficiency virus-specific neutralizing activity in sera: clustering analysis  
843 and association with clinical variables. *Journal of virology* 84, 1631-1636.
- 844 Easterhoff, D., Moody, M.A., Fera, D., Cheng, H., Ackerman, M., Wiehe, K., Saunders,  
845 K.O., Pollara, J., Vandergrift, N., Parks, R., *et al.* (2017). Boosting of HIV envelope CD4

- 846 binding site antibodies with long variable heavy third complementarity determining  
847 region in the randomized double blind RV305 HIV-1 vaccine trial. *PLoS Pathog* *13*,  
848 e1006182.
- 849 Gray, G.E., Bekker, L.G., Laher, F., Malahleha, M., Allen, M., Moodie, Z., Grunenber,  
850 N., Huang, Y., Grove, D., Prigmore, B., *et al.* (2021). Vaccine Efficacy of ALVAC-HIV  
851 and Bivalent Subtype C gp120-MF59 in Adults. *N Engl J Med* *384*, 1089-1100.
- 852 Gristick, H.B., von Boehmer, L., West, A.P., Jr., Schamber, M., Gazumyan, A.,  
853 Golijanin, J., Seaman, M.S., Fatkenheuer, G., Klein, F., Nussenzweig, M.C., *et al.*  
854 (2016). Natively glycosylated HIV-1 Env structure reveals new mode for antibody  
855 recognition of the CD4-binding site. *Nat Struct Mol Biol* *23*, 906-915.
- 856 Havenar-Daughton, C., Sarkar, A., Kulp, D.W., Toy, L., Hu, X., Deresa, I., Kalyuzhniy,  
857 O., Kaushik, K., Upadhyay, A.A., Menis, S., *et al.* (2018). The human naive B cell  
858 repertoire contains distinct subclasses for a germline-targeting HIV-1 vaccine  
859 immunogen. *Sci Transl Med* *10*.
- 860 Hraber, P., Seaman, M.S., Bailer, R.T., Mascola, J.R., Montefiori, D.C., and Korber,  
861 B.T. (2014). Prevalence of broadly neutralizing antibody responses during chronic HIV-1  
862 infection. *Aids* *28*, 163-169.
- 863 Huang, C.C., Venturi, M., Majeed, S., Moore, M.J., Phogat, S., Zhang, M.Y., Dimitrov,  
864 D.S., Hendrickson, W.A., Robinson, J., Sodroski, J., *et al.* (2004). Structural basis of  
865 tyrosine sulfation and VH-gene usage in antibodies that recognize the HIV type 1  
866 coreceptor-binding site on gp120. *Proceedings of the National Academy of Sciences of*  
867 *the United States of America* *101*, 2706-2711.
- 868 Huang, J., Kang, B.H., Ishida, E., Zhou, T., Griesman, T., Sheng, Z., Wu, F., Doria-  
869 Rose, N.A., Zhang, B., McKee, K., *et al.* (2016). Identification of a CD4-Binding-Site  
870 Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity* *45*, 1108-  
871 1121.
- 872 Jardine, J., Julien, J.P., Menis, S., Ota, T., Kalyuzhniy, O., McGuire, A., Sok, D., Huang,  
873 P.S., MacPherson, S., Jones, M., *et al.* (2013). Rational HIV immunogen design to  
874 target specific germline B cell receptors. *Science* *340*, 711-716.
- 875 Jardine, J.G., Kulp, D.W., Havenar-Daughton, C., Sarkar, A., Briney, B., Sok, D.,  
876 Sesterhenn, F., Ereno-Orbea, J., Kalyuzhniy, O., Deresa, I., *et al.* (2016a). HIV-1  
877 broadly neutralizing antibody precursor B cells revealed by germline-targeting  
878 immunogen. *Science* *351*, 1458-1463.
- 879 Jardine, J.G., Ota, T., Sok, D., Pauthner, M., Kulp, D.W., Kalyuzhniy, O., Skog, P.D.,  
880 Thinnes, T.C., Bhullar, D., Briney, B., *et al.* (2015). HIV-1 VACCINES. Priming a broadly  
881 neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science*  
882 *349*, 156-161.
- 883 Jardine, J.G., Sok, D., Julien, J.P., Briney, B., Sarkar, A., Liang, C.H., Scherer, E.A.,  
884 Henry Dunand, C.J., Adachi, Y., Diwanji, D., *et al.* (2016b). Minimally Mutated HIV-1  
885 Broadly Neutralizing Antibodies to Guide Reductionist Vaccine Design. *PLoS Pathog*  
886 *12*, e1005815.

- 887 Kim, S.I., Noh, J., Kim, S., Choi, Y., Yoo, D.K., Lee, Y., Lee, H., Jung, J., Kang, C.K.,  
888 Song, K.H., *et al.* (2021). Stereotypic neutralizing VH antibodies against SARS-CoV-2  
889 spike protein receptor binding domain in patients with COVID-19 and healthy  
890 individuals. *Sci Transl Med* 13.
- 891 Kwong, P.D., and Mascola, J.R. (2012). Human antibodies that neutralize HIV-1:  
892 identification, structures, and B cell ontogenies. *Immunity* 37, 412-425.
- 893 Kwong, P.D., Mascola, J.R., and Nabel, G.J. (2013). Broadly neutralizing antibodies and  
894 the search for an HIV-1 vaccine: the end of the beginning. *Nature reviews Immunology*  
895 13, 693-701.
- 896 Lefranc, M.P., Giudicelli, V., Ginestoux, C., Jabado-Michaloud, J., Folch, G.,  
897 Bellahcene, F., Wu, Y., Gemrot, E., Brochet, X., Lane, J., *et al.* (2009). IMGT, the  
898 international ImMunoGeneTics information system. *Nucleic Acids Res* 37, D1006-1012.
- 899 Li, Y., Spellerberg, M.B., Stevenson, F.K., Capra, J.D., and Potter, K.N. (1996). The I  
900 binding specificity of human VH 4-34 (VH 4-21) encoded antibodies is determined by  
901 both VH framework region 1 and complementarity determining region 3. *J Mol Biol* 256,  
902 577-589.
- 903 Liao, H.X., Levesque, M.C., Nagel, A., Dixon, A., Zhang, R., Walter, E., Parks, R.,  
904 Whitesides, J., Marshall, D.J., Hwang, K.K., *et al.* (2009). High-throughput isolation of  
905 immunoglobulin genes from single human B cells and expression as monoclonal  
906 antibodies. *J Virol Methods* 158, 171-179.
- 907 McGuire, A.T. (2019). Targeting broadly neutralizing antibody precursors: a naive  
908 approach to vaccine design. *Curr Opin HIV AIDS* 14, 294-301.
- 909 McGuire, A.T., Gray, M.D., Dosenovic, P., Gitlin, A.D., Freund, N.T., Petersen, J.,  
910 Correnti, C., Johnsen, W., Kegel, R., Stuart, A.B., *et al.* (2016). Specifically modified  
911 Env immunogens activate B-cell precursors of broadly neutralizing HIV-1 antibodies in  
912 transgenic mice. *Nat Commun* 7, 10618.
- 913 McGuire, A.T., Hoot, S., Dreyer, A.M., Lippy, A., Stuart, A., Cohen, K.W., Jardine, J.,  
914 Menis, S., Scheid, J.F., West, A.P., *et al.* (2013). Engineering HIV envelope protein to  
915 activate germline B cell receptors of broadly neutralizing anti-CD4 binding site  
916 antibodies. *J Exp Med* 210, 655-663.
- 917 Mikell, I., Sather, D.N., Kalams, S.A., Altfeld, M., Alter, G., and Stamatatos, L. (2011).  
918 Characteristics of the Earliest Cross-Neutralizing Antibody Response to HIV-1. *PLoS*  
919 *Pathog* 7, e1001251.
- 920 Moncunill, G., Dobano, C., McElrath, M.J., and De Rosa, S.C. (2015). OMIP-025:  
921 evaluation of human T- and NK-cell responses including memory and follicular helper  
922 phenotype by intracellular cytokine staining. *Cytometry A* 87, 289-292.
- 923 Murugan, R., Imkeller, K., Busse, C.E., and Wardemann, H. (2015). Direct high-  
924 throughput amplification and sequencing of immunoglobulin genes from single human B  
925 cells. *Eur J Immunol* 45, 2698-2700.
- 926 Piantadosi, A., Panteleeff, D., Blish, C.A., Baeten, J.M., Jaoko, W., McClelland, R.S.,  
927 and Overbaugh, J. (2009). Breadth of neutralizing antibody response to human



- 928 immunodeficiency virus type 1 is affected by factors early in infection but does not  
929 influence disease progression. *Journal of virology* 83, 10269-10274.
- 930 Robb, M.L., Rerks-Ngarm, S., Nitayaphan, S., Pitisuttithum, P., Kaewkungwal, J.,  
931 Kunasol, P., Khamboonruang, C., Thongcharoen, P., Morgan, P., Benenson, M., *et al.*  
932 (2012). Risk behaviour and time as covariates for efficacy of the HIV vaccine regimen  
933 ALVAC-HIV (vCP1521) and AIDSVAX B/E: a post-hoc analysis of the Thai phase 3  
934 efficacy trial RV 144. *Lancet Infect Dis* 12, 531-537.
- 935 Roskin, K.M., Jackson, K.J.L., Lee, J.Y., Hoh, R.A., Joshi, S.A., Hwang, K.K.,  
936 Bonsignori, M., Pedroza-Pacheco, I., Liao, H.X., Moody, M.A., *et al.* (2020). Aberrant B  
937 cell repertoire selection associated with HIV neutralizing antibody breadth. *Nat Immunol*  
938 21, 199-209.
- 939 Sather, D.N., Carbonetti, S., Kehayia, J., Kraft, Z., Mikell, I., Scheid, J.F., Klein, F., and  
940 Stamatatos, L. (2012). Broadly neutralizing antibodies developed by an HIV-positive  
941 elite neutralizer exact a replication fitness cost on the contemporaneous virus. *Journal*  
942 *of virology* 86, 12676-12685.
- 943 Scheid, J.F., Mouquet, H., Ueberheide, B., Diskin, R., Klein, F., Oliveira, T.Y., Pietzsch,  
944 J., Fenyo, D., Abadir, A., Velinzon, K., *et al.* (2011). Sequence and structural  
945 convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 333,  
946 1633-1637.
- 947 Setliff, I., McDonnell, W.J., Raju, N., Bombardi, R.G., Murji, A.A., Scheepers, C., Ziki,  
948 R., Mynhardt, C., Shepherd, B.E., Mamchak, A.A., *et al.* (2018). Multi-Donor  
949 Longitudinal Antibody Repertoire Sequencing Reveals the Existence of Public Antibody  
950 Clonotypes in HIV-1 Infection. *Cell Host Microbe* 23, 845-854 e846.
- 951 Shingai, M., Donau, O.K., Plishka, R.J., Buckler-White, A., Mascola, J.R., Nabel, G.J.,  
952 Nason, M.C., Montefiori, D., Moldt, B., Poignard, P., *et al.* (2014). Passive transfer of  
953 modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block  
954 SHIV infection in macaques. *J Exp Med* 211, 2061-2074.
- 955 Simek, M.D., Rida, W., Priddy, F.H., Pung, P., Carrow, E., Laufer, D.S., Lehrman, J.K.,  
956 Boaz, M., Tarragona-Fiol, T., Miiro, G., *et al.* (2009). Human immunodeficiency virus  
957 type 1 elite neutralizers: individuals with broad and potent neutralizing activity identified  
958 by using a high-throughput neutralization assay together with an analytical selection  
959 algorithm. *Journal of virology* 83, 7337-7348.
- 960 Smith, S.A., Burton, S.L., Kilembe, W., Lakhi, S., Karita, E., Price, M., Allen, S., and  
961 Derdeyn, C.A. (2018). VH1-69 Utilizing Antibodies Are Capable of Mediating Non-  
962 neutralizing Fc-Mediated Effector Functions Against the Transmitted/Founder gp120.  
963 *Front Immunol* 9, 3163.
- 964 Sok, D., Briney, B., Jardine, J.G., Kulp, D.W., Menis, S., Pauthner, M., Wood, A., Lee,  
965 E.C., Le, K.M., Jones, M., *et al.* (2016). Priming HIV-1 broadly neutralizing antibody  
966 precursors in human Ig loci transgenic mice. *Science* 353, 1557-1560.
- 967 Sok, D., van Gils, M.J., Pauthner, M., Julien, J.P., Saye-Francisco, K.L., Hsueh, J.,  
968 Briney, B., Lee, J.H., Le, K.M., Lee, P.S., *et al.* (2014). Recombinant HIV envelope  
969 trimer selects for quaternary-dependent antibodies targeting the trimer apex.

- 970 Proceedings of the National Academy of Sciences of the United States of America *111*,  
971 17624-17629.
- 972 Tomaras, G.D., Yates, N.L., Liu, P., Qin, L., Fouda, G.G., Chavez, L.L., Decamp, A.C.,  
973 Parks, R.J., Ashley, V.C., Lucas, J.T., *et al.* (2008). Initial B-cell responses to  
974 transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M  
975 (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective  
976 control of initial viremia. *J Virol* *82*, 12449-12463.
- 977 van Gils, M.J., Euler, Z., Schweighardt, B., Wrin, T., and Schuitemaker, H. (2009).  
978 Prevalence of cross-reactive HIV-1-neutralizing activity in HIV-1-infected patients with  
979 rapid or slow disease progression. *Aids* *23*, 2405-2414.
- 980 Vander Heiden, J.A., Yaari, G., Uduman, M., Stern, J.N., O'Connor, K.C., Hafler, D.A.,  
981 Vigneault, F., and Kleinstein, S.H. (2014). pRESTO: a toolkit for processing high-  
982 throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics* *30*,  
983 1930-1932.
- 984 West, A.P., Jr., Diskin, R., Nussenzweig, M.C., and Bjorkman, P.J. (2012). Structural  
985 basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding  
986 site of HIV-1 gp120. *Proceedings of the National Academy of Sciences of the United*  
987 *States of America* *109*, E2083-2090.
- 988 Yates, N.L., deCamp, A.C., Korber, B.T., Liao, H.X., Irene, C., Pinter, A., Peacock, J.,  
989 Harris, L.J., Sawant, S., Hraber, P., *et al.* (2018). HIV-1 Envelope Glycoproteins from  
990 Diverse Clades Differentiate Antibody Responses and Durability among Vaccinees. *J*  
991 *Virol* *92*.
- 992