

1 **Naive human B cells engage the receptor binding domain of SARS-CoV-2,**
2 **variants of concern, and related sarbecoviruses**
3
4
5
6

7 Jared Feldman^{1†}, Julia Bals^{1†}, Clara G. Altomare², Kerri St. Denis¹, Evan C. Lam¹, Blake M.
8 Hauser¹, Larance Ronsard¹, Maya Sangesland¹, Thalia Bracamonte Moreno¹, Vintus Okonkwo¹,
9 Nathania Hartojo¹ Alejandro B. Balazs¹, Goran Bajic², Daniel Lingwood^{1*} and Aaron G.
10 Schmidt^{1,3*}
11
12
13
14
15

16 ¹Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, 02139, USA
17

18 ²Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029
19

20 ³Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA
21

22 *Correspondence to: Email: dlingwood@mgh.harvard.edu (D.L.); aschmidt@crystal.harvard.edu
23 (A.G.S)
24

25 †These authors contributed equally to this work.

26 **ABSTRACT**

27 Exposure to a pathogen elicits an adaptive immune response aimed to control and eradicate.
28 Interrogating the abundance and specificity of the naive B cell repertoire contributes to
29 understanding how to potentially elicit protective responses. Here, we isolated naive B cells from
30 8 seronegative human donors targeting the SARS-CoV-2 receptor-binding domain (RBD). Single
31 B cell analysis showed diverse gene usage with no restricted complementarity determining region
32 lengths. We show that recombinant antibodies engage SARS-CoV-2 RBD, circulating variants,
33 and pre-emergent coronaviruses. Representative antibodies signal in a B cell activation assay and
34 can be affinity matured through directed evolution. Structural analysis of a naive antibody in
35 complex with spike shows a conserved mode of recognition shared with infection-induced
36 antibodies. Lastly, both naive and affinity-matured antibodies can neutralize SARS-CoV-2.
37 Understanding the naive repertoire may inform potential responses recognizing variants or
38 emerging coronaviruses enabling the development of pan-coronavirus vaccines aimed at engaging
39 germline responses.

40

41

42 **One Sentence Summary:** Isolation of antibody germline precursors targeting the receptor binding
43 domain of coronaviruses.

44

45 **MAIN TEXT**

46 Initial exposure to viral antigens by natural infection or vaccination primes an immune response
47 and often establishes an immune memory which can prevent or control future infections. The naive
48 repertoire contains potential B cell receptor (BCR) rearrangements capable of recognizing these
49 antigens, often the surface-exposed glycoproteins. An early step in generating humoral immunity
50 involves activation of these naive B cells through recognition of a cognate antigen (*1*) which in
51 turn can lead to affinity maturation through somatic hypermutation (SHM) and subsequent
52 differentiation (*2*). The initial engagement of the naive repertoire begins this cascade and often
53 coincides with the eventual generation of a protective or neutralizing antibody response (*3, 4*).

54

55 For SARS-CoV-2, the etiological agent of COVID-19, the development of a neutralizing antibody
56 response after primary infection or vaccination is associated with protection against reinfection in
57 non-human primates (*5-9*). In humans, the presence of neutralizing antibodies can predict disease
58 severity and survival after primary SARS-CoV-2 infection (*10*) or vaccination (*11*) and correlates
59 with protection from symptomatic secondary infection (*12, 13*). Further, the two arms of humoral
60 immune memory, long-lived bone marrow plasma cells (*14*) and circulating memory B cells (*15-*
61 *19*), were induced by natural infection in humans and may persist for at least 8 months after
62 primary infection providing potentially durable long-term protection. Comparable levels of
63 neutralizing antibody titers were present in convalescent COVID-19 subjects and vaccine
64 recipients (*20-22*) further supporting the role of adaptive immune responses in helping to control
65 and prevent disease severity.

66

67 Both infection- and vaccine-elicited antibodies target the major envelope glycoprotein, spike,
68 present on the virion surface (23). A substantial component of the neutralizing response engages
69 the receptor binding domain (RBD) (24-29) and does so by directly blocking interactions with the
70 viral receptor ACE2 (30-35). Isolated RBD-directed monoclonal antibodies derive from diverse
71 heavy- and light-chain variable gene segments suggesting that multiple biochemical solutions for
72 developing RBD-directed antibodies are encoded within the human B-cell repertoire (24, 26, 29,
73 36). Potential immunogenicity of this antigenic site is based on the human naive B cell repertoire,
74 and the overall frequency of naive BCRs that have some level of intrinsic affinity to stimulate their
75 elicitation (37-40). However, antigen-specificity of naive B cells is largely undefined.

76

77 Traditional approaches for studying antigen-specific naive B cells include bioinformatic mining
78 of available BCR datasets and inference of likely germline precursors by “germline-reverting”
79 mature BCR sequences, which can be limited by the availability of heavy and light chain paired
80 sequence data and unreliable CDR3 (complementarity-determining region 3) loop approximation,
81 respectively. Here, we address this limitation by characterizing human naive B cells specific for
82 the SARS-CoV-2 RBD directly from the peripheral blood of seronegative donors to understand
83 their relative abundance, intrinsic affinity, and potential for activation. Furthermore, we asked
84 whether the SARS-CoV-2 specific naive repertoire could also engage related circulating variants
85 of concern and pre-pandemic CoVs. We find that SARS-CoV-2 RBD-specific naive B cells were
86 of unrestricted gene usage and several isolated B cells had affinity for circulating SARS-CoV-2
87 variants and related CoV-RBDs. We determined the structure of a representative naive antibody
88 that binds the SARS-CoV-2 RBD with a mode of recognition similar to a multi-donor class of
89 antibodies prevalent in human responses to SARS-CoV-2 infection (41). Further, we improved the

90 affinity for two representative naive antibodies to RBD and showed that the starting naive
91 specificity dictated the breadth of evolved clones to circulating variants. The analysis of the human
92 naive antigen-specific B cell repertoire for the SARS-CoV-2 RBD and its capacity to recognize
93 related variants and emerging CoVs may inform the rational design of epitope-focused
94 immunogens for next generation vaccines.

95

96 **Isolated SARS-CoV-2-specific naive B cells are genetically diverse**

97 To measure the reactivity of naive human B cells specific for the SARS-CoV-2 RBD we adapted
98 an *ex vivo* B cell profiling approach used previously to study epitope-specific naive precursors
99 targeting neutralizing sites on HIV (42-44) and influenza virus surface glycoproteins (37). We first
100 designed a SARS-CoV-2 RBD construct that positions two glycans at residues 475 and 501 to
101 selectively block binding to ACE2 and the receptor-binding motif (RBM)-directed antibody, B38
102 (**fig. S1**) (45). Using this “ Δ RBM” probe, in addition to wildtype SARS-CoV-2 spike, and RBD
103 probes, we isolated naive (CD19⁺/IgD⁺/IgG⁻) B cells specific to the RBD and, more finely, the
104 RBM from the peripheral blood of 8 SARS-CoV-2 seronegative human donors (**Fig. 1A** and **fig.**
105 **S1E**). We defined RBM-specificity as B cells that bound to fluorescently labeled spike and RBD,
106 but not the Δ RBM probe (**fig. S2A**). Although rare, all 8 donors had detectable populations of
107 RBM-specific naive B cells (**fig. S2B**). The median frequency of RBM-specific B cells among
108 total and naive B cells was 0.0021% and 0.0023%, respectively (**Fig. 1B**). Within spike-reactive,
109 naive cells, the median frequency of RBM-specific B cells was 3.2% (**Fig. 1C**); this potentially
110 suggests that a large proportion of spike epitopes targeted by naive responses reside outside of the
111 RBD. The majority of IgD⁺ RBM-specific B cells were CD27⁻ (mean frequency ~97%), in
112 agreement with the naive B cell phenotype (**fig. S2C**).

113

114 To understand in more detail the properties of this naive repertoire, we obtained 163 paired heavy-
115 and light-chain antibody sequences from 5 of the 8 donors (**Fig. 1D** and **Table S1**). Sequence
116 analysis showed that all clones were unique with diverse gene usage for both heavy and light
117 chains and minimal gene pairing preferences (**Table S1**). These data reflect the polyclonal gene
118 usage observed in RBD-specific memory B cells sequenced from COVID-19 convalescent
119 individuals (26, 29, 36) and vaccine recipients (23), suggesting that a diverse pool of antibody
120 precursors can be activated upon antigen exposure. In comparing this naive repertoire to gene
121 usage distribution from non-SARS-CoV-2-specific repertoires (46), we observed an increase in
122 mean repertoire frequency of ~20% for IGHV3-9 in 4 out of 5 sequenced donors (**fig. S3A**).
123 Notably, this enrichment of IGHV3-9 was also observed in isolated memory B cells from
124 convalescent individuals (47) and vaccine recipients (23), as well as in expanded IgG⁺ B cells
125 sequenced from a cohort of COVID-19 subjects during acute infection (36). These expanded
126 clones detected shortly after symptom onset displayed low levels of SHM (36), suggesting
127 potential IGHV3-9 usage in an early extrafollicular response in which naive B cells differentiate
128 into short-lived plasma cells (48). Additionally, IGHV3-53 and 3-30 gene segments, over-
129 represented in RBD-specific antibodies isolated from convalescent subjects (27, 35, 49), were
130 recovered from three sequenced donors (13 total clones; ~8.0% of total). The amino acid length of
131 heavy and light chain third complementarity-determining regions (CDR3) ranged from 8 to 27
132 (average length ~16) for HCDR3 and 4 to 13 (average length ~10) for LCDR3 (**Fig. 1E**). These
133 lengths are normally distributed relative to both unselected human repertoires (46, 50) and RBD-
134 specific memory B cell repertoires (23, 26, 27, 29); this is in contrast to antibody precursors
135 targeting the influenza and HIV receptor binding sites which have strict requirements for length

136 (51) or gene usage (52, 53). These data suggest that overall HCDR3 length does not restrict
137 precursor frequency and there appears no inherent bias for CDR3 length conferring RBM-
138 specificity. The majority of obtained sequences were at germline in both the variable heavy (V_H)
139 and light (V_L) chains. However, despite sorting B cells with a naive phenotype, some sequences
140 were recovered that deviated from germline. Specifically, the V_H ranged from 91.4 to 100%
141 identity to germline, with a median of 99.7%; the V_L ranged from 93.6 to 100%, with a median of
142 99.3% (**Fig. 1F, fig. 2B, C**).

143

144 **Naive antibodies engage SARS-CoV-2 RBD with high affinity**

145 To obtain affinities of the isolated naive antibodies, we cloned and recombinantly expressed 44
146 IgGs selected to reflect the polyclonal RBD-specific repertoire with representatives from diverse
147 variable region gene segments (**Table S1**). Additionally, we ensured diversity in terms of HCDR3
148 length, kappa and lambda usage, as well as representation from all 5 donors. By ELISA, we
149 identified IgGs with detectable binding to SARS-CoV-2 RBD; we summarize these results for all
150 antibodies (**Fig. 2A**) and parsed by donor (**fig. S3D**). Across 5 donors, 36 (~81%) bound to
151 monomeric SARS-CoV-2 RBD (**Fig. 2A**) with EC_{50} values ranging from 3.3 to 410 nM and a
152 mean of 62 nM (**Fig. 2A** and **fig. S3E**). These antibodies included 32 unique variable heavy and
153 light chain pairings (**Table S1**). Of the binding population, there is no apparent predisposition for
154 HCDR3 length or light chain pairing (**Fig. 2 C, D**). We further defined the epitopic region of these
155 IgGs using the Δ RBM construct and the individual glycan variants, Δ 501 and Δ 475, both of which
156 independently block ACE2 cell-surface binding but are on opposite sides of the RBM (**fig. S1E,**
157 **F**). 11 IgGs had no detectable Δ RBM binding (e.g., ab079, ab119), while 21 IgGs had reduced
158 ELISA binding relative to wild-type RBD, reflected in the reduced Δ RBM median EC_{50} values

159 **(fig. S3E)**. We also identified examples of antibodies sensitive to only $\Delta 475$ (e.g., ab185) and only
160 $\Delta 501$ (e.g., ab007) (**Fig. 2A** and **fig. S3E**).

161
162 To obtain binding kinetics independent of avidity effects from bivalent IgGs, 12 antibodies were
163 selected for expression as Fabs to determine monovalent binding affinity (K_{DS}) by biolayer
164 interferometry (BLI). Using monomeric RBD as the analyte, 10 of the 12 Fabs had detectable
165 binding with K_{DS} ranging from ~ 6.5 to ~ 75 μM ; the other two remaining Fabs (ab177, ab185),
166 gave unreliable affinity measurements (i.e., >100 μM) (**fig. S4**). Notably, all Fabs had
167 characteristically fast off rates (k_{off}). This observation is consistent for germline B cells where fast
168 off-rates are compensated by avidity due to overall BCR surface density (54); subsequent affinity
169 gains via SHM often result in slowing of the off-rate and is a canonical mechanism of improved
170 antigen binding (55-57).

171
172 **Naive antibodies engage SARS-CoV-2 variants of concern**

173 The emergence of SARS-CoV-2 variants with mutations in RBD has raised significant concern
174 that antigenic evolution will impair recognition of RBD-directed antibodies elicited by prior
175 infection and vaccination with an antigenically distinct SARS-CoV-2 variant (58-61). We
176 therefore asked whether these naive antibodies, isolated using wild-type SARS-CoV-2 RBD, could
177 recognize circulating viral variants, B.1.1.7 (mutations N501Y) (62) and B.1.351 (mutations
178 K417N/E484K/N501Y) (63); the former has now become the most prevalent circulating variant
179 in the US and many other countries (64). we find that 89% of the antibodies with wild-type RBD
180 affinity also bound to the B.1.1.7 variant with a comparable mean affinity of 68.9 nM (**Fig. 2D,**
181 **F**). For B.1.351, a concerning variant prevalent in South Africa (64), 62% of the wild-type SARS-

182 CoV-2 RBD binding IgGs also bound to the B.1.351 variant, many of which displayed reduced
183 ELISA binding relative to wild-type RBD with a mean affinity of 262 nM (**Fig. 2E, F**). A more
184 pronounced reduction in cross-reactivity to the B.1.351 variant may be predictive of reduced sera
185 binding and neutralization titers from convalescent individuals and vaccine-recipients (22, 58, 65).

186

187 **Naive antibodies engage pre-emerging CoVs**

188 We next tested the cross-reactivity of these naive antibodies to related sarbecovirus RBDs, which
189 also use ACE2 as a host receptor (66). Our panel included the previously circulating SARS-CoV
190 RBD and representative preemergent bat CoV RBDs from WIV1 (67), RaTG13 (68), and SHC014
191 (69). These RBDs share 73 to 90% paired-sequence identity with the highest degree of amino acid
192 conservation in residues outside of the RBM (70). 13 antibodies cross-reacted with at least one
193 additional RBD in our panel, with decreasing affinity for RBDs with more divergent amino acid
194 sequence identity (**Fig. 2A, G**). Notably, ab017, ab072, ab109, and ab114 had broad reactivity to
195 all tested sarbecovirus RBDs, suggesting binding to highly conserved epitopes. Of these cross-
196 reactive antibodies, ab017 and ab114, derive from the same IGHV3-33 and IGVL2-14 pairing but
197 were isolated from different donors, suggesting a shared or public clonotype.

198

199 **Naive antibodies are not polyreactive and do not engage seasonal coronaviruses.**

200 Prior studies have shown that germline antibodies are more likely to display polyreactivity relative
201 to affinity-matured antibodies with higher levels of SHM from mature B cell compartments (71-
202 74). We therefore tested the polyreactivity of all 44 naive antibodies using three common
203 autoantigens, double-stranded DNA (dsDNA), *Escherichia coli* lipopolysaccharide (LPS), and
204 human insulin in ELISA (**Fig. 2A**) We observed no polyreactivity of any naive antibody, including

205 those that are broadly reactive. Furthermore, none of the naive antibodies bound RBDs from the
206 human seasonal betacoronaviruses (hCoVs), OC43 and HKU1 (**Fig. 2A**), which share 22 and 19%
207 paired-sequence identity to SARS-CoV-2 RBD, respectively. Together, these results suggest that
208 the isolated naive B cells encode BCRs with specificity to sarbecoviruses.

209

210 ***In vitro* reconstitution of naive B cell activation**

211 Physiological interactions between a naive BCR and cognate antigen occurs at the B cell surface.
212 Naive BCRs are displayed as a bivalent membrane-bound IgM and multivalent antigen binding
213 can initiate intracellular signaling resulting in an activated B cell with the capacity to differentiate
214 to antibody secreting plasma cells or memory cells (75). To determine whether the isolated RBD-
215 specific naive BCRs have the capacity to be activated, we generated stable Ramos B cell lines
216 expressing ab090 or ab072 as cell-surface BCRs and measured their activation by monitoring
217 calcium flux *in vitro* (76). These antibodies were selected to represent divergent germline gene
218 usage and specificities: 1) ab090 (IGHV1-2/IGKV3-15) bound SARS-CoV-2 and variant B.1.1.7
219 RBDs, but not variant B.1.351 and WIV1 RBDs (**Fig. 3A**); and 2) ab072 (IGHV3-23/IGLV2-14)
220 had broad reactivity to all RBDs (**Fig. 3B**). To assess BCR activation, we generated ferritin-based
221 nanoparticles (NPs) for multivalent RBD display using SpyTag-SpyCatcher (70, 77, 78); these
222 RBD NPs included SARS-CoV-2, B.1.1.7, B.1.351 and WIV RBDs. We found that ab090
223 expressing Ramos B cells were only activated by SARS-CoV-2 RBD and variant B.1.1.7 RBD
224 NPs (**Fig. 3C**), while ab072 Ramos B cells were activated by all RBD-NPs (**Fig. 3D**). Notably,
225 these data parallel the observed recombinant binding specificity of each antibody. Importantly,
226 neither ab090 nor ab072 Ramos B cell lines were activated by influenza hemagglutinin NPs,
227 suggesting that this activation is sarbecovirus RBD-specific (**Fig. 3C, D**).

228

229 **ab090 engages the SARS-CoV-2 RBM**

230 To further characterize the epitope specificity of a representative naive antibody, we determined
231 the structure of ab090 in complex with SARS-CoV-2 spike (S) by electron cryomicroscopy (cryo-
232 EM). A $\sim 6.7\text{-\AA}$ structure showed one Fab bound to an RBD in the “up” conformation (**Fig. 4A, B**
233 **and fig. S5**). Based on this modest resolution structure, we make the following general descriptions
234 of the antibody-antigen interface. The interaction between ab090 and the RBD is mediated
235 primarily by the antibody heavy chain, with the germline encoded HCDR1, HCDR2, and the
236 framework 3 DE-loop centered over the RBM epitope (**Fig. 4B**). The ab090 light chain is oriented
237 distal to the RBD and does not appear to substantially contribute to the paratope (**Fig. 4B**). IGHV1-
238 2 antibodies represent a prevalent antibody class in human responses to SARS-CoV-2 infection,
239 many of which display high neutralization potency (41, 79). ab090 shares a V_H -centric mode of
240 contact and angle of approach similar to members of this class of infection-elicited antibodies (**Fig.**
241 **4D**), despite varying HCDR3 lengths and diverse light chain pairings (**Fig. 4D**) (41, 79).
242 Additionally, members of the IGHV1-2 antibody class contain relatively few SHMs (**fig. S5B**).
243 We note that many of the infection-elicited IGHV1-2 RBD-specific memory B cells derive from
244 the IGHV1-2*02 allele, while ab090 is encoded by the IGHV1-2*06 allelic variant (**fig. S5B**). The
245 IGHV1-2*06 allele is represented by a single nucleotide polymorphism encoding an arginine
246 rather than a tryptophan at position 50 (80) (**fig. S5B**). Notably, a potent neutralizing antibody,
247 H4, derives from the same *06 allele (34). In conjunction with the structure, we biochemically
248 defined the sensitivity of ab090 to variant B.1.351 by testing the binding to individual mutations.
249 Binding affinity was detected to SARS-CoV-2 RBDs with either N501Y or K417N mutations, but
250 not to E484K alone (**Fig. 4C**). Based on the structure, the E484K mutation, is grossly positioned

251 proximal to the CDRH2 loop (**Fig. 4C**), which has a germline-encoded motif critical for IGHV1-
252 2 antibody binding to RBD (**fig. S5B**) (41). Indeed, infection-elicited IGHV1-2 antibodies are
253 susceptible to escape by E484K alone, which disrupts a CDRH2 hydrogen binding network (81).
254 Together, the cryo-EM structure and binding data suggest that ab090 represents a precursor of a
255 class of RBM directed SARS-CoV-2 neutralizing antibodies. More generally, structural
256 characterization of germline antibody complexes has been limited to hapten antigens (82), simple
257 peptides (83) and to protein antigens bearing engineered affinity to inferred germline
258 sequences(84). We present, to our knowledge, the first structure of a naturally occurring naive
259 human antibody bound to non-engineered viral protein.

260

261 ***in vitro* affinity-matured naive antibodies retain intrinsic specificity**

262 After initial antigen recognition and subsequent activation, naive B cells can undergo successive
263 rounds of somatic hypermutation within the germinal center (GC) that ultimately result in higher
264 affinity antibodies for the cognate antigen. To determine how somatic hypermutation might
265 influence overall affinity and specificity, we used yeast surface display to *in vitro* mature ab072
266 and ab090. We randomly mutagenized the single chain variable fragment (scFv) variable heavy
267 and light chain regions to generate ab072 and ab090 variant display libraries (85). After two rounds
268 of selections using SARS-CoV-2 RBD, we enriched the ab072 and ab090 libraries for improved
269 binding over their respective parental clones (**Fig. 5A, D and fig. S6A**). We also observed
270 increased binding to B.1.351 for the ab072 library but not for ab090; notably this corresponded
271 with the respective specificity of the parent clones (**Fig. 5A, D**).

272

273 We next isolated and sequenced individual clones from the enriched libraries. For ab090, we
274 observed a dominant mutation, R72H, in the FRWH3 region present in 60% of sequenced clones
275 (**fig. S6B**). Notably, multiple mutations at position 72 conferred a ~3- to 5-fold improvement in
276 monovalent affinity relative to parental ab090 for wild-type and B.1.1.7 RBDs, with no detectable
277 B.1.351 binding for affinity matured progeny (**Fig. 5B, C**). We observed no mutations within the
278 light chain which appears to be consistent with the V_H -centric binding mode in the cryo-EM
279 structure (**Fig. 4**). For the broadly reactive ab072, isolated clones had mutations in both the V_H and
280 V_L ; ~35% of the sequenced clones had mutation S31P in the HCDR1 (**fig. S6B, C**). There was 3-
281 to ~5-fold improvement in monovalent affinity of ab072 progeny relative to parent for SARS-
282 CoV-2, B.1.1.7 and B.1.351 RBDs (**Fig. 5E, F**). Collectively, these data identify potential
283 mutations that can improve affinity while retaining initial parental antigen specificity.

284

285 **SARS-CoV-2 pseudovirus neutralization by naive and affinity-matured Abs**

286 We next used a SARS-CoV-2 pseudovirus assay (10) to ask whether any of the isolated naive
287 antibodies and affinity matured clones were capable of blocking transduction of target cells. We
288 found that of the 36 RBD-binding antibodies tested in this assay, 5 had detectable levels of
289 neutralization (~14%) (**Fig. 6A**). These antibodies, obtained from multiple donors, have no
290 commonality with respect to their gene usages and HCDR3 lengths (**Fig. 6B**). While these naive
291 antibodies were not as potent as B38, isolated from a memory B cell (34), the observation,
292 nevertheless, that the naive repertoire has antibodies that neutralize is noteworthy.

293

294 To determine whether improved affinity correlated with enhanced neutralization potency, we
295 evaluated the affinity matured progeny of ab090 in a SARS-CoV-2 pseudovirus neutralization

296 assay (**Fig. 6B**). We find that all three ab090 progeny that had higher affinity for SARS-CoV-2
297 RBD also had increased neutralization potency. ab090_A08 bearing the R72H mutation had the
298 highest affinity gain and was the most potent neutralizer with a K_D of 1.7 μ M and an IC50 of 0.37
299 μ g/ml, respectively. Notably, ab090 progeny had IC50 values similar to other IGHV1-2 memory
300 B cells isolated from convalescent donors (41); this increase in potency is conferred through
301 minimal somatic hypermutation.

302

303 **DISCUSSION**

304 The development of a protective humoral immune response upon infection or vaccination relies
305 on the recruitment, activation, and maturation of antigen-specific naive B cells. However, the
306 specificity of the naive B cell repertoire remains largely undefined. Here, we showed that
307 coronavirus-specific naive B cells are present across distinct seronegative donors, are of
308 unrestricted gene usage and when recombinantly expressed as IgGs, have affinity for SARS-CoV-
309 2 RBD, circulating variants of concern, and at least four related coronaviruses. These data suggest
310 that RBD-specific precursors are likely present across a large fraction of individual human naive
311 repertoires, consistent with longitudinal studies of SARS-CoV-2 infected individuals in which
312 most convalescent individuals seroconverted with detectable RBD serum antibodies and
313 neutralization titers (17, 86, 87). The naive B cells characterized here engage epitopes across the
314 RBM with a range of angles of approach as defined by our glycan variant probes and cross-
315 reactivity profiles; this is also consistent with infection and vaccine elicited, RBD-specific
316 repertoire characterized by epitope-mapping, deep mutational scanning and structural analyses
317 (30, 32, 88). Having naive BCRs recognizing distinct or partially overlapping epitopes across the

318 RBM may be advantageous for eliciting a polyclonal response more able to recognize variants of
319 concern.

320

321 The presence of broadly reactive naive B cells inherently capable of recognizing sarbecovirus
322 RBDs and circulating variants suggests that these precursors could be vaccine-amplified. Recent
323 work showed that uninfected individuals have pre-existing SARS-CoV-2 S-reactive serum
324 antibodies (89-91) and memory B cells (28, 92) which cross-react with hCoVs and can be boosted
325 upon SARS-CoV-2 infection. These cross-reactive antibodies appear to be specific to the S2
326 domain and are predominantly IgG or IgA. Notably, this observation contrasts the cross-reactive
327 B cells described here that engage the RBD, have no reactivity to hCoV and are IgG⁻ naive B cells
328 suggesting that they are distinct from previously described S-reactive pre-existing antibodies.

329

330 Data suggests that the competitive success of a naive B cell within a GC is influenced by precursor
331 frequencies and antigen affinities (40). However, the biologically relevant affinities necessary for
332 activation remain unclear—indeed several studies suggest that B cell activation and affinity
333 maturation is not restricted by immeasurably low affinity BCR interactions (93-95). Recently, two
334 studies involving naive precursors of receptor-binding site (RBS) directed HIV-1 broadly
335 neutralizing antibodies (bnAbs) contributed to our understanding of these parameters (38, 39).
336 Using an *in vivo* murine adoptive transfer model, these RBS-directed precursors were recruited
337 into a GC reaction at a precursor frequency of ~1:10,000 and a monovalent antigen affinity of
338 14 μ M (39). For comparison, here we defined the SARS CoV-2 RBM-specific naive precursor
339 frequency as 1:41,000 by flow cytometric gating (**fig. S2**) with monovalent affinities ranging from
340 6.5 to >100 μ M. These data suggest that these isolated naive B cells, especially those with

341 demonstrable monomeric affinity, could be readily elicited upon antigen exposure. However,
342 longitudinal studies tracking antigen specific naive B cells pre- and post-exposure are required to
343 determine the fate (i.e., plasma cell, memory, or germinal center B cell compartments) of potential
344 precursors and define relevant naive affinities for elicitation by SARS-CoV-2.

345
346 Through biochemical and structural analyses, we characterized a naive antibody, ab090, which
347 resembles a commonly elicited class of potent neutralizing antibodies utilizing the IGHV1-2 gene
348 (41). This class of antibodies share restricted binding specificity for wild-type SARS-CoV RBD
349 (the vaccine strain) and the prevalent B.1.1.7 variant. This recombinant binding pattern also
350 paralleled the reconstituted *in vitro* B cell activation dynamics of ab090 in the highly avid assay
351 with the capacity to detect immeasurably low affinity interactions (54). *In vitro* affinity maturation
352 of ab090 against corresponded to a single H-FR3 mutation, which improved monovalent affinity
353 ~5-fold to wild-type SARS-CoV-2 and B.1.1.7 RBDs relative to parent and pseudovirus
354 neutralization to IC₅₀ values less than 1µg/ml. This observation is consistent with the low levels
355 of SHM within IGHV1-2 neutralizing antibodies (41) and with reports of other potent RBD-
356 directed neutralizing antibodies with a limited level of somatic hypermutation (24, 26, 29, 36, 96).
357 Further, a recent study monitoring RBD-specific memory B cell evolution up to 12 months after
358 SARS-CoV-2 infection revealed examples of affinity matured clones with increased neutralizing
359 breadth over time against circulating RBD variants (97). While *in vitro* affinity gains and
360 neutralization potency are generally correlated (98), we note that affinity does not necessarily
361 correlate to neutralization potency for all SARS-CoV-2 RBD targeting antibodies, where fine
362 epitope specificity appears to be most relevant (28, 99).

363

364 Probing and characterizing the human naive B cell antigen-specific repertoire can identify
365 precursors for vaccine or infection-specific naive B cells and expand our understanding of basic B
366 cell biology. Germline-endowed specificity for neutralizing antibody targets on the RBD may also
367 contribute to the strong clinical efficacy observed for the current SARS-CoV-2 vaccines (*100*,
368 *101*). Furthermore, understanding the naive B cell repertoire to potential pandemic coronaviruses
369 may reveal commonalties in antigen-specific precursors, enabling the development of pan-
370 coronavirus vaccines aimed at engaging broadly protective germline responses.

371

372

373

374 **METHODS**

375 **Donor Samples**

376 PBMCs were isolated from blood donors obtained from the MGH blood donor center (8 donors
377 total). Prior to donating blood, subjects were required to sign a donor attestation/consent statement,
378 as per hospital requirements, stating “I give permission for my blood to be used for transfusion to
379 patients or for research”. The gender and age are not recorded, however eligible donors are of at
380 least 16 years old and weigh a minimum of 110lbs. All experiments were conducted with MGH
381 Institutional Biosafety Committee approval (MGH protocol 2014B000035). Isolated PBMCs were
382 used for B cell enrichment and single cell sorting (described below); plasma was aliquoted and
383 stored at -80 °C until further use. Additionally, the control convalescent sera used for ELISA was
384 obtained under the approved Partners Institutional Review Board (protocol 2020P000895) for use
385 of patient samples for the development and validation of SARS-CoV-2 diagnostic tests (10).

386

387 **Expression and purification of recombinant CoV Antigens**

388 Plasmids encoding the receptor binding domains (RBDs) were designed based on GenBank
389 sequences MN975262.1 (SARS-CoV-2), ABD72970.1 (SARS-CoV), AGZ48828.1 (WIV-1),
390 MN996532.2 (RaTG13), QJE50589.1 (SHC014), AAT98580.1 (HKU1), and AAT84362 (OC43).
391 Constructs were codon optimized and synthesized by IDT. QuikChange Mutagenesis (Agilent)
392 was used to insert glycosylation sites at SARS-CoV-2 RBD residues 501 and/or 475 as well as for
393 RBD variant mutations, B.1.351 (K417N/E484K/N501Y) and B.1.1.7 (N501Y). SARS-CoV-2
394 spike contained a C-terminal foldon trimerization domain and HRV 3C-cleavable 6xHis and
395 2xStrep II tags (102). All proteins were transiently expressed in Expi293F cells (ThermoFisher).
396 5 to 7 days post-transfection, supernatants were harvested by centrifugation and further purified

397 using immobilized metal affinity chromatography (IMAC) with cobalt-TALON resin (Takara)
398 followed by Superdex 200 Increase 10/300 GL size exclusion column (GE Healthcare).

399

400 **Expression and purification IgGs and Fabs**

401 IgG and Fab genes for the heavy- and light-chain variable domains were synthesized and codon
402 optimized by IDT and subcloned into pVRC protein expression vectors and sequence confirmed
403 (Genewiz). Fabs and IgGs were similarly expressed and purified as described above for RBDs.
404 IgGs were buffer exchanged into PBS while Fabs were concentrated and further purified by
405 Superdex 200 Increase 10/300 GL size exclusion column.

406

407 **ELISA**

408 Both sera and monoclonal antibody reactivity to CoV antigens were assayed by ELISA. Briefly,
409 96-well plates (Corning) were coated with 5 µg/ml of monomeric RBDs in PBS at 100µl/well and
410 incubated overnight at 4°C. Plates were blocked with 1% BSA in PBS containing 1% Tween-20
411 (PBS-T) for 1hr at room temperature (RT). Blocking solution was discarded and 4-fold serial
412 dilutions of human plasma (1:20 starting dilution) or isolated monoclonal antibodies (150 µg/ml
413 starting concentration) in PBS were added to wells and incubated for 1hr at RT. Plates were then
414 washed three times with PBS-T. Secondary, anti-human IgG-HRP (Abcam), was added to each
415 well at 1:20,000 dilution in PBS-T and incubated for 1hr at RT. Plates were then washed three
416 times with PBS-T and developed with 1-Step ABTS substrate (ThermoFisher) per manufacturer
417 recommendations. Absorbance was measured using a plate reader at 405nm. EC₅₀ values were
418 determined for monoclonal antibodies by non-linear regression (sigmoidal) using GraphPad Prism

419 8.4.3 software. ELISAs against OC43 and HKU1 RBDs were done at a single IgG concentration
420 (150 µg/ml) in replicate. Positive binding was defined by an $OD_{405} \geq 0.30$.

421
422 For polyreactivity ELISAs against human insulin (MilliporeSigma) and dsDNA (Calf Thymus
423 DNA; Invitrogen), plates were coated with 2µg/ml and 50µg/ml, respectively, in PBS at
424 100µl/well and incubated overnight at 4°C. Plates were then blocked and incubated with IgGs as
425 described above for CoV antigens. LPS ELISAs were measured according to a previously
426 described method (103, 104). Briefly, plates were coated with 30µg/ml LPS (*Escherichia coli*
427 O55:B5; MilliporeSigma) in carbonate buffer (100mM Na₂CO₂, 20mM EDTA, pH 9.6) at
428 100µl/well for 3hrs at 37°C, washed three times with water, and air-dried overnight at RT. Coated
429 plates were blocked with 200µl/well of HS buffer (50mM HEPES, 0.15mM NaCl, pH 7.4) plus
430 10mg/ml. Plates were incubated with IgGs diluted in HS buffer containing 1mg/ml BSA for 3hrs
431 at 37°C, washed three times with HS buffer, and developed as detailed above for CoV antigens.
432 All polyreactivity ELISAs were performed at a single IgG concentration (15µg/ml) in replicate
433 with positive binding was defined by an $OD_{405} \geq 0.30$.

434

435 **ACE-2 cell binding assay**

436 ACE-2 expressing 293T cells were incubated with 200 nM of RBD antigen in PBS for 1hr on ice.
437 Cells were resuspended in 50µL of secondary stain containing streptavidin-PE (Invitrogen) at a
438 1:200 dilution and incubated for 30 min on ice. Cell binding was analyzed by flow cytometry using
439 a Stratadigm S1300Exi Flow Cytometer equipped with a 96 well plate high throughput sampler.
440 Resulting data were analyzed using FlowJo (10.7.1).

441

442 **Probe Generation**

443 SARS-CoV-2 RBD and Δ RBM constructs were expressed as dimeric murine-Fc (mFc; IgG1)
444 fusion proteins containing a HRV 3C-cleavable C-terminal 8xHis and SBP tags and purified as
445 described above. SBP-tagged RBD- and Δ RBM-mFc dimers were individually mixed with
446 fluorescently labeled streptavidin, SA-BV650 and SA-BV786 (BioLegend), to form RBD-mFc-
447 BV650 and Δ RBM-mFc-BV786 tetramers. SARS-CoV-2 spike with a C-terminal Strep II tag was
448 labeled separately with StrepTactin PE and APC (IBA) to form spike-PE and -APC tetramers,
449 respectively. Both labeling steps were performed for 30 min at 4 °C prior to sorting.

450

451 **Single B Cell Sorting**

452 Naive B cells were purified from PBMCs using the MACS Human B Cell isolation kit (Miltenyi
453 Biotec) and incubated with 25nM of each SARS-CoV-2 probe (RBD-mFc-BV650, Δ RBM-mFc-
454 BV786, spike-PE, and spike-APC) for 30 min at 4°C. Cells were stained with anti-human CD19
455 (Alexa-700), CD3 (PerCP-Cy5), IgD (PE-Cy7), IgG (BV711), CD27 (BV510), LiveDead Violet
456 (Invitrogen), and Calcein (Invitrogen) for an additional 30 min. RBM-specific naive B cells,
457 defined as CD19⁺/CD3⁻/IgG⁻/IgD⁺/spike PE⁺/spike APC⁺/RBD⁺/ Δ RBM⁻, were single-cell sorted
458 using BD FACS Aria II (BD Biosciences) into 96-well plates containing lysis buffer supplemented
459 with 1% BME. Within the CD19⁺/IgG⁻/IgD⁺ gated cells, we also confirmed that 97% of the events
460 were CD27 negative. Plates were stored at -80 °C for subsequent analysis. Flow cytometry data
461 was analyzed using FlowJo software version 10.7.1.

462

463 **BCR Sequencing**

464 BCR Sequencing was carried out as described previously (37). Briefly, whole transcriptome
465 amplification (WTA) was performed on the sorted cell-lysates according to the Smart-Seq2
466 protocol (105). We then amplified heavy and light chain sequences from the WTA products
467 utilizing pools of partially degenerate pools of V region specific primers (Qiagen HotStar Taq
468 Plus). Heavy and light chain amplifications were carried out separately, with each pool containing
469 pooled primers against human IGHV and heavy chain constant region genes, or human IGLV,
470 IGKV, and light chain constant region genes. Cellular barcodes and index adapters (based on
471 Nextera XT Index Adapters, Illumina Inc.) were added using a step-out PCR method. Amplicons
472 were then pooled and sequenced using a 250x250 paired end 8x8 index reads on an Illumina Miseq
473 System. The data were then demultiplexed, heavy and light chain reads were paired, and
474 overlapping sequence reads were obtained (Panda-Seq) (106) and aligned against the human
475 IMGT database (107).

476

477 **Interferometry binding experiments**

478 Interferometry experiments were performed using a BLItz instrument (ForteBio). Fabs (0.1 mg/ml)
479 were immobilized on Ni-NTA biosensors. The SARS-CoV-2 RBD analyte was titrated (10 μ M,
480 5 μ M, 2.5 μ M, and 1 μ M) to acquire binding affinities; the K_D was obtained through global fit of the
481 titration curves by applying a 1:1 binding isotherm using vendor-supplied software.

482

483 **Pseudotyped neutralization assay**

484 SARS-CoV-2 neutralization was assessed using lentiviral particles pseudotyped as previously
485 described (10, 108). Briefly, lentiviral particles were produced via transient transfection of 293T
486 cells. The titers of viral supernatants were determined via flow cytometry on 293T-ACE2 cells

487 (108) and via the HIV-1 p24^{CA} antigen capture assay (Leidos Biomedical Research, Inc.). Assays
488 were performed in 384-well plates (Grenier) using a Fluent Automated Workstation (Tecan). IgGs
489 starting at 150 µg/ml, were serially diluted (3-fold) in 20µL followed by addition of 20 µL of
490 pseudovirus containing 250 infectious units and incubated at room temperature for 1 hr. Finally,
491 10,000 293T-ACE2 cells (108) in 20 µL cell media containing 15 µg/ml polybrene were added to
492 each well and incubated at 37 °C for 60-72 hrs. Following transduction, cells were lysed using a
493 previously described assay buffer (109) and shaken for 5 min prior to quantitation of luciferase
494 expression using a Spectramax L luminometer (Molecular Devices). Percent neutralization was
495 determined by subtracting background luminescence measured from cells control wells (cells only)
496 from sample wells and dividing by virus control wells (virus and cells only). Data were analyzed
497 using Graphpad Prism.

498

499 **cryo-EM sample preparation, data collection and processing**

500 SARS-CoV-2 spike HexaPro was incubated with ab090 Fab at 0.6 mg/mL at a molar ratio of 1.5:1
501 Fab:Spike for 20 minutes at 4°C and two 3 µl aliquots were applied to UltrAuFoil gold R0.6/1
502 grids and subsequently blotted for 3 seconds at blot force 3 twice, then plunge-frozen in liquid
503 ethane using an FEI Vitrobot Mark IV. Grids were imaged on a Titan Krios microscope operated
504 at 300 kV and equipped with a Gatan K3 Summit direct detector. 10,690 movies were collected in
505 counting mode at 16e⁻/pix/s at a magnification of 81,000, corresponding to a calibrated pixel size
506 of 1.058 Å. Defocus values were at around -2.00 µm. Micrographs were aligned and dose weighted
507 using Relion's (110) implementation of MotionCorr2 (111). Contrast transfer function estimation
508 was done in GCTF (112). Particles were picked with crYOLO (113) with a model trained with 12
509 manually picked micrographs with particle diameter value of 330Å. Initial processing was

510 performed in Relion. The picked particles were binned to $\sim 12\text{\AA}/\text{pixel}$ and subjected to a 2D
511 classification. Selected particles were then extracted to $\sim 6\text{\AA}/\text{pixel}$ then subjected to a second round
512 of 2D classification. An initial model was generated on the selected particles at $\sim 6\text{\AA}/\text{pixel}$ and used
513 as a reference for two rounds of 3D classification; first to select particles containing SARS-CoV-
514 2 spike then to select particles containing both spike and ab090. Selected particles were unbinned
515 then aligned using 3D auto-refine and subjected to a third round of 3D classification to select for
516 a single class with SARS-CoV-2 spike bound with one ab090 Fab. Selected particles were aligned
517 using 3D auto-refine before undergoing CTF refinement and Bayesian polishing. Polished
518 particles were then simultaneously focus-aligned relative to the RBD and ab090 region (Figure S5
519 A) to aid in model building of this region of interest and imported to cryoSPARC (114). Imported
520 particles were aligned using non-uniform refinement and local resolution estimation (Figure S5B).
521 Non-uniform refined maps were then sharpened with DeepEMhancer then used to dock a
522 previously built SARS-CoV-2-spike model (PDB ID 7LQW).

523

524 **cryo-EM model building**

525 Backbone models were built by docking the variable regions of structurally similar Fabs (PDB ID
526 2D2P and 6FG1 for heavy and light chains, respectively) and a previously built RBD (6M0J) into
527 the focus refined maps using UCSF Chimera (115) variable regions were then mutated and
528 manually built using COOT (116). For the remainder of the spike, a previously published model
529 (PDB ID 6VXX) was docked into the full, sharpened map in UCSF Chimera.

530

531 **RBD nanoparticle production and conjugation**

532 Monomeric SARS-CoV-2 wild-type, B.1.1.7, B.1.351, and WIV1 RBDs were recombinantly
533 produced and purified as described above with an 8xHis and SpyTag (cite) at the C-terminus.
534 *Helicobacter pylori* ferritin nanoparticles (NP) were expressed separately with N-terminal 8xHis
535 and SpyCatcher tags. SpyTag-SpyCatcher conjugations were performed overnight at 4°C with a
536 4-fold molar excess of SpyTag-RBD relative to SpyCatcher-NP. The conjugated RBD-NPs were
537 repurified by size-exclusion chromatography to remove excess RBD-SpyTag.

538

539 ***In vitro* BCR triggering**

540 The capacity of RBD-NPs to trigger naive BCR signaling was determined through activation of
541 Ramos cells engineered to display mono-specific IgM BCRs of interest, as previously described
542 (76). Briefly, BCRs for ab090 and ab072 were stably expressed in an IgM negative Ramos B cell
543 clone by lentiviral transduction. Five to seven days post transduction, confluent BCR-expressing
544 B cells were FACS sorted on IgM (APC anti-human IgM; BioLegend) and kappa light chain (PE
545 anti-human kappa light chain; BD Biosciences) double positivity using a SH800S Cell Sorter
546 (Sony Biotechnology). Sorted cells were expanded in RPMI (GIBCO) and evaluated for B cell
547 activation by labeling 10 million cells with 0.5µg/ml Fura red dye (Invitrogen) in 2ml of RPMI at
548 37°C for 30 min. Cells were then washed and resuspended to 4 million cells/ml in RPMI. BCR
549 triggering was measured in response to the RBD-NPs described above by flow cytometry (LSR II,
550 BD Biosciences) as the ratio of Ca²⁺ bound/unbound states of Fura red. Ratiometric measures for
551 individual B cell lines were normalized to the maximum Ca²⁺ flux as measured by exposure to
552 10µg/ml ionomycin.

553

554 ***in vitro* affinity maturation of ab090 and ab072**

555 To build yeast display libraries for ab090 and ab072, variable heavy and light chains were
556 reformatted into an scFv and synthesized as gBlocks (IDT). gBlocks were amplified by polymerase
557 chain reaction (PCR) using Q5 polymerase (New England BioLabs) following the manufacturer's
558 protocol. The amplified DNA purified and subsequently mutagenized by error-prone PCR (ePCR)
559 via the GeneMorph II Random Mutagenesis Kit (Agilent Technologies) with a target nucleotide
560 mutation frequency of 0-4.5 mutations/kb. Mutagenized scFv DNA products were combined with
561 the linearized yeast display vector pCHA (117) and electroporated into EBY100 grown to mid-log
562 phase in YPD media, where the full plasmid was reassembled by homologous recombination (85).
563 The final library size was estimated to be 4×10^7 .

564

565 The scFv libraries and selection outputs were passaged in selective SDCAA media (20 g/L
566 dextrose, 6.7 g/L Yeast Nitrogen base, 5 g/L Bacto casamino acids, 5.4 g/L Na_2HPO_4 and 8.56 g/L
567 $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$) at shaking at 30°C and induced in SGCAA media (same as SDCAA with 20 g/L
568 galactose instead of dextrose) at 20°C. The scFv libraries were induced covering at least 10-fold
569 of their respective diversities and subject to three rounds of selection for binding to SBP-tagged
570 SARS-CoV-2 RBD-Fc. Induced yeast libraries were stained for antigen binding (RBD-Fc APC
571 tetramers) and scFv expression (chicken anti-c-myc IgY; Invitrogen). Following two washes in
572 PBSF (1x PBS, 0.1% w/v BSA), yeast was stained with donkey anti-chicken IgY AF488 (Jackson
573 ImmunoResearch). Two gates were drawn for cells with improved RBD binding over parental
574 clones, a more stringent “edge” gate represented ~1% and a “diversity” gate represented ~3-5% of
575 the improved output. Yeast from the final round of selection were resuspended in SDCAA media
576 and plated on SDCAA agar plates for single colony isolation and Sanger sequencing from which
577 IgGs and Fabs were cloned and recombinantly expressed as described above.

578

579 **Acknowledgements**

580 We thank members of the Schmidt and Lingwood Labs for helpful discussions, especially Tim
581 Caradonna, Catherine Jacob-Dolan, and Daniel Maurer. We thank Samuel Kazer, James Gatter
582 and Alex Shalek for BCR sequencing advice, Jason McLellan for the SARS-CoV-2 spike plasmid,
583 and Nir Hacohen and Michael Farzan for ACE2 expressing 293T cells. Some of this work was
584 performed at the National Center for CryoEM Access and Training (NCCAT) and the Simons
585 Electron Microscopy Center located at the New York Structural Biology Center, supported by the
586 NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy program (U24
587 GM129539), and by grants from the Simons Foundation (SF349247) and NY State Assembly.

588

589 **Funding:** We acknowledge support from NIH (R01AI146779, R01AI124378, R01AI137057,
590 R01AI153098, R01AI155447, DP2DA042422, DP2DA040254, T32 AI007245), a Massachusetts
591 Consortium on Pathogenesis Readiness (MassCPR) grant to A.G.S. and a MGH Transformative
592 Scholars Program and Charles H. Hood Foundation to A.B.B.

593

594 **Author contributions:** J.F., J.B., A.B.B., G.B., D.L., A.G.S. designed research; J.F., J.B., C.G.A.,
595 K.S.D., E.C.L, B.M.H., L.R., M.S., T.B.M. ,V.O., N.H., performed research; J.F., J.B., C.G.A.,
596 G.B., D.L., A.G.S analyzed data; J.F. and A.G.S. wrote the paper. J.F., J.B., C.G.A., B.M.H,
597 A.B.B., G.B., D.L., A.G.S edited and commented on the paper.

598

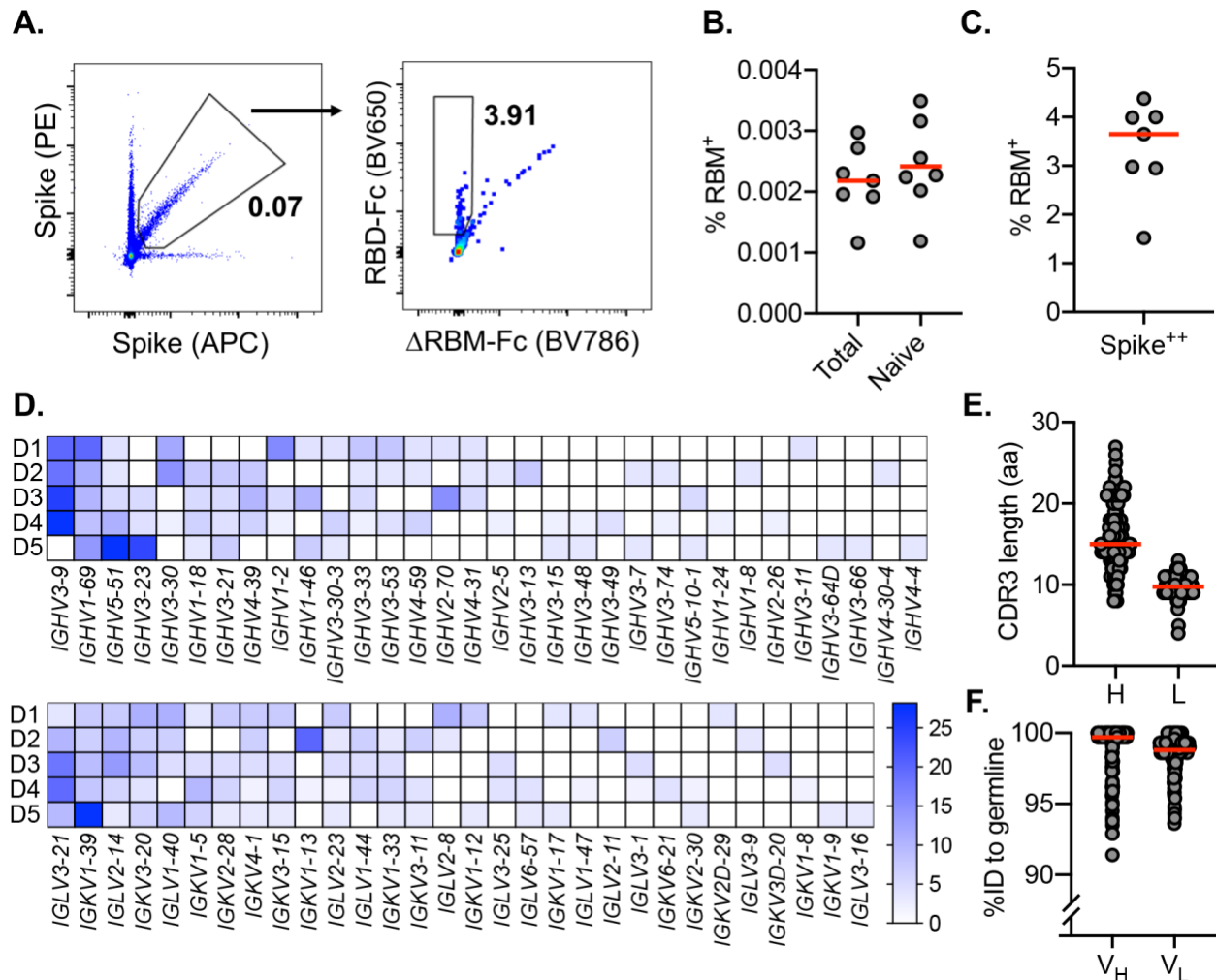
599 **Competing interests:** Authors declare no competing interests.

600

601 **Data and materials availability:** All data are provided in the Supplementary Materials. Requests
602 for material should be addressed to Daniel Lingwood (dlingwood@mgh.harvard.edu) or Aaron G.
603 Schmidt (aschmidt@crystal.harvard.edu). This work is licensed under a Creative Commons
604 Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and
605 reproduction in any medium, provided the original work is properly cited. To view a copy of this
606 license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to
607 figures/photos/artwork or other content included in the article that is credited to a third party;
608 obtain authorization from the rights holder before using such material. The EM maps have been
609 deposited in the Electron Microscopy Data Bank (EMDB) under accession code: EMD-24279.
610

611 MAIN TEXT FIGURES

612



613

614 **Fig. 1. SARS-CoV-2-specific naive B cells isolation and characterization.** (A) RBM-specific

615 naive B cells from seronegative human donors were isolated by fluorescence-activated cell sorting

616 gated on CD19⁺IgD⁺IgG⁻; representative plot from donors 1 and 2 is shown. Δ RBM is a sorting

617 probe with N-linked glycans at residues 501 and 475. RBM-specific B cell frequency among (B)

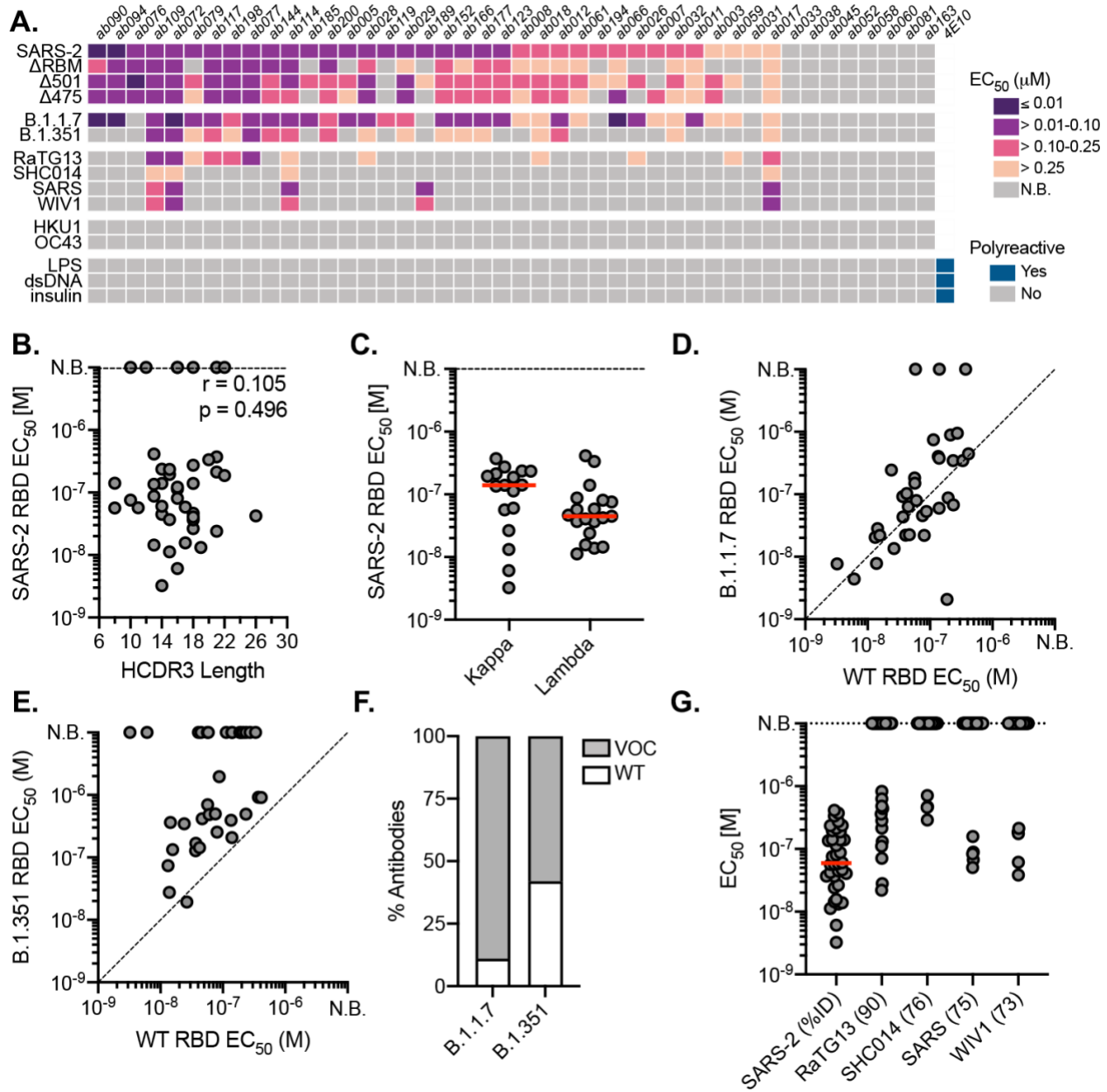
618 total, naive, and (C) spike-positive cells from each donor ($n = 8$). (D) Heat map showing variable-

619 gene usage for all paired B cell sequences. Scale indicates percent of total sequences for each

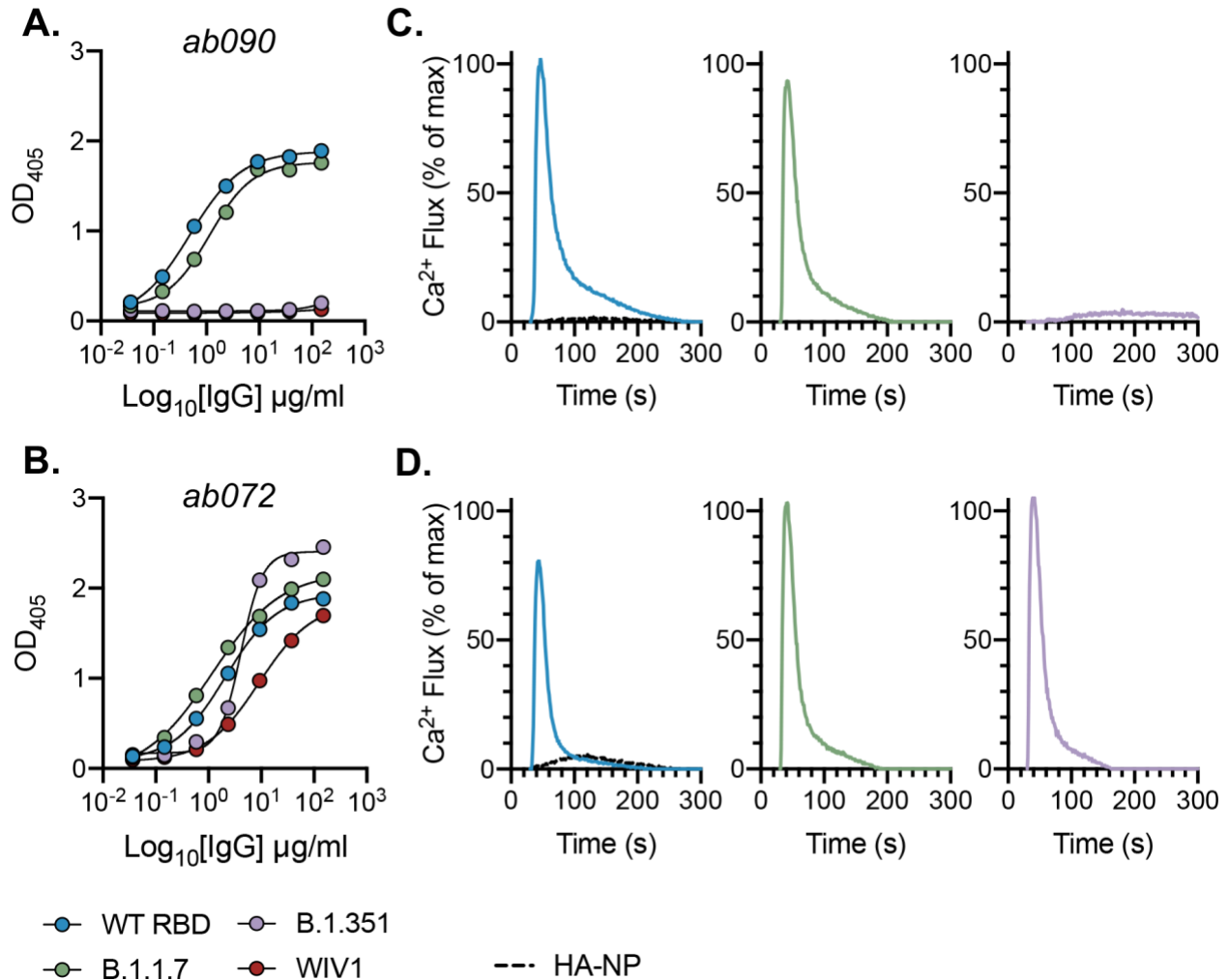
620 donor separately. (E) Heavy (H) and light (L) CDR3 amino acid length distribution determined

621 using IMGT numbering. Red bars indicate median amino acid length. (F) Divergence from

622 inferred germline gene sequences. Red bars indicate the median percent values.

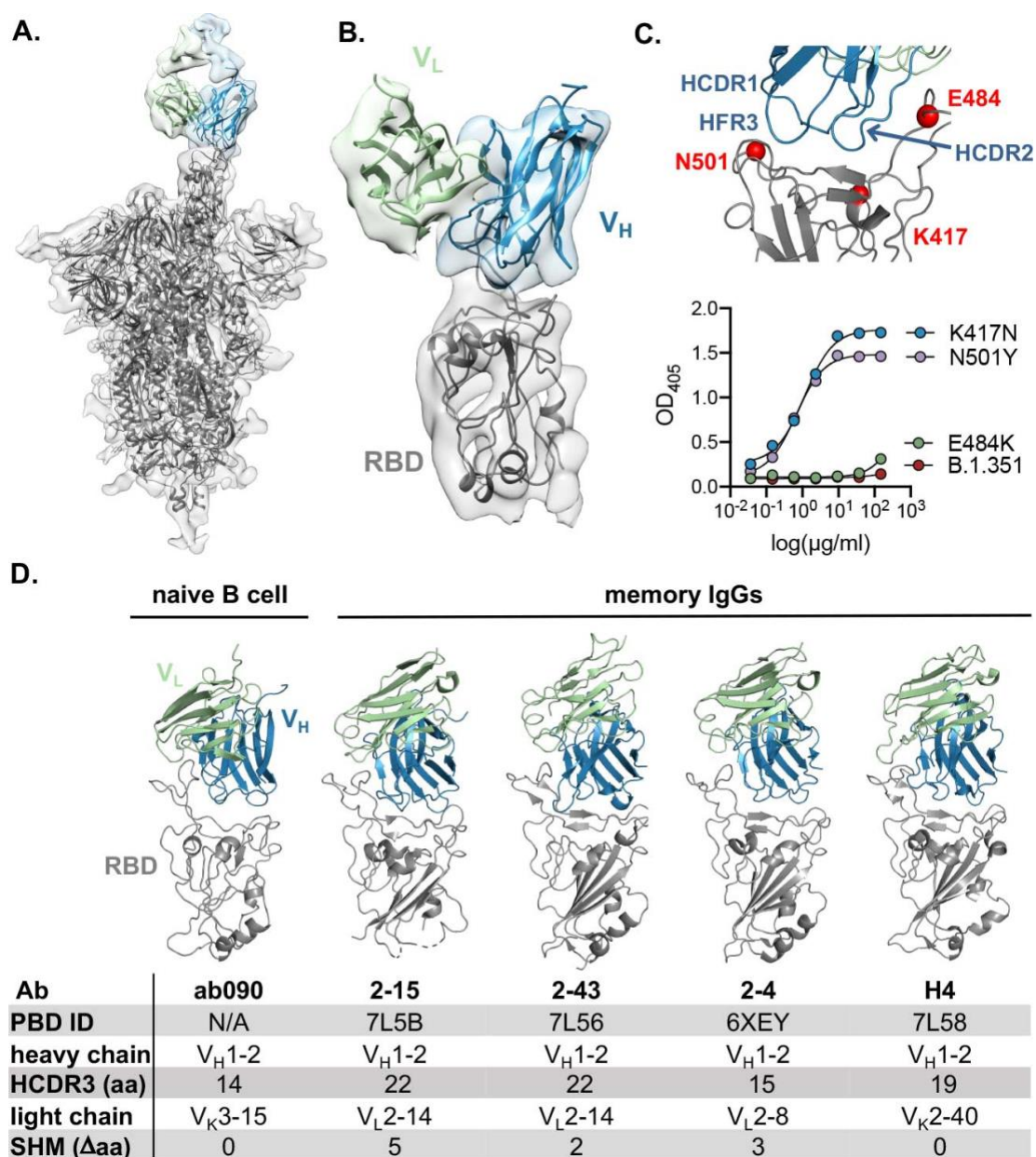


623
 624 **Fig. 2. Binding properties and specificity of isolated naive antibodies.** (A) ELISA binding heat
 625 map of 44 naive IgGs. Binding to wildtype SARS-CoV-2 RBD (SARS-2), Δ RBM, individual RBD
 626 glycan variants, circulating variants, related CoVs, hCoVs and polyreactivity antigens. (B) Pearson
 627 correlation analysis of SARS-CoV-2 RBD affinities and HCDR3 length. (C) ELISA EC_{50} s for
 628 IgGs with detectable SARS-CoV-2 RBD binding ($n = 36$) based on kappa or lambda gene usage.
 629 Red bars indicate the mean EC_{50} values. (D) Wildtype SARS-CoV-2 RBD ELISA EC_{50} s plotted
 630 against EC_{50} s for B.1.1.7 RBD (E) B.1.351 RBD. (F) Proportion of SARS-CoV-2 RBD binders
 631 with detectable ELISA affinity for variants of concern (VOC) B.1.1.7 and B.1.351 RBDs. (G)
 632 ELISA EC_{50} values to related sarbecovirus RBDs displayed in decreasing order of paired-sequence
 633 identity.



634
635

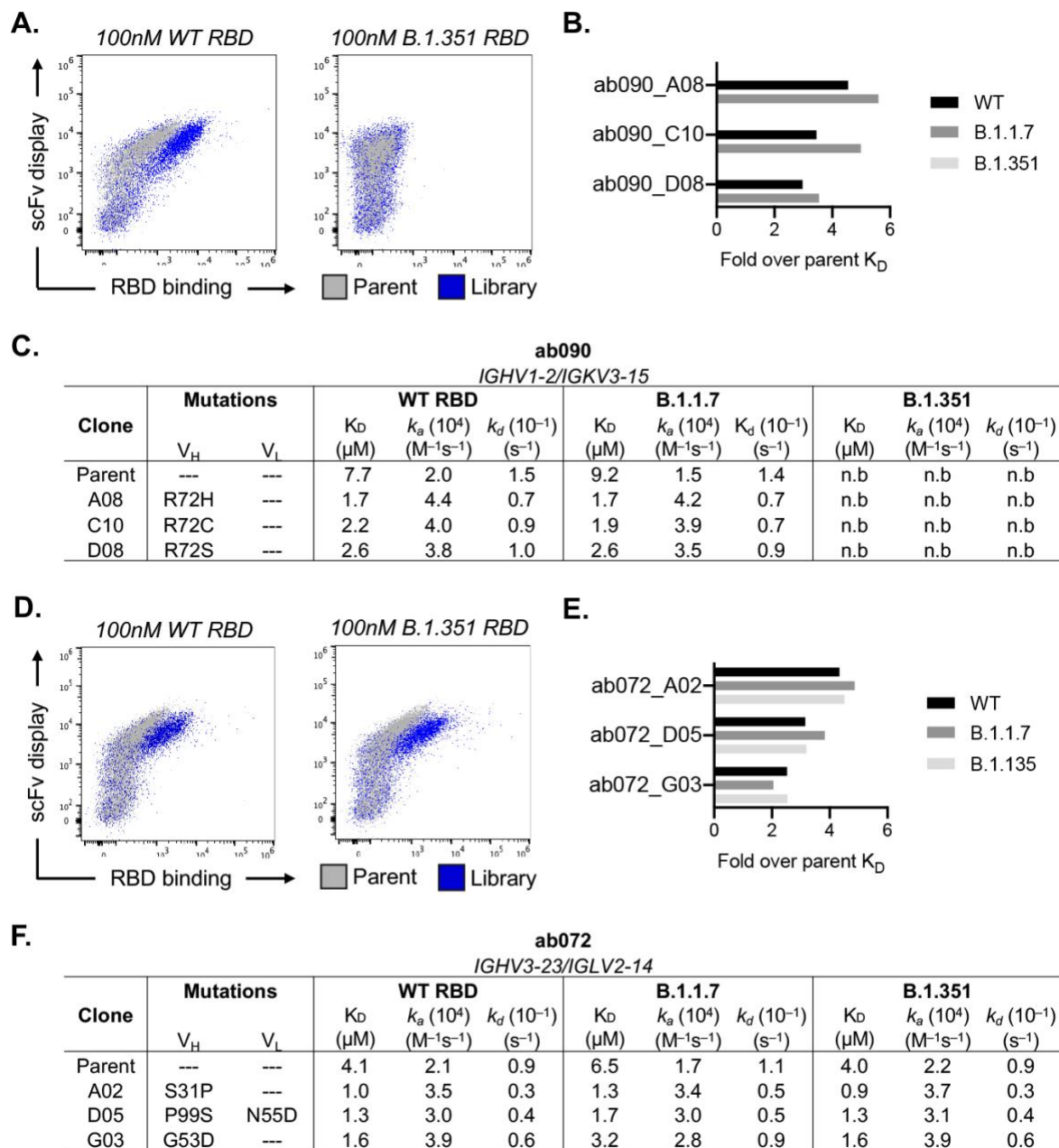
636 **Fig. 3. *In vitro* reconstitution of naive B cell activation.** (A) ELISA binding reactivity shows
637 restricted specificity of ab090 and (B) broad binding of ab072 to wildtype (WT) SARS-CoV-2,
638 B.1.1.7, B.1.351, and WIV1 RBDs. (C) BCR activation as measured by calcium flux in a Ramos
639 B cell line expressing ab090 membrane-anchored IgM (mIgM) and (D) and ab072 mIgM in
640 response to ferritin nanoparticles (NPs) displaying WT SARS-CoV-2, B.1.1.7, B.1.351, and WIV1
641 RBDs. Influenza hemagglutinin (HA) NP was used as a negative control.



642
643

644 **Fig. 4. ab090 recognizes the SARS-CoV-2 RBM.** (A) Cryo-EM structure of the SARS-CoV-2
645 spike trimer (grey) with ab090 Fab bound to one RBD in the up position. (B) ab090 recognizes the
646 SARS-CoV-2 RBM with through a paratope centered on the V_H (blue). (C) Close-up view showing
647 the approximate locations of HCDR loops proximal to the RBM epitope and B.1.351 RBD
648 mutations highlighted in red (top). ELISA binding reactivity of ab090 to individual mutations from
649 B.1.351 RBD (bottom). (D) ab090 binds to the RBM with a similar mode and angle of approach
650 to IGHV1-2 neutralizing antibodies isolated from memory B cells from convalescent COVID-19
651 donors. RBDs (grey) are shown in the same relative orientation in each panel with PBD codes and
652 sequence attributes listed in the below.

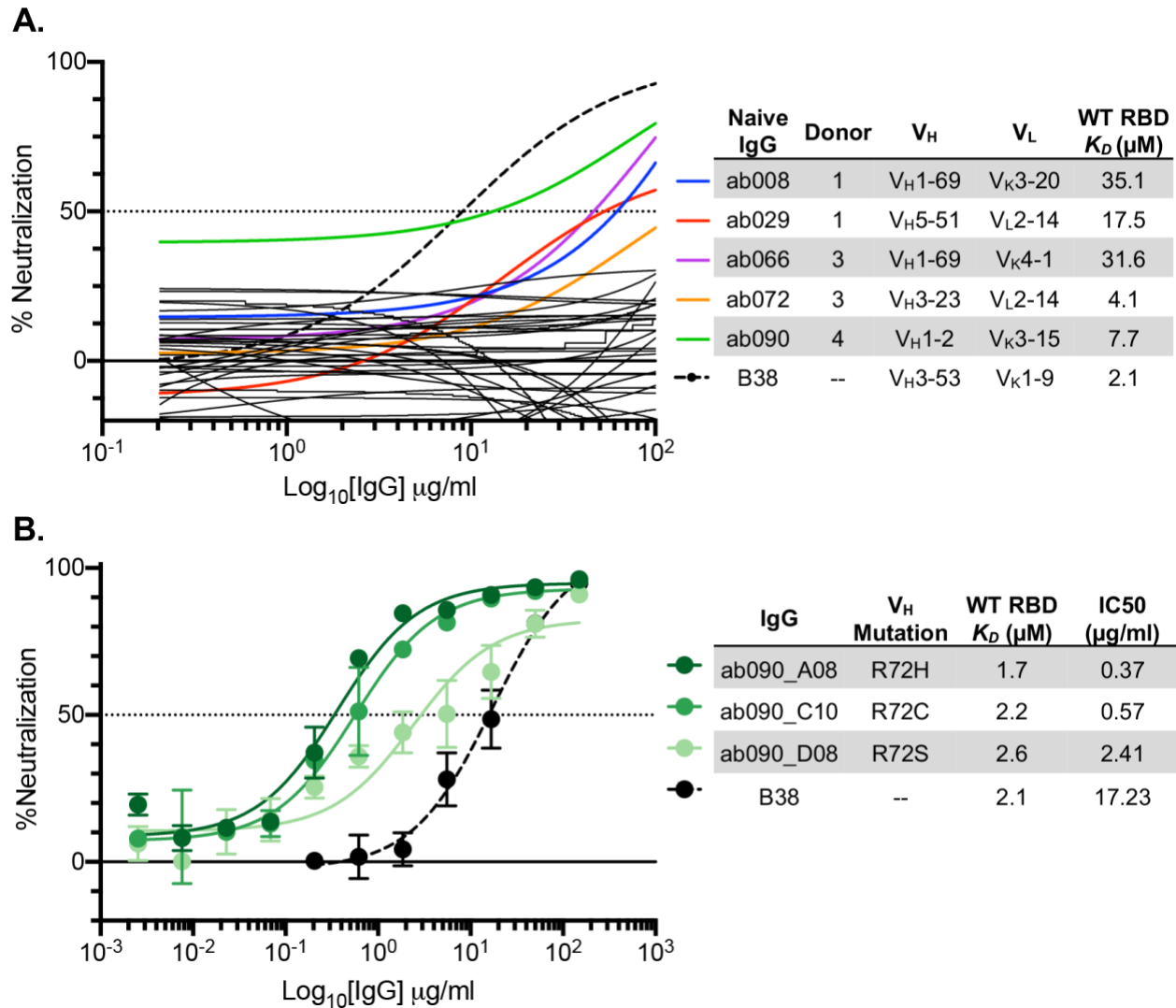
653



654

655

656 **Fig. 5. *in vitro* affinity-matured naive antibodies retain intrinsic specificity.** (A) Enrichment
 657 of ab090 parent (grey) and affinity matured (blue) libraires to 100nM SARS-CoV-2 or B.1.351
 658 RBD using flow cytometry (B) Fold enrichment in monovalent K_D over ab090 parent for selected
 659 affinity matured progeny. (C) Kinetic analysis using biolayer interferometry (BLI) for ab090
 660 parent and progeny Fabs to monomeric WT and variant RBDs. (D) Enrichment of ab072 parent
 661 (grey) and affinity matured (blue) libraires to 100nM SARS-CoV-2 or B.1.351 RBD using flow
 662 cytometry (E) Fold enrichment in monovalent K_D over ab072 parent for selected affinity matured
 663 progeny. (F) Kinetic analysis using biolayer interferometry (BLI) for ab072 parent and progeny
 664 Fabs to monomeric WT and variant RBDs.



665
666

667 **Fig. 6. SARS-CoV-2 pseudovirus neutralization by naive and affinity-matured Abs.** (A)
668 SARS-CoV-2 pseudovirus neutralization assay for 36 purified IgGs. Curves in color highlighted
669 antibodies with neutralizing activity with donor and monovalent wild-type RBD affinity listed for
670 this subset of antibodies. The neutralizing monoclonal antibody, B38, was used as a positive
671 control. Dashed lines indicate IC₅₀ values and data represent means ± SD of two technical
672 replicates. (B) SARS-CoV-2 pseudovirus neutralization for select affinity matured progeny from
673 the ab090 lineage with respective mutations relative to ab090 parent sequence, monovalent wild-
674 type RBD affinity, and IC₅₀ listed.

675 **REFERENCES**

- 676
- 677 1. N. E. Harwood, F. D. Batista, Early events in B cell activation. *Annual review of*
- 678 *immunology* **28**, 185-210 (2009).
- 679 2. G. D. Victora, M. C. Nussenzweig, Germinal centers. *Annual review of immunology* **30**,
- 680 429-457 (2012).
- 681 3. C. Havenar-Daughton, R. K. Abbott, W. R. Schief, S. Crotty, When designing vaccines,
- 682 consider the starting material: the human B cell repertoire. *Current opinion in*
- 683 *immunology* **53**, 209-216 (2018).
- 684 4. F. Sallusto, A. Lanzavecchia, K. Araki, R. Ahmed, From vaccines to memory and back.
- 685 *Immunity* **33**, 451-463 (2010).
- 686 5. W. Deng *et al.*, Primary exposure to SARS-CoV-2 protects against reinfection in rhesus
- 687 macaques. *Science* **369**, 818-823 (2020).
- 688 6. A. Chandrashekar *et al.*, SARS-CoV-2 infection protects against rechallenge in rhesus
- 689 macaques. *Science*, (2020).
- 690 7. N. B. Mercado *et al.*, Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus
- 691 macaques. *Nature* **586**, 583-588 (2020).
- 692 8. Q. Gao *et al.*, Development of an inactivated vaccine candidate for SARS-CoV-2.
- 693 *Science* **369**, 77-81 (2020).
- 694 9. K. S. Corbett *et al.*, Immune Correlates of Protection by mRNA-1273 Immunization
- 695 against SARS-CoV-2 Infection in Nonhuman Primates. *bioRxiv*, 2021.2004.2020.440647
- 696 (2021).
- 697 10. W. F. Garcia-Beltran *et al.*, COVID-19 neutralizing antibodies predict disease severity
- 698 and survival. *Cell*, (2020).
- 699 11. D. S. Khoury *et al.*, Neutralizing antibody levels are highly predictive of immune
- 700 protection from symptomatic SARS-CoV-2 infection. *Nature Medicine*, (2021).
- 701 12. A. Addetia *et al.*, Neutralizing Antibodies Correlate with Protection from SARS-CoV-2 in
- 702 Humans during a Fishery Vessel Outbreak with a High Attack Rate. *Journal of Clinical*
- 703 *Microbiology* **58**, e02107-02120 (2020).
- 704 13. S. F. Lumley *et al.*, Antibodies to SARS-CoV-2 are associated with protection against
- 705 reinfection. *medRxiv*, 2020.2011.2018.20234369 (2020).
- 706 14. J. S. Turner *et al.*, SARS-CoV-2 infection induces long-lived bone marrow plasma cells
- 707 in humans. *Nature*, 1-8 (2021).
- 708 15. A. Sokal *et al.*, Maturation and persistence of the anti-SARS-CoV-2 memory B cell
- 709 response. *bioRxiv*, 2020.2011.2017.385252 (2020).
- 710 16. C. Gaebler *et al.*, Evolution of antibody immunity to SARS-CoV-2. *Nature*, (2021).
- 711 17. J. M. Dan *et al.*, Immunological memory to SARS-CoV-2 assessed for up to 8 months
- 712 after infection. *Science*, eabf4063 (2021).
- 713 18. L. B. Rodda *et al.*, Functional SARS-CoV-2-specific immune memory persists after mild
- 714 COVID-19. *Cell*, (2020).
- 715 19. M. Sakharkar *et al.*, Prolonged evolution of the human B cell response to SARS-CoV-2
- 716 infection. *Science immunology* **6**, (2021).
- 717 20. E. J. Anderson *et al.*, Safety and immunogenicity of SARS-CoV-2 mRNA-1273 vaccine
- 718 in older adults. *New England Journal of Medicine*, (2020).
- 719 21. E. E. Walsh *et al.*, Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine
- 720 Candidates. *New England Journal of Medicine* **383**, 2439-2450 (2020).
- 721 22. Z. Wang *et al.*, mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating
- 722 variants. *Nature*, (2021).
- 723 23. Z. Wang *et al.*, mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating
- 724 variants. *bioRxiv*, 2021.2001.2015.426911 (2021).

- 725 24. P. J. M. Brouwer *et al.*, Potent neutralizing antibodies from COVID-19 patients define
726 multiple targets of vulnerability. *Science* **369**, 643-650 (2020).
- 727 25. B. Ju *et al.*, Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature*
728 **584**, 115-119 (2020).
- 729 26. C. Kreer *et al.*, Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing
730 antibodies from COVID-19 patients. *Cell* **182**, 843-854. e812 (2020).
- 731 27. D. F. Robbiani *et al.*, Convergent antibody responses to SARS-CoV-2 in convalescent
732 individuals. *Nature* **584**, 437-442 (2020).
- 733 28. A. Z. Wec *et al.*, Broad neutralization of SARS-related viruses by human monoclonal
734 antibodies. *Science* **369**, 731-736 (2020).
- 735 29. S. J. Zost *et al.*, Rapid isolation and profiling of a diverse panel of human monoclonal
736 antibodies targeting the SARS-CoV-2 spike protein. *Nature Medicine* **26**, 1422-1427
737 (2020).
- 738 30. C. O. Barnes *et al.*, SARS-CoV-2 neutralizing antibody structures inform therapeutic
739 strategies. *Nature*, (2020).
- 740 31. C. O. Barnes *et al.*, Structures of Human Antibodies Bound to SARS-CoV-2 Spike
741 Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **182**, 828-
742 842.e816 (2020).
- 743 32. L. Piccoli *et al.*, Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2
744 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell*
745 **183**, 1024-1042.e1021 (2020).
- 746 33. R. Shi *et al.*, A human neutralizing antibody targets the receptor binding site of SARS-
747 CoV-2. *Nature*, 1-8 (2020).
- 748 34. Y. Wu *et al.*, A noncompeting pair of human neutralizing antibodies block COVID-19
749 virus binding to its receptor ACE2. *Science* **368**, 1274-1278 (2020).
- 750 35. M. Yuan *et al.*, Structural basis of a shared antibody response to SARS-CoV-2. *Science*
751 **369**, 1119-1123 (2020).
- 752 36. S. C. A. Nielsen *et al.*, Human B Cell Clonal Expansion and Convergent Antibody
753 Responses to SARS-CoV-2. *Cell Host & Microbe* **28**, 516-525.e515 (2020).
- 754 37. M. Sangesland *et al.*, Germline-encoded affinity for cognate antigen enables vaccine
755 amplification of a human broadly neutralizing response against influenza virus. *Immunity*
756 **51**, 735-749. e738 (2019).
- 757 38. P. Dosenovic *et al.*, Anti-HIV-1 B cell responses are dependent on B cell precursor
758 frequency and antigen-binding affinity. *Proceedings of the National Academy of*
759 *Sciences*, 201803457 (2018).
- 760 39. R. K. Abbott *et al.*, Precursor frequency and affinity determine B cell competitive fitness
761 in germinal centers, tested with germline-targeting HIV vaccine immunogens. *Immunity*
762 **48**, 133-146. e136 (2018).
- 763 40. A. Amitai *et al.*, Defining and Manipulating B cell Immunodominance Hierarchies to Elicit
764 Broadly Neutralizing Antibody Responses Against Influenza Virus. *Cell Systems* **11**, 573-
765 588. e579 (2020).
- 766 41. M. Rapp *et al.*, Modular basis for potent SARS-CoV-2 neutralization by a prevalent VH1-
767 2-derived antibody class. *Cell Reports* **35**, 108950 (2021).
- 768 42. C. Havenar-Daughton *et al.*, The human naive B cell repertoire contains distinct
769 subclasses for a germline-targeting HIV-1 vaccine immunogen. *Science translational*
770 *medicine* **10**, eaat0381 (2018).
- 771 43. J. G. Jardine *et al.*, HIV-1 broadly neutralizing antibody precursor B cells revealed by
772 germline-targeting immunogen. *Science* **351**, 1458-1463 (2016).
- 773 44. J. M. Steichen *et al.*, A generalized HIV vaccine design strategy for priming of broadly
774 neutralizing antibody responses. *Science* **366**, (2019).

- 775 45. B. M. Hauser *et al.*, Engineered receptor binding domain immunogens elicit pan-
776 coronavirus neutralizing antibodies. *bioRxiv*, 2020.2012.2007.415216 (2020).
- 777 46. B. Briney, A. Inderbitzin, C. Joyce, D. R. Burton, Commonality despite exceptional
778 diversity in the baseline human antibody repertoire. *Nature* **566**, 393-397 (2019).
- 779 47. B. Ju *et al.*, Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature*
780 **584**, 115-119 (2020).
- 781 48. I. C. MacLennan *et al.*, Extrafollicular antibody responses. *Immunological reviews* **194**,
782 8-18 (2003).
- 783 49. Y. Cao *et al.*, Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-
784 Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **182**, 73-
785 84.e16 (2020).
- 786 50. C. Soto *et al.*, High frequency of shared clonotypes in human B cell receptor repertoires.
787 *Nature* **566**, 398-402 (2019).
- 788 51. A. G. Schmidt *et al.*, Viral receptor-binding site antibodies with diverse germline origins.
789 *Cell* **161**, 1026-1034 (2015).
- 790 52. T. Zhou *et al.*, Multidonor analysis reveals structural elements, genetic determinants,
791 and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity*
792 **39**, 245-258 (2013).
- 793 53. A. P. West, R. Diskin, M. C. Nussenzweig, P. J. Bjorkman, Structural basis for germ-line
794 gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1
795 gp120. *Proceedings of the National Academy of Sciences* **109**, E2083-E2090 (2012).
- 796 54. D. Lingwood *et al.*, Structural and genetic basis for development of broadly neutralizing
797 influenza antibodies. *Nature* **489**, 566 (2012).
- 798 55. A. G. Schmidt *et al.*, Immunogenic stimulus for germline precursors of antibodies that
799 engage the influenza hemagglutinin receptor-binding site. *Cell reports* **13**, 2842-2850
800 (2015).
- 801 56. F. D. Batista, M. S. Neuberger, Affinity dependence of the B cell response to antigen: a
802 threshold, a ceiling, and the importance of off-rate. *Immunity* **8**, 751-759 (1998).
- 803 57. J. M. Tas *et al.*, Visualizing antibody affinity maturation in germinal centers. *Science* **351**,
804 1048-1054 (2016).
- 805 58. S. Cele *et al.*, Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent
806 plasma. *Nature* **593**, 142-146 (2021).
- 807 59. R. E. Chen *et al.*, Resistance of SARS-CoV-2 variants to neutralization by monoclonal
808 and serum-derived polyclonal antibodies. *Nature Medicine* **27**, 717-726 (2021).
- 809 60. P. Wang *et al.*, Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7.
810 *Nature* **593**, 130-135 (2021).
- 811 61. C. K. Wibmer *et al.*, SARS-CoV-2 501Y.V2 escapes neutralization by South African
812 COVID-19 donor plasma. *Nature Medicine* **27**, 622-625 (2021).
- 813 62. N. G. Davies *et al.*, Estimated transmissibility and impact of SARS-CoV-2 lineage
814 B.1.1.7 in England. *Science* **372**, eabg3055 (2021).
- 815 63. H. Tegally *et al.*, Emergence and rapid spread of a new severe acute respiratory
816 syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in
817 South Africa. *medRxiv*, 2020.2012.2021.20248640 (2020).
- 818 64. E. B. Hodcroft, CoVariants: SARS-CoV-2 Mutations and Variants of Interest.
819 <https://covariants.org/>, (2021).
- 820 65. W. F. Garcia-Beltran *et al.*, Multiple SARS-CoV-2 variants escape neutralization by
821 vaccine-induced humoral immunity. *Cell* **184**, 2372-2383. e2379 (2021).
- 822 66. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage
823 for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology* **5**, 562-569
824 (2020).

- 825 67. V. D. Menachery *et al.*, SARS-like WIV1-CoV poised for human emergence.
826 *Proceedings of the National Academy of Sciences* **113**, 3048-3053 (2016).
- 827 68. A. G. Wrobel *et al.*, SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform
828 on virus evolution and furin-cleavage effects. *Nature Structural & Molecular Biology* **27**,
829 763-767 (2020).
- 830 69. V. D. Menachery *et al.*, A SARS-like cluster of circulating bat coronaviruses shows
831 potential for human emergence. *Nature Medicine* **21**, 1508-1513 (2015).
- 832 70. A. A. Cohen *et al.*, Mosaic nanoparticles elicit cross-reactive immune responses to
833 zoonotic coronaviruses in mice. *Science* **371**, 735-741 (2021).
- 834 71. E. Goodwin *et al.*, Infants Infected with Respiratory Syncytial Virus Generate Potent
835 Neutralizing Antibodies that Lack Somatic Hypermutation. *Immunity* **48**, 339-349.e335
836 (2018).
- 837 72. H.-X. Liao *et al.*, Initial antibodies binding to HIV-1 gp41 in acutely infected subjects are
838 polyreactive and highly mutated. *Journal of Experimental Medicine* **208**, 2237-2249
839 (2011).
- 840 73. Z.-H. Zhou, A. G. Tzioufas, A. L. Notkins, Properties and function of polyreactive
841 antibodies and polyreactive antigen-binding B cells. *Journal of Autoimmunity* **29**, 219-
842 228 (2007).
- 843 74. L. Shehata *et al.*, Affinity Maturation Enhances Antibody Specificity but Compromises
844 Conformational Stability. *Cell Reports* **28**, 3300-3308.e3304 (2019).
- 845 75. N. E. Harwood, F. D. Batista, Early Events in B Cell Activation. *Annual Review of*
846 *Immunology* **28**, 185-210 (2010).
- 847 76. G. C. Weaver *et al.*, In vitro reconstitution of B cell receptor–antigen interactions to
848 evaluate potential vaccine candidates. *Nature protocols* **11**, 193 (2016).
- 849 77. B. Zakeri *et al.*, Peptide tag forming a rapid covalent bond to a protein, through
850 engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences* **109**,
851 E690-E697 (2012).
- 852 78. B. Zhang *et al.*, A platform incorporating trimeric antigens into self-assembling
853 nanoparticles reveals SARS-CoV-2-spike nanoparticles to elicit substantially higher
854 neutralizing responses than spike alone. *Scientific Reports* **10**, 18149 (2020).
- 855 79. M. Yuan, H. Liu, N. C. Wu, I. A. Wilson, Recognition of the SARS-CoV-2 receptor
856 binding domain by neutralizing antibodies. *Biochemical and Biophysical Research*
857 *Communications* **538**, 192-203 (2021).
- 858 80. V. Giudicelli *et al.*, IMGT/LIGM-DB, the IMGT® comprehensive database of
859 immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Research* **34**,
860 D781-D784 (2006).
- 861 81. M. Yuan *et al.*, Structural and functional ramifications of antigenic drift in recent SARS-
862 CoV-2 variants. *Science*, eabh1139 (2021).
- 863 82. D. K. Sethi, A. Agarwal, V. Manivel, K. V. Rao, D. M. Salunke, Differential epitope
864 positioning within the germline antibody paratope enhances promiscuity in the primary
865 immune response. *Immunity* **24**, 429-438 (2006).
- 866 83. L. C. James, P. Roversi, D. S. Tawfik, Antibody multispecificity mediated by
867 conformational diversity. *Science* **299**, 1362-1367 (2003).
- 868 84. Z. Guo, E. Kraka, D. Cremer, Description of local and global shape properties of protein
869 helices. *J Mol Model* **19**, 2901-2911 (2013).
- 870 85. G. Chao *et al.*, Isolating and engineering human antibodies using yeast surface display.
871 *Nature Protocols* **1**, 755-768 (2006).
- 872 86. D. Stadlbauer *et al.*, Repeated cross-sectional sero-monitoring of SARS-CoV-2 in New
873 York City. *Nature*, 1-5 (2020).
- 874 87. K. Vanshilla *et al.*, Kinetics and correlates of the neutralizing antibody response to
875 SARS-CoV-2 infection in humans. *Cell Host & Microbe* **29**, 917-929.e914 (2021).

- 876 88. A. J. Greaney *et al.*, Complete Mapping of Mutations to the SARS-CoV-2 Spike
877 Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host & Microbe*,
878 (2020).
- 879 89. E. M. Anderson *et al.*, Seasonal human coronavirus antibodies are boosted upon SARS-
880 CoV-2 infection but not associated with protection. *Cell* **184**, 1858-1864.e1810 (2021).
- 881 90. K. W. Ng *et al.*, Preexisting and de novo humoral immunity to SARS-CoV-2 in humans.
882 *Science* **370**, 1339-1343 (2020).
- 883 91. E. Shrock *et al.*, Viral epitope profiling of COVID-19 patients reveals cross-reactivity and
884 correlates of severity. *Science* **370**, eabd4250 (2020).
- 885 92. P. Nguyen-Contant *et al.*, S Protein-Reactive IgG and Memory B Cell Production after
886 Human SARS-CoV-2 Infection Includes Broad Reactivity to the S2 Subunit. *mBio* **11**,
887 e01991-01920 (2020).
- 888 93. C. Viant *et al.*, Antibody Affinity Shapes the Choice between Memory and Germinal
889 Center B Cell Fates. *Cell* **183**, 1298-1311. e1211 (2020).
- 890 94. R. Di Niro *et al.*, Salmonella infection drives promiscuous B cell activation followed by
891 extrafollicular affinity maturation. *Immunity* **43**, 120-131 (2015).
- 892 95. M. Kuraoka *et al.*, Complex antigens drive permissive clonal selection in germinal
893 centers. *Immunity* **44**, 542-552 (2016).
- 894 96. S. I. Kim *et al.*, Stereotypic neutralizing VH antibodies against SARS-CoV-2 spike
895 protein receptor binding domain in COVID-19 patients and healthy individuals. *Science*
896 *Translational Medicine*, (2021).
- 897 97. Z. Wang *et al.*, Naturally enhanced neutralizing breadth against SARS-CoV-2 one year
898 after infection. *Nature*, (2021).
- 899 98. P. W. H. I. Parren, D. R. Burton, in *Advances in Immunology*. (Academic Press, 2001),
900 vol. 77, pp. 195-262.
- 901 99. C. G. Rappazzo *et al.*, Broad and potent activity against SARS-like viruses by an
902 engineered human monoclonal antibody. *Science*, eabf4830 (2021).
- 903 100. L. R. Baden *et al.*, Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N Engl*
904 *J Med* **384**, 403-416 (2021).
- 905 101. D. M. Skowronski, G. De Serres, Safety and Efficacy of the BNT162b2 mRNA Covid-19
906 Vaccine. *N Engl J Med* **384**, 1576-1577 (2021).
- 907 102. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion
908 conformation. *Science* **367**, 1260-1263 (2020).
- 909 103. M. Sangesland *et al.*, A Single Human VH-gene Allows for a Broad-Spectrum Antibody
910 Response Targeting Bacterial Lipopolysaccharides in the Blood. *Cell Reports* **32**,
911 108065 (2020).
- 912 104. P. S. Tobias, K. Soldau, R. J. Ulevitch, Identification of a Lipid A Binding Site in the
913 Acute Phase Reactant Lipopolysaccharide Binding Protein. *Journal of Biological*
914 *Chemistry* **264**, 10867-10871 (1989).
- 915 105. J. J. Trombetta *et al.*, Preparation of single-cell RNA-seq libraries for next generation
916 sequencing. *Current protocols in molecular biology* **107**, 4.22. 21-24.22. 17 (2014).
- 917 106. A. P. Masella, A. K. Bartram, J. M. Truszkowski, D. G. Brown, J. D. Neufeld, PANDAseq:
918 paired-end assembler for illumina sequences. *BMC bioinformatics* **13**, 1-7 (2012).
- 919 107. B. Shi *et al.*, Comparative analysis of human and mouse immunoglobulin variable heavy
920 regions from IMGT/LIGM-DB with IMGT/HighV-QUEST. *Theoretical Biology and Medical*
921 *Modelling* **11**, 1-11 (2014).
- 922 108. M. J. Moore *et al.*, Retroviruses pseudotyped with the severe acute respiratory
923 syndrome coronavirus spike protein efficiently infect cells expressing angiotensin-
924 converting enzyme 2. *Journal of virology* **78**, 10628-10635 (2004).
- 925 109. E. Siebring-van Olst *et al.*, Affordable luciferase reporter assay for cell-based high-
926 throughput screening. *Journal of biomolecular screening* **18**, 453-461 (2013).

- 927 110. S. H. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure
928 determination. *J Struct Biol* **180**, 519-530 (2012).
- 929 111. S. Q. Zheng *et al.*, MotionCor2: anisotropic correction of beam-induced motion for
930 improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
- 931 112. K. Zhang, Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12
932 (2016).
- 933 113. T. Wagner *et al.*, SPHIRE-crYOLO is a fast and accurate fully automated particle picker
934 for cryo-EM. *Commun Biol* **2**, 218 (2019).
- 935 114. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for
936 rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
- 937 115. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research
938 and analysis. *J Comput Chem* **25**, 1605-1612 (2004).
- 939 116. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta*
940 *Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
- 941 117. J. Mata-Fink *et al.*, Rapid Conformational Epitope Mapping of Anti-gp120 Antibodies with
942 a Designed Mutant Panel Displayed on Yeast. *Journal of Molecular Biology* **425**, 444-
943 456 (2013).
- 944 118. M. I. J. Raybould, A. Kovaltsuk, C. Marks, C. M. Deane, CoV-AbDab: the coronavirus
945 antibody database. *Bioinformatics* **37**, 734-735 (2020).
- 946

947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977

Supplemental Information

Naive human B cells engage the receptor binding domain of SARS-CoV-2, variants of concern, and related sarbecoviruses

Jared Feldman^{1†}, Julia Bals^{1†}, Clara G. Altomare², Kerri St. Denis¹, Evan C. Lam¹, Blake M. Hauser¹, Larance Ronsard¹, Maya Sangesland¹, Thalia Bracamonte Moreno¹, Vintus Okonkwo¹, Nathania Hartojo¹, Alejandro B. Balazs¹, Goran Bajic², Daniel Lingwood^{1*} and Aaron G. Schmidt^{1,3*}

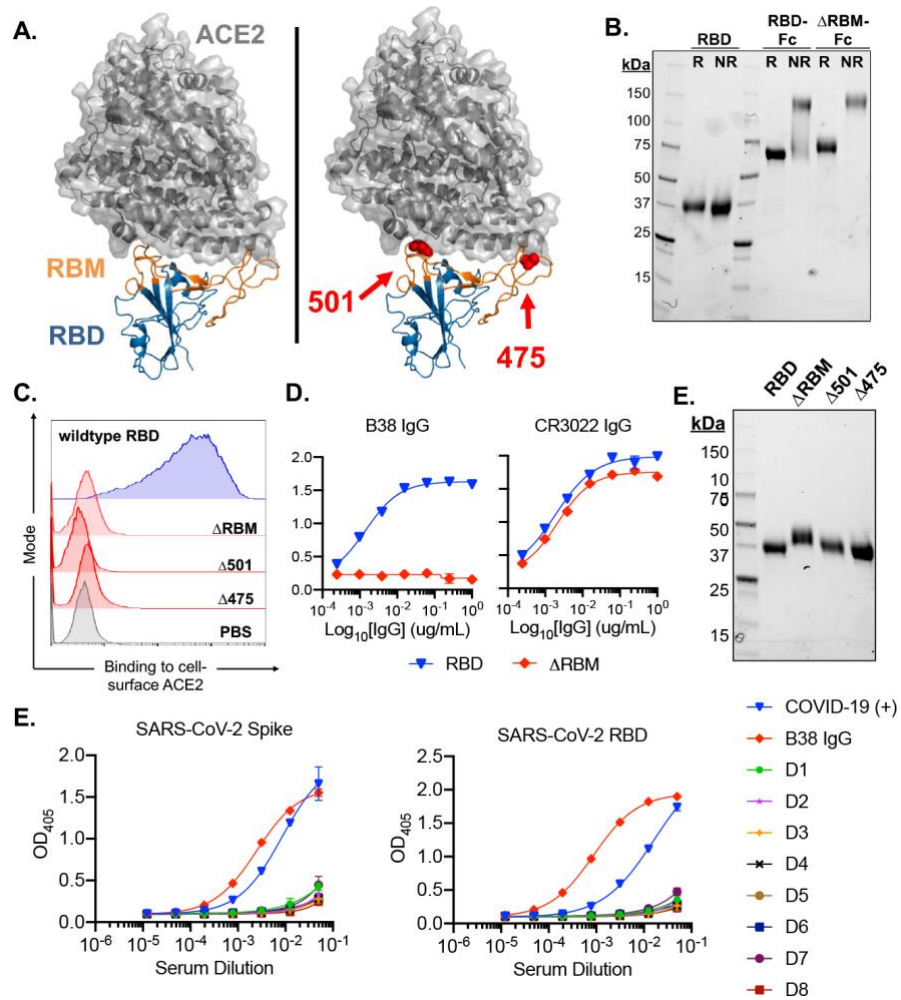
¹Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, 02139, USA

²Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029

³Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

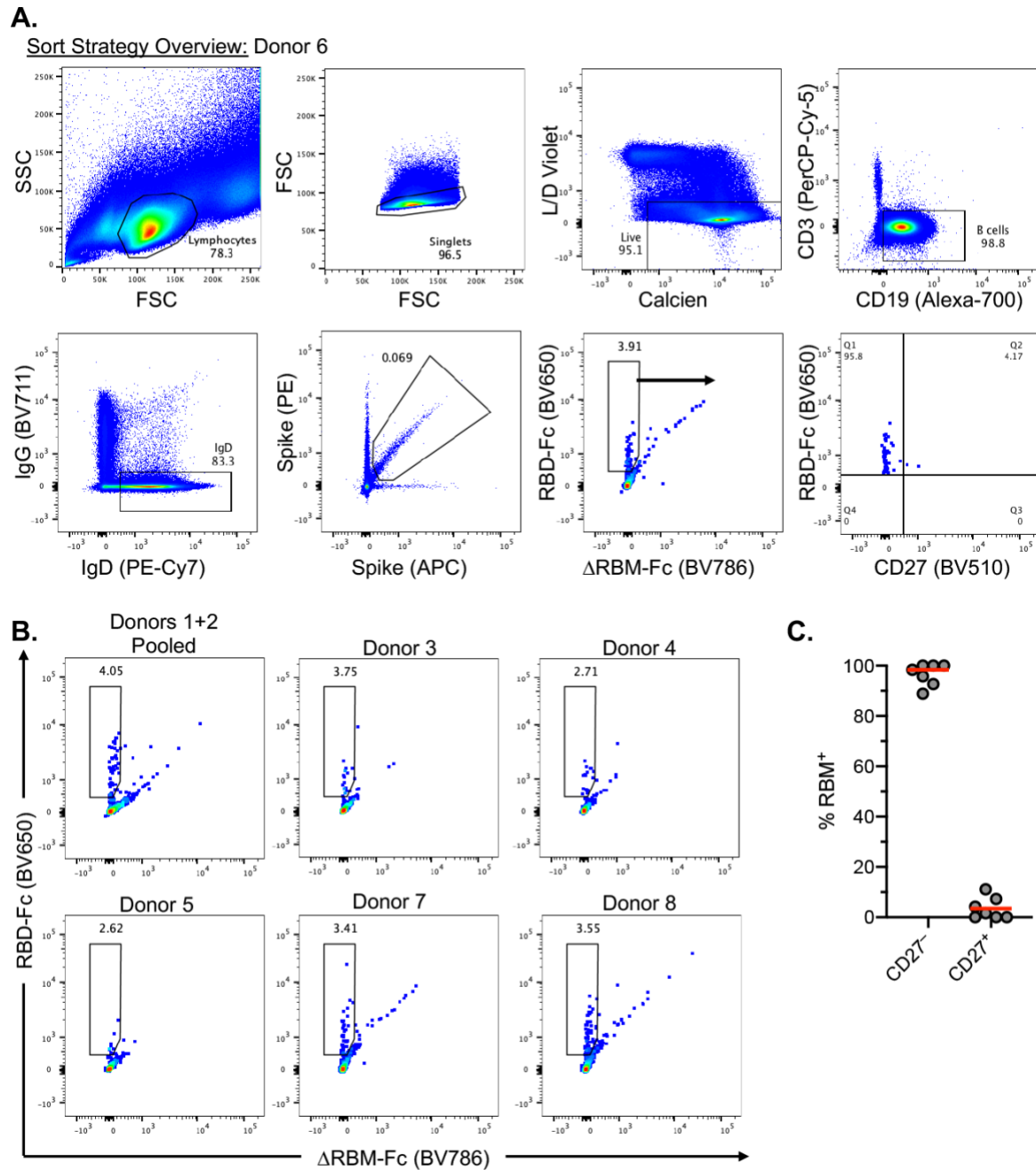
*Correspondence to: Email: dlingwood@mgh.harvard.edu (D.L.); aschmidt@crystal.harvard.edu (A.G.S)

†These authors contributed equally to this work.



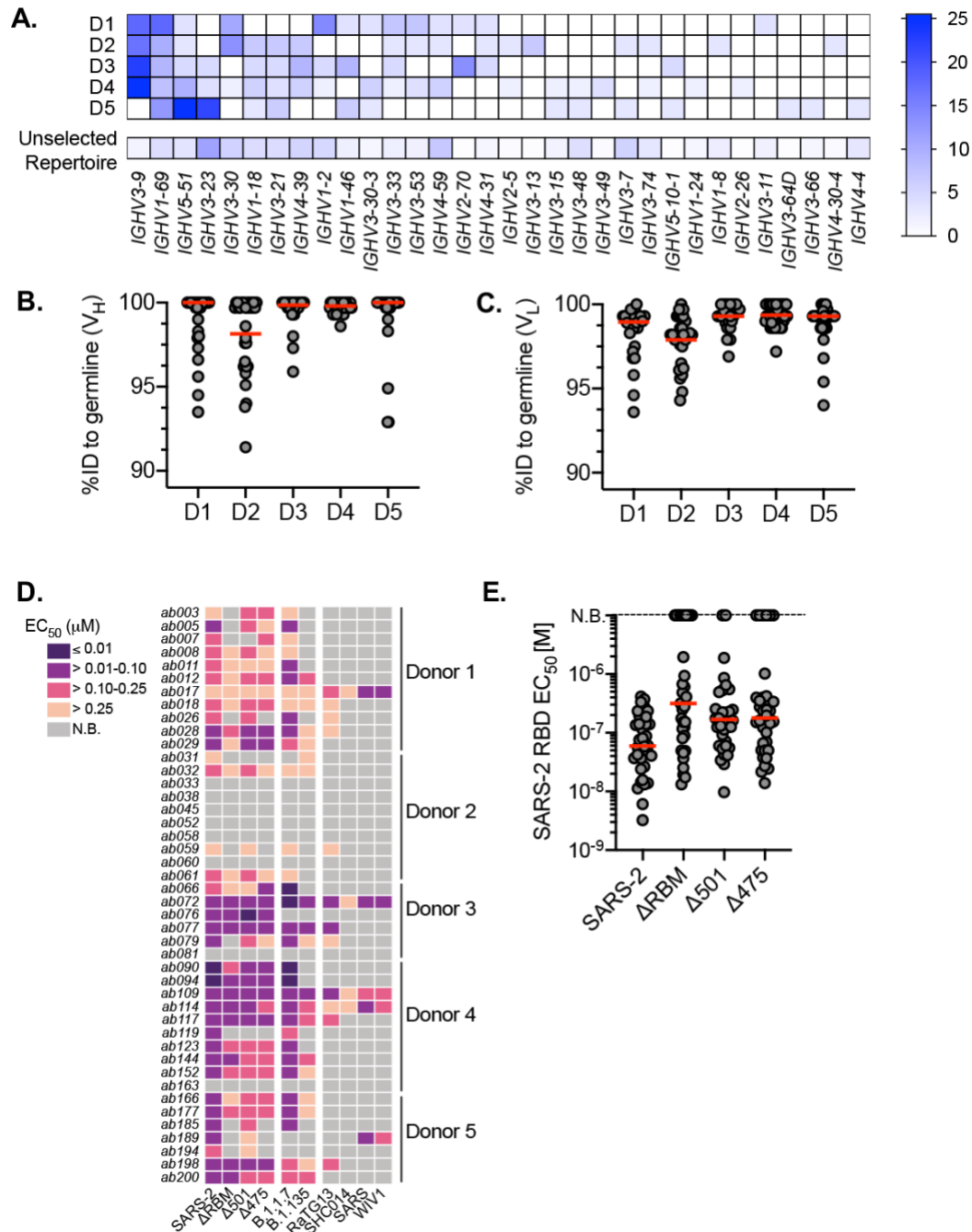
978
979

980 **fig. S1. Design and characterization of SARS-CoV-2 antigens and healthy donor sera**
 981 **binding.** (A) SARS-CoV-2 RBD in complex with viral receptor, ACE2 shown in blue and grey,
 982 respectively (PDB 6M0J). Wild-type RBD with, the receptor binding motif (RBM), shown in
 983 orange (left panel). Structural model of the Δ RBM probe designed to abrogate binding to ACE2
 984 (right panel). Putative N-linked glycosylation sites engineered onto the RBM are shown in red
 985 spheres at amino acid positions 501 and 475. (B) SDS-PAGE gel under reducing (R) and non-
 986 reducing (NR) conditions for monomeric RBD, RBD-Fc and Δ RBM-Fc. (C) Wildtype RBD,
 987 Δ RBM and single glycan variant binding to ACE2-expressing 293T cells by flow cytometry. Wild-
 988 type RBD binding shown in blue, glycan variant binding shown in red. Streptavidin-PE was used
 989 to detect the relative intensity of antigen binding to cell-surface ACE2. A PBS control (gray)
 990 indicates secondary-only staining. (D) Control antibody ELISA binding to RBD and Δ RBM
 991 antigens. RBM-specific antibody, B38 (left). Non-RBM-specific control antibody, CR3022
 992 (right). (E) Δ RBM and Δ 501 and Δ 475 variants analyzed by SDS-PAGE gel under reducing
 993 conditions; wildtype RBD is shown for comparison. (E) SARS-CoV-2 spike (left) and RBD (right)
 994 sera ELISA from human subjects 1-8. Sera from a COVID-19 convalescent patient and control
 995 antibody, B38, were included as positive controls.



996
997

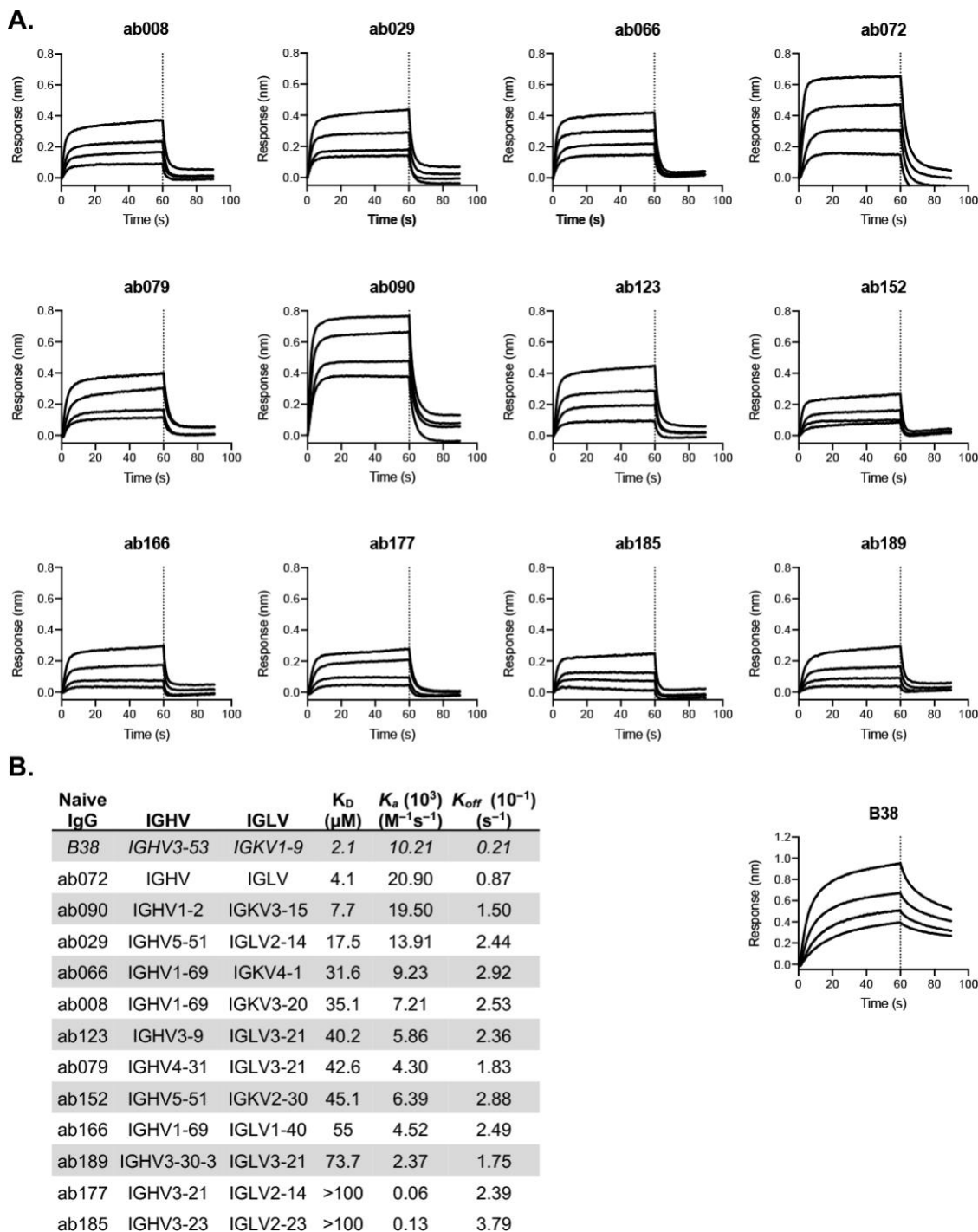
998 **fig. S2. PBMC flow cytometry analyses.** (A) Representative gating strategy used for FACS of
 999 PBMCs pooled from donors 1 and 2. Gating was on naive B cells defined by single living
 1000 lymphocytes that were CD19⁺CD3⁻IgD⁺IgG⁻. Sorted cells were RBM-specific as defined by spike-
 1001 PE⁺/spike-APC⁺/RBD-Fc-BV650⁺/ΔRBM-Fc-BC650⁻. Sort gate is denoted by the blue arrow. The
 1002 bottom right plot shows CD27 staining of sorted RBM-specific naive B cells. (B) Flow cytometry
 1003 showing the sort gate and percentage of RBM-specific B cells for the remaining 6 healthy human
 1004 donors. (C) RBM-specific B cell frequency among CD27⁺ and CD27⁻ cells. Each symbol
 1005 represents a different donor ($n = 8$).



1006
1007

1008 **fig. S3. Repertoire comparison, germline identity, and IgG binding by individual donor.** (A)
1009 Heatmap showing V_H-gene usage of isolated antibodies derived from donors 1-5. Unselected
1010 repertoire gene usage derive from a high-throughput sequencing data set of circulating B cells
1011 across 10 human subjects (46). Heatmap scale represents percent of total paired sequence from
1012 each donor. Divergence from inferred germline gene sequences separated by individual donor for
1013 (B) V_H and (C) V_L. Red bars indicate the median percent values, and each dot represents an
1014 individual paired sequence. (D) Heatmap showing IgG binding to RBDs (n = 44) sorted by donor.
1015 (E) ELISA EC₅₀ values for IgGs with detectable SARS-CoV-2 RBD binding (n = 36) against RBM
1016 glycan probes. Red bars indicate the mean EC₅₀ values.

1017

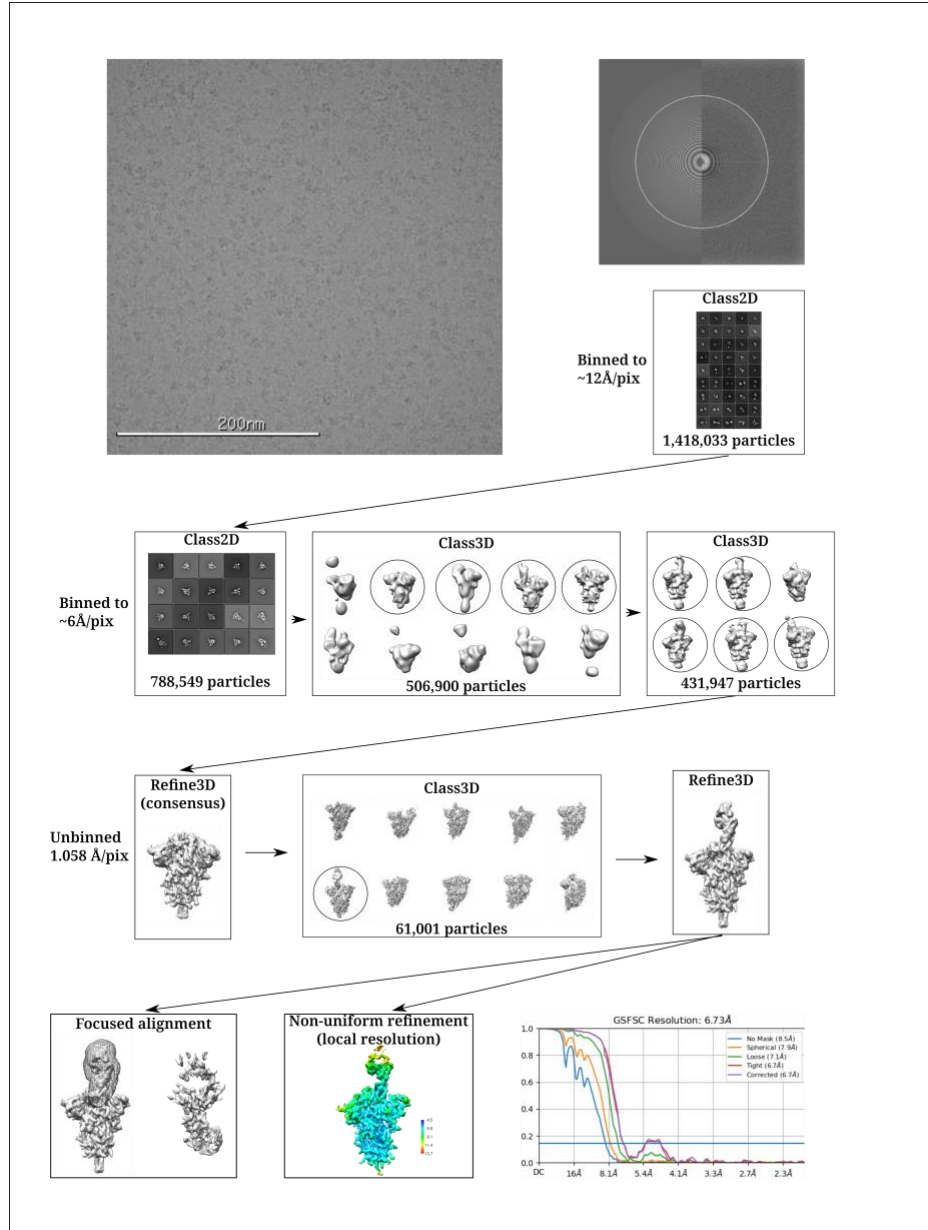


1018

1019

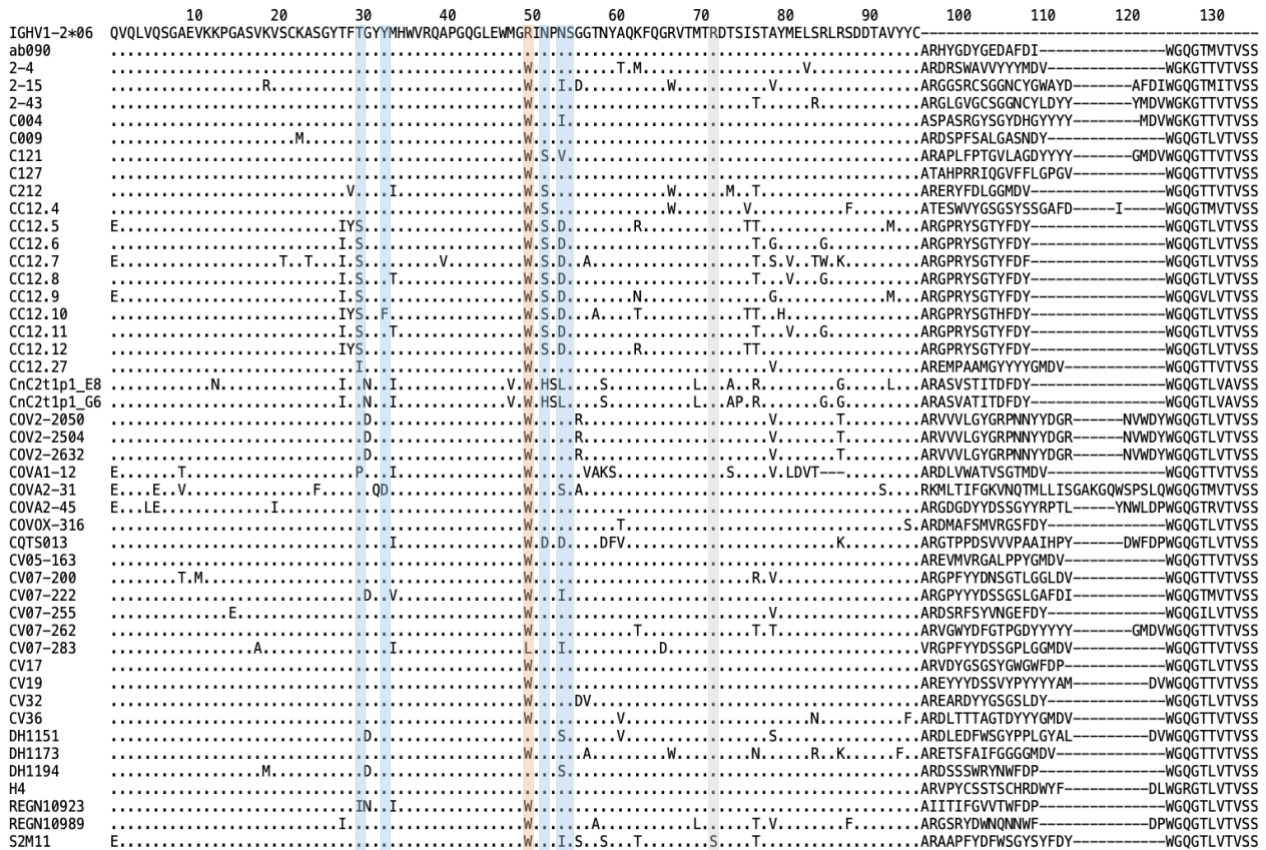
1020 **fig. S4. SARS-CoV-2 RBD-binding kinetics of isolated naive antibodies.** (A) Biolayer
 1021 interferometry (BLI) binding kinetic analysis of titrated SARS-CoV-2 RBD to immobilized Fabs.
 1022 Dotted line at 60 s denotes the start of the dissociation phase. (B) Kinetic and equilibrium constants
 1023 for binding to RBD calculated from a 1:1 binding model using a global fit to all curves for each
 1024 Fab using vendor supplied software. B38 Fab is used as a positive control.

1025
1026 **A.**



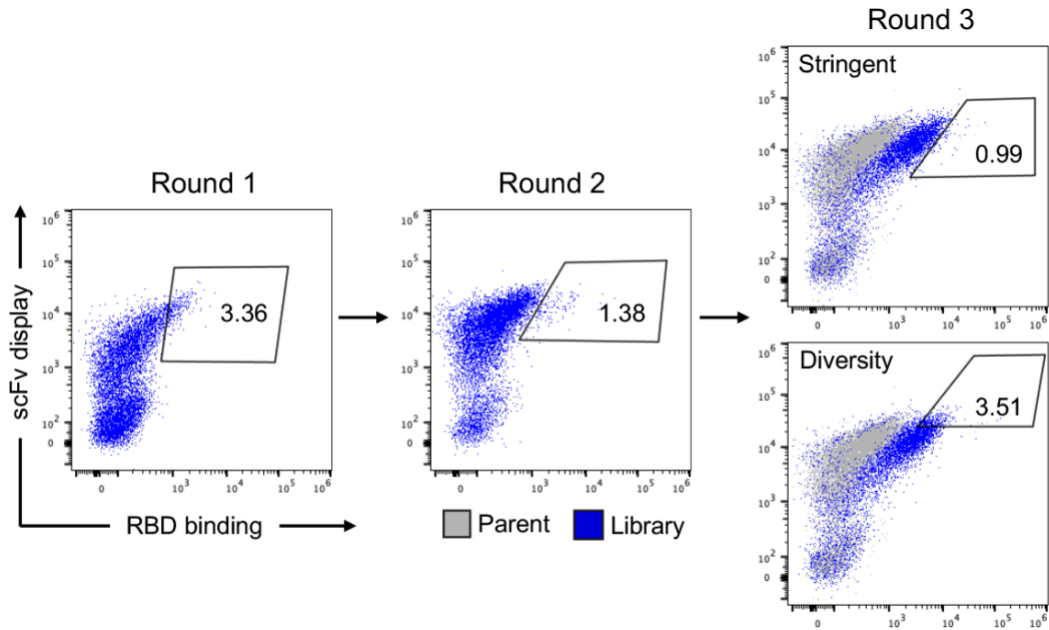
1027
1028

1029 **B.**

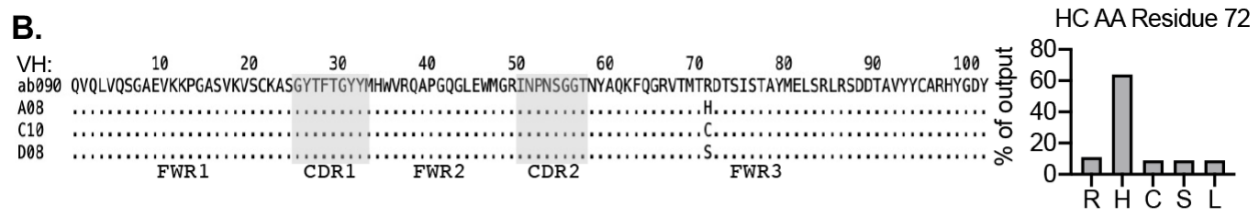


1030 **fig. S5. Structural characterization and analysis. (A)** Cryo-EM data processing scheme of
1031 ab090 Fab bound with SARS-CoV-2 spike. See the Methods section for more details. **(B)** Heavy
1032 chain amino acid sequence alignment of ab090 with IGHV1-2 derived antibodies from
1033 convalescent COVID-19 patients. Sequences were obtained from CoV-AbDab (118) and aligned
1034 to the IGHV1-2*06 reference. Residues forming the germline-encoded HCDR1 and HCDR2 motif
1035 contacting the SARS-CoV-2 RBD are highlighted in blue. The single nucleotide polymorphism in
1036 the *06 allele at position 50 is highlighted red. The site of the dominant mutation from *in vitro*
1037 affinity maturation efforts with ab090 is highlighted in green.

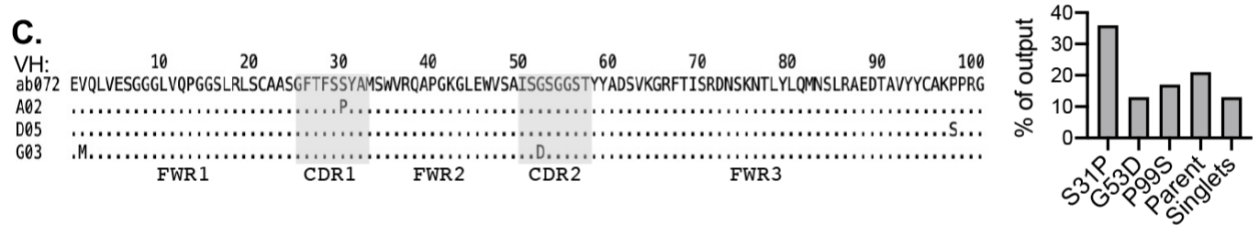
A.



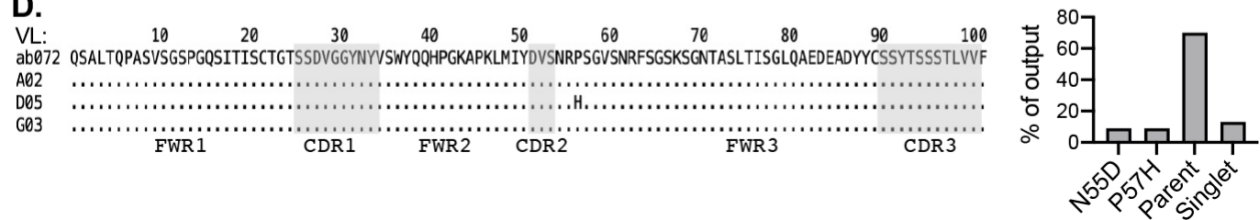
B.



C.



D.



1038
1039
1040
1041
1042
1043
1044
1045
1046
1047

fig. S6. Representative affinity maturation selection strategy and output sequence overview.

(A) Flow cytometric sorting of diversified single chain variable fragment (scFv) libraries of ab090. Gates represent the yeast population sorted for subsequent selections. After 2 rounds of enrichment for wildtype SARS-CoV-2 binding, a “stringent” and “diversity gate were sorted in round 3 indicating the yeast populations sorted for individual colony isolation and sequencing. Alignment of the V_H sequencing output clones for ab090 (B) and ab072 (C) with the output frequency of each mutation from a total of 48 single colonies. (D) Alignment of the V_L sequencing output clones ab072 with the output frequency of each mutation from a total of 48 single colonies. The V_L output for ab090 was exclusively parent.