

1 **Germinal centre-driven maturation of B cell response to SARS-CoV-2 vaccination**

2

3 Wooseob Kim^{1*}, Julian Q. Zhou^{1*}, Alexandria J. Sturtz¹, Stephen C. Horvath¹, Aaron J. Schmitz¹,
4 Tingting Lei¹, Elizaveta Kalaidina², Mahima Thapa¹, Wafaa B. Alsoussi¹, Alem Haile³, Michael K.
5 Klebert³, Teresa Suessen⁴, Luis Parra-Rodriguez⁵, Philip A. Mudd^{6,7}, William D. Middleton⁴,
6 Sharlene A. Teefey⁴, Iskra Pusic⁸, Jane A. O'Halloran⁵, Rachel M. Presti^{5,7}, Jackson S. Turner¹,
7 and Ali H. Ellebedy^{1,7,9#}

8

9 ¹Department of Pathology and Immunology, Washington University School of Medicine, St.
10 Louis, MO, USA; ²Division of Allergy and Immunology, Department of Internal Medicine,
11 Washington University School of Medicine, St. Louis, MO, USA; ³Clinical Trials Unit,
12 Washington University School of Medicine, St. Louis, MO, USA; ⁴Mallinckrodt Institute of
13 Radiology, Washington University School of Medicine, St. Louis, MO, USA; ⁵Division of
14 Infectious Diseases, Department of Internal Medicine, Washington University School of
15 Medicine, St. Louis, MO, USA; ⁶Department of Emergency Medicine, Washington University
16 School of Medicine, St. Louis, MO, USA; ⁷Center for Vaccines and Immunity to Microbial
17 Pathogens, Washington University School of Medicine, St. Louis, MO; ⁸Division of Oncology,
18 Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA; ⁹The
19 Andrew M. and Jane M. Bursky Center for Human Immunology & Immunotherapy Programs,
20 Washington University School of Medicine, St. Louis, MO, USA.

21

22 *These authors contributed equally to this work

23 # Corresponding author: Ali H. Ellebedy, Ph.D., ellebedy@wustl.edu

24

25 **Abstract**

26 Germinal centres (GC) are lymphoid structures where vaccine-responding B cells acquire
27 affinity-enhancing somatic hypermutations (SHM), with surviving clones differentiating into
28 memory B cells (MBCs) and long-lived bone marrow plasma cells (BMPCs)¹⁻⁴. Induction of the
29 latter is a hallmark of durable immunity after vaccination⁵. SARS-CoV-2 mRNA vaccination
30 induces a robust GC response in humans⁶⁻⁸, but the maturation dynamics of GC B cells and
31 propagation of their progeny throughout the B cell diaspora have not been elucidated. Here we
32 show that anti-SARS-CoV-2 spike (S)-binding GC B cells were detectable in draining lymph
33 nodes for at least six months in 10 out of 15 individuals who had received two doses of
34 BNT162b2, a SARS-CoV-2 mRNA vaccine. Six months after vaccination, circulating S-binding
35 MBCs were detected in all participants (n=42) and S-specific IgG-secreting BMPCs were
36 detected in 9 out of 11 participants. Using a combined approach of single-cell RNA sequencing
37 of responding blood and lymph node B cells from eight participants and expression of the
38 corresponding monoclonal antibodies, we tracked the evolution of 1540 S-specific B cell clones.
39 SHM accumulated along the B cell differentiation trajectory, with early blood plasmablasts
40 showing the lowest frequencies, followed by MBCs and lymph node plasma cells whose SHM
41 largely overlapped with GC B cells. By three months after vaccination, the frequency of SHM
42 within GC B cells had doubled. Strikingly, S⁺ BMPCs detected six months after vaccination
43 accumulated the highest level of SHM, corresponding with significantly enhanced anti-S
44 polyclonal antibody avidity in blood at that time point. This study documents the induction of
45 affinity-matured BMPCs after two doses of SARS-CoV-2 mRNA vaccination in humans,
46 providing a foundation for the sustained high efficacy observed with these vaccines.

47

48 **Main text**

49 Phase 3 trials of SARS-CoV-2 mRNA-based vaccines demonstrated the remarkable
50 efficacy of these vaccines against coronavirus disease 2019 (COVID-19)^{9,10}. The Pfizer-

51 BioNTech SARS-CoV-2 mRNA vaccine (BNT162b2) remains highly efficacious at preventing
52 symptomatic infection for at least six months after the initial two-dose vaccine series in
53 individuals with no history of SARS-CoV-2 infection¹¹. This is despite a decrease in S-binding
54 antibody concentrations and neutralizing antibodies that occur during the same period^{12,13}. We
55 have previously shown that vaccination of human subjects with BNT162b2 induces a robust but
56 transient circulating plasmablast (PB) response followed by remarkably persistent germinal
57 centre (GC) reactions in the draining lymph nodes⁶. Whether these persistent GC responses
58 lead to the generation of affinity-matured bone marrow-resident plasma cells remains unclear.
59 To address this question, we undertook longitudinal follow-up of participants enrolled in our
60 previously described observational study of 43 healthy participants (13 with a history of
61 confirmed SARS-CoV-2 infection) who received two doses of the Pfizer-BioNTech SARS-CoV-2
62 mRNA vaccine (BNT162b2) (**Extended Data Tables 1, 2**)^{6,7}. Long-term blood samples (n=42),
63 fine needle aspirates (FNAs) of the draining axillary lymph nodes (n=15) and bone marrow
64 aspirates (n=11) were collected 29 weeks after the primary vaccination (**Fig. 1a**). None of the
65 participants who contributed FNA or bone marrow specimens had a history of SARS-CoV-2
66 infection.

67 GC B cells were detected in FNAs from all 15 participants (**Fig. 1b, c, left panels,**
68 **Extended Data Fig. 1a, Extended Data Table 3**). All 14 participants with FNAs collected prior
69 to week 29 generated SARS-CoV-2 S-binding antigen-specific GC B cell responses of varying
70 magnitudes (**Fig 1b, c, right panels, and Extended Data Table 3**). Intriguingly, S-binding GC
71 B cells were detected in FNAs from 10 of 15 participants at week 29 (**Fig. 1b, c, right panels,**
72 **Extended Data Table 3**), demonstrating that more than half of the sampled subjects maintained
73 an antigen-specific GC B cell response more than 6 months after vaccination. S-binding LNPCs
74 were also detected in FNAs from all 15 participants and exhibited kinetics similar to S-binding
75 GC B cells, albeit at lower frequencies within the total B cell population (**Extended Data Fig. 1a-**
76 **b, Extended Data Table 3**). None of the FNAs demonstrated significant contamination with

77 peripheral blood based upon the nearly complete absence of CD14+ myeloid cells (**Extended**
78 **Data Table 4**).

79 To determine whether SARS-CoV-2 mRNA vaccination induces antigen-specific bone
80 marrow plasma cells (BMPCs), we examined bone marrow aspirates collected 29 weeks after
81 primary vaccination. We first magnetically enriched BMPCs and then measured the frequencies
82 of BMPCs secreting IgG or IgA antibodies against the 2019-2020 inactivated influenza virus
83 vaccine, the tetanus-diphtheria vaccine or SARS-CoV-2 S protein by enzyme-linked
84 immunosorbent spot assay (ELISpot) (**Fig. 1d-e, Extended Data Fig. 1c**). Influenza- and
85 tetanus-diphtheria vaccine-specific IgG-secreting BMPCs (median frequencies of 1.4% and
86 0.15%, respectively) were detected in all 11 participants (**Fig. 1e**). SARS-CoV-2 S-binding IgG-
87 secreting BMPCs were detected in 9 of 11 participants (median frequency of 0.06%). IgA-
88 secreting BMPCs specific for the seasonal influenza vaccine were detected in 10 of 11
89 participants, but IgA-secreting BMPCs directed against tetanus-diphtheria-vaccine and SARS-
90 CoV-2 S protein were largely below the limit of detection (**Extended Data Fig. 1c**). Importantly,
91 none of the participants had detectable S-specific antibody-secreting cells (ASCs) in blood at
92 the time of bone marrow collection, indicating that the detected BMPCs represent bona fide
93 bone marrow-resident cells and not contamination from circulating blood PBs.

94 All study participants had readily detectable plasma anti-S IgG antibodies and circulating S-
95 specific memory B cells (MBCs) at the 29 week time point (**Fig. 1f, g, and h**), consistent with
96 the presence of S-specific BMPCs and GC B cells in the bone marrow and lymph node,
97 respectively. Notably, anti-S IgG titers 29 weeks following vaccination were higher than titers
98 observed in a separate cohort of unvaccinated SARS-CoV-2-infected convalescent subjects
99 measured 29 weeks post-infection¹⁴⁻¹⁶ (**Extended Data Fig. 1d, Extended Data Table 1**).
100 Vaccinated participants with a history of confirmed SARS-CoV-2 infection showed significantly
101 higher titers of plasma anti-S IgG 5 weeks and 29 weeks following vaccination than their naive
102 counterparts^{14,16,17} (**Fig. 1f**). Similar trends were observed for plasma anti-S IgM and IgA

103 **(Extended Data Fig. 1e)**. SARS-CoV-2 S-binding MBCs were observed in all participants, with
104 a median frequency of 0.23% of total CD19⁺ circulating B cells(**Fig. 1g, h, Extended Data Fig.**
105 **1f)**. . Collectively, these results demonstrate that SARS-CoV-2 mRNA vaccination generates
106 robust, persistent humoral responses.

107 To investigate SARS-CoV-2 S-specific B cell composition and clonal distribution within
108 blood, lymph node and bone marrow, we performed single-cell RNA sequencing (scRNA-seq)
109 for both transcriptomic and B cell receptor (BCR) repertoire profiling on 8 participants who
110 contributed samples from all three of these tissue compartments. To maximize the number of
111 blood PBs characterized by scRNA-seq, we sorted PBs from blood samples collected at the
112 peak of the circulating PB response, 4 weeks after primary vaccination⁶(**Extended Data Fig.**
113 **2a)**. We selected 7 and 15 weeks post-vaccination as representative early and late time points
114 to interrogate immune responses in the lymph node, except that we were unable to include
115 samples from 2 of the 8 participants at the late time points due to insufficient sample remaining.
116 For these two subjects, we analyzed two separate relatively early time points - weeks 5 and 7
117 for participant 02a, and weeks 4 and 7 for participant 04. Single-cell transcriptome analysis of
118 lymph nodes revealed distinct populations of T and B cells, NK cells, monocytes, plasmacytoid
119 dendritic cells, and follicular dendritic cells, as previously described¹⁸⁻²¹(**Fig. 2a, bottom left**
120 **panel, Extended Data Fig. 2b, c, Extended Data Table 5)**. To further distinguish between
121 transcriptionally distinct subsets of B cells in the lymph node following SARS-CoV-2 vaccination,
122 we performed unbiased clustering of the B cell populations from the total cellular analysis (**Fig**
123 **2a, bottom left panel, Extended Data Fig. 2b, c)** based upon gene expression (**Fig. 2a,**
124 **bottom right panel, Extended Data Fig. 2d, e, Extended Data Table 5)**. Consistent with our
125 flow cytometry results, considerable numbers of cells clustered into transcriptional groups
126 defining GC B cells (33.8%) and LNPCs (7%) in the lymph node.

127 To achieve reasonably comprehensive estimation of the vaccine-responding B cell clones
128 in the lymph nodes, we selected one representative BCR from either the GC B cell or the LNPC

129 transcriptional clusters for recombinant monoclonal antibody (mAb) generation from every
130 single-cell sequenced BCR clone containing at least 3 cells and present in one or both
131 compartments. A total of 2099 recombinant mAb from 8 participants were generated and
132 subsequently tested for binding to SARS-CoV-2 S by ELISA. We found that 1503 of the 2099
133 recombinant mAbs bound to SARS-CoV-2 S (**Fig. 2b, Extended Data Table 6**). In subsequent
134 analyses of these S-binding BCRs, we also included 15, 5 and 17 previously identified and
135 published S-binding mAbs generated from GC B cells at week 4 from, respectively, participants
136 07, 20, and 22⁶. Clonal relationships among S-specific BCRs were analyzed using heavy chains
137 from scRNA-seq libraries (**Extended Data Table 7**), bulk-seq libraries for sorted GC B cells and
138 LNPs (**Extended Data Fig. 2f, Extended Data Table 8**), as well as previously published bulk-
139 seq libraries of sorted PBs and GC B cells⁶ and magnetically enriched IgD^{low} activated or
140 memory B cells from PBMC²². B cell clones consisting of experimentally-validated S-binding B
141 cells accounted for 44.5% and 67.3%, respectively of the single-cell profiled GC B cells and
142 LNPs (**Fig. 2c, Extended Data Fig. 2g, Extended Data Table 5**). B cells that were clonally
143 related to S-binding B cells were also found in the PB cluster in blood (6.7%) and resting
144 memory B cell (RMB) cluster in lymph nodes (0.3%) (**Fig. 2c, Extended Data Fig. 2g,**
145 **Extended Data Table 5**). Thus, our results clearly demonstrate robust induction of SARS-CoV-
146 2 S-binding GC B cells and LNPs in the lymph node at the single-cell level following mRNA
147 vaccination.

148 We next analyzed the proportion of S-binding GC B cells clonally related to PBs in the
149 blood at week 4. The frequencies of PB-related S-binding GC B cells varied broadly among
150 participants and over time from 11.5% to 82.5% (**Fig. 3a**). BCR sequences of S-binding GC B
151 cells had significantly higher levels of somatic hypermutation (SHM) compared to clonally
152 related PB, and this difference increased over time (**Fig. 3b**). We observed a 2.5-fold increase
153 in SHM frequency among all S-binding GC B cells between weeks 4 and 15 (**Fig. 3c, Extended**
154 **Data Fig. 3a**). We also observed that the relative proportion of IgA-expressing S-binding GC B

155 cells increased in the lymph node over time (**Fig. 3d**). Collectively, these results suggest that
156 GC B cells continue to evolve for at least 15 weeks following mRNA vaccination.

157 Clonal analysis revealed a very high degree of overlap between S-binding GC and LNPC
158 compartments (**Fig. 3e**). Furthermore, SHM frequencies of both S-binding LNPCs and GC B
159 cells increased over time at a remarkably similar rate with small differences in contrast to those
160 between S-binding PB and GC B cells (**Fig. 3f**). Taken together, the near-complete clonal
161 overlap of S-binding LNPCs with GC B cells and their lockstep accumulation of SHM suggests
162 that LNPCs differentiate from GC B cells.

163 To determine whether the increase in SHM frequencies of S-specific GC B cells and LNPCs
164 over time is reflected in increased circulating anti-S antibody binding affinity, we measured the
165 avidity of plasma anti-S IgG at 5 and 29 weeks post-vaccination. In participants without a history
166 of SARS-CoV-2 infection, avidity of plasma anti-S IgG increased at 29 weeks compared to 5
167 weeks post-vaccination. Interestingly, participants with a history of SARS-CoV-2 infection had
168 comparable plasma anti-S IgG avidity at 5 and 29 weeks post-vaccination (**Fig. 4a**). As
169 expected, S-binding ASCs sampled at various time-points post-vaccination exhibited higher
170 frequencies of SHM over time, and BMPCs sampled 29 weeks post-vaccination exhibited the
171 highest degree of SHM (**Fig. 4b, Extended Data Fig. 4a**). To understand the evolutionary
172 relationships between S-specific ASC lineages, we analyzed S-binding clones using a
173 hierarchical phylogenetic model. The resulting phylogenetic trees revealed a close relationship
174 between BMPCs and LNPCs. The closely-related BMPCs and LNPCs shared common PB
175 ancestors (**Fig. 4c**). Together these results provide strong support for a model where SARS-
176 CoV-2 S-specific BMPCs are the products of affinity-matured GCs.

177 **Discussion**

178 Upon antigenic stimulation, naïve B cells and pre-existing MBCs can differentiate into short-
179 lived PBs that secrete relatively low-affinity antibodies or into GC B cells where they accumulate
180 SHM and undergo affinity-based selection, with surviving clones differentiating into MBCs and
181 long-lived BMPCs^{1-3,18,23,24}. This study evaluated whether the persistent GC response induced
182 by SARS-CoV-2 mRNA-based vaccines in humans⁶ results in generation of affinity-matured
183 antigen-specific BMPCs. The two-dose series of BNT162b2 induced a robust SARS-CoV-2 S-
184 binding GC B cell response that was detected in 10 of the 15 participants' draining lymph nodes
185 for at least six months after vaccination. The fruits of such persistent GC reactions were evident
186 in the form of circulating SARS-CoV-2 S⁺ MBCs in all participants and SARS-CoV-2 S⁺ BMPCs
187 detected in bone marrow specimens from 9 of the 11 participants examined six months after
188 vaccination. It is likely that S⁺ BMPCs in those two participants are present at very low
189 frequencies below the assay detection limit. Anti-S IgG plasma antibodies declined from an
190 initial peak 2 weeks after secondary vaccination but remained readily detectable for at least six
191 months after vaccination. Longitudinal tracking of over 1500 vaccine-induced B cell clones
192 revealed the gradual accumulation of SHM within the GC B cell compartment over time.
193 Additionally, it revealed the differentiation of those GC B cells into LNPCs within the lymph
194 node, and then into affinity matured BMPCs. The maturity of BMPCs was reflected in the
195 significantly enhanced avidity of circulating anti-S IgG antibodies detected at six months
196 compared to that measured at 2 weeks after the second vaccination. Our data corroborate
197 multiple reports demonstrating the maturation of B cell responses after SARS-CoV-2 mRNA
198 vaccination in humans^{14,15,17,25-27}. A potential limitation to our analyses of S-binding clones is that
199 our selection strategy may have excluded some low-abundance or low-affinity S-specific clones.
200 Nonetheless, we were able to account for 45% and 67% of all GC B cell and LNPC clones,
201 respectively identified by scRNA-seq.

202 This is the first study to provide direct evidence for the induction of antigen-specific
203 BMPCs by an mRNA-based vaccine in humans. Notably, none of the 11 participants from whom
204 post-vaccination bone marrow specimens were examined had a history of SARS-CoV-2
205 infection. BMPCs that recognized contemporary seasonal influenza virus vaccine antigens and
206 diphtheria/tetanus vaccine antigens were present at frequencies roughly 10- and 2-fold greater
207 than those against SARS-CoV-2 S, respectively. This is likely due to both the greater number of
208 antigenic targets contained in the former vaccines and the repeated exposures to influenza and
209 tetanus/diphtheria vaccine antigens our study participants likely experienced in comparison to
210 the initial exposure to the novel SARS-CoV-2 S antigen. There are some epitopes within the S
211 protein that are conserved between human seasonal coronaviruses and SARS-CoV-2^{28,29}.
212 Cross-reactive B cells targeting these epitopes participate in PB and GC B cell responses to
213 SARS-CoV-2 vaccination^{6,30}. It is unlikely, however, that a substantial proportion of the SARS-
214 CoV-2 S⁺ BMPCs we observed six months after immunization were part of a pre-existing pool of
215 BMPCs, as in a previous study we could not detect any SARS-CoV-2 S⁺ BMPCs in specimens
216 from 11 healthy individuals with no history of SARS-CoV-2 infection or vaccination²⁴. It is yet to
217 be determined, however, whether vaccine-induced SARS-CoV-2 S⁺ BMPCs we detected are
218 indeed long-lived.

219 An intriguing finding in our study is that the SARS-CoV-2 S⁺ BMPCs detected six months
220 after vaccination exhibited high SHM frequencies relative to the other B cell compartments.
221 These data corroborate similar observations made in the mouse model^{31,32}. The murine data led
222 to a proposal of a division of labor between memory B cells and long-lived BMPCs^{33,34}. Under
223 that framework, BMPCs secrete highly specific, high-affinity antibodies that provide the first
224 layer of protection against the invading pathogen while MBCs would only be engaged in the
225 event that a pathogen is not fully neutralized by BMPC-derived antibodies. Consistent with this
226 notion, multiple reports have recently documented the evolution of circulating MBCs induced by
227 SARS-CoV-2 mRNA vaccination in humans^{14,15,17,26}. These reports have shown that not only the

228 frequency of circulating S⁺ MBCs increased over time, but their ability to recognize S proteins
229 from emerging SARS-CoV-2 variants seems to have expanded as well^{25,26}. These data indicate
230 an important role for affinity maturation of responding B cell clones beyond increasing binding
231 affinity to the immunizing antigen. It is yet to be elucidated whether a similar division of labor is
232 established in human B cell responses to rapidly changing pathogens.

233 Our study raises a number of important questions that will need to be addressed in future
234 studies concerning the effects of an additional homologous or heterologous immunization on the
235 dynamics and products of ongoing germinal centres, particularly with respect to affinity and
236 specificity to original SARS-CoV-2 S and variant strains. Overall, our data demonstrate the
237 remarkable capacity of SARS-CoV-2 mRNA-based vaccines to induce robust and persistent GC
238 reactions that culminate in highly matured BMPC populations. mRNA-based vaccines represent
239 a relatively new platform that can be leveraged to rationally design new vaccines against
240 historically challenging targets, such as influenza, RSV, HIV, and malaria. Our studies provide a
241 foundation for how such vaccines can enhance the affinity and durability of elicited antibody
242 responses against these challenging human pathogens.

243 **Materials and Methods**

244 **Sample collection, preparation, and storage**

245 All studies were approved by the Institutional Review Board of Washington University in St
246 Louis. Written consent was obtained from all participants. Forty-three healthy volunteers were
247 enrolled, of whom 13 had a history of confirmed SARS-CoV-2 infection (Extended Data Table 1,
248 2). Fifteen out of 43 healthy participants provided FNAs of draining axillary lymph nodes. In 6
249 out of the 15 participants, a second draining lymph node was identified and sampled following
250 secondary immunization. One participant (15) received the boost vaccination in the contralateral
251 arm; draining lymph nodes were identified and sampled on both sides. Eleven out of 43 healthy
252 participants provided bone marrow aspirates. Forty-eight participants who had recovered from
253 mild SARS-CoV-2 infection but had not been vaccinated within 7 months of illness were
254 previously described²⁴ (Extended Data Table 1).

255 Peripheral blood samples were collected in EDTA tubes, and PBMCs were enriched by
256 density gradient centrifugation over Ficoll-Paque PLUS (Cytiva) or Lymphopure (BioLegend).
257 The residual red blood cells were lysed with ammonium chloride lysis buffer, and cells were
258 immediately used or cryopreserved in 10% dimethyl sulfoxide in fetal bovine serum (FBS).

259 Ultrasound-guided FNA of draining axillary lymph nodes was performed by a radiologist or
260 a qualified physician's assistant under the supervision of a radiologist. Scans were performed
261 with a commercially available ultrasound unit (Loqic E10, General Electric, Milwaukee, WI)
262 using an L2-9 linear array transducer with transmit frequencies of 7, 8, and 9 MHz or a L6-15
263 linear array transducer with transmit frequencies of 10, 12, and 15 MHz. Lymph node
264 dimensions and cortical thickness were measured, and the presence and degree of cortical
265 vascularity and location of the lymph node relative to the axillary vein were determined before
266 each FNA. For each FNA sample, six passes were made under continuous real-time ultrasound
267 guidance using 25-gauge needles, each of which was flushed with 3 ml of RPMI 1640
268 supplemented with 10% FBS and 100 U/ml penicillin-streptomycin, followed by three 1 ml

269 rinses. Red blood cells were lysed with ammonium chloride buffer (Lonza), washed with
270 washing buffer (phosphate-buffered saline supplemented with 2% FBS and 2 mM EDTA), and
271 immediately used or cryopreserved in 10% dimethyl sulfoxide in FBS. Participants reported no
272 adverse effects from phlebotomies or serial FNAs.

273 Bone marrow aspirates of approximately 30ml were collected in EDTA tubes from the iliac
274 crest. Bone marrow mononuclear cells (BMMCs) were enriched by density gradient
275 centrifugation over Ficoll-Paque PLUS, and then the remaining red blood cells were lysed with
276 ammonium chloride buffer (Lonza) and washed with washing buffer. BMPCs were enriched from
277 bone marrow mononuclear cells using EasySep Human CD138 Positive Selection Kit II
278 (StemCell Technologies) and immediately used for ELISpot or cryopreserved in 10% dimethyl
279 sulfoxide in FBS.

280

281 **Antigens**

282 Recombinant soluble spike (S) protein derived from SARS-CoV-2 was expressed as
283 previously described³⁵. In brief, a mammalian cell codon-optimized nucleotide sequences
284 coding for the soluble version of S (GenBank: MN908947.3, amino acids 1-1,213) including a C-
285 terminal thrombin cleavage site, T4 fold trimerization domain and hexahistidine tag was cloned
286 into the mammalian expression vector pCAGGS. The S protein sequence was modified to
287 remove the polybasic cleavage site (RRAR to A) and two stabilizing mutations were introduced
288 (K986P and V987P, wild-type numbering). Recombinant proteins were produced in Expi293F
289 cells (Thermo Fisher Scientific) by transfection with purified plasmid using the ExpiFectamine
290 293 Transfection Kit (Thermo Fisher Scientific). Supernatants from transfected cells were
291 collected 3 days after transfection, and recombinant proteins were purified using Ni-NTA
292 agarose (Thermo Fisher Scientific), then buffer-exchanged into PBS and concentrated using
293 Amicon Ultra centrifugal filters (MilliporeSigma). For flow cytometry staining, recombinant S was
294 labeled with Alexa Fluor 7647-NHS ester or biotinylated using the EZ-Link Micro NHS-PEG4-

295 Biotinylation Kit (Thermo Fisher Scientific); excess Alexa Fluor 647 and biotin were removed
296 using 7-kDa Zeba desalting columns (Thermo Fisher Scientific).

297

298 **Flow cytometry and cell sorting**

299 Staining for flow cytometry analysis and sorting was performed using freshly isolated or
300 cryo-preserved PBMCs or FNAs. For FNA staining, cells were incubated for 30 min on ice with
301 biotinylated and Alexa Fluor 647-conjugated recombinant soluble S and PD-1-BB515 (EH12.1,
302 BD Horizon, 1:100) in 2% FBS and 2 mM EDTA in PBS (P2), washed twice, then stained for 30
303 min on ice with IgG-BV480 (goat polyclonal, Jackson ImmunoResearch, 1:100), IgA-FITC
304 (M24A, Millipore, 1:500), CD45-A532 (HI30, Thermo, 1:50), CD38-BB700 (HIT2, BD Horizon,
305 1:500), CD20-Pacific Blue (2H7, 1:400), CD27-BV510 (O323, 1:50), CD8-BV570 (RPA-T8,
306 1:200), IgM-BV605 (MHM-88, 1:100), HLA-DR-BV650 (L243, 1:100), CD19-BV750 (HIB19,
307 1:100), CXCR5-PE-Dazzle 594 (J252D4, 1:50), IgD-PE-Cy5 (IA6-2, 1:200), CD14-PerCP
308 (HCD14, 1:50), CD71-PE-Cy7 (CY1G4, 1:400), CD4-Spark685 (SK3, 1:200), streptavidin-APC-
309 Fire750, CD3-APC-Fire810 (SK7, 1:50) and Zombie NIR (all BioLegend) diluted in Brilliant
310 Staining buffer (BD Horizon). Cells were washed twice with P2, fixed for 1 h at 25 °C using the
311 True Nuclear fixation kit (BioLegend), washed twice with True Nuclear Permeabilization/Wash
312 buffer, stained with FOXP3-BV421 (206D, BioLegend, 1:15), Ki-67-BV711 (Ki-67, BioLegend,
313 1:200), T-bet-BV785 (4B10, BioLegend, 1:400), BCL6-PE (K112-91, BD Pharmingen, 1:25), and
314 BLIMP1-A700 (646702, R&D, 1:50) for 1 h at 25 °C, washed twice with True Nuclear
315 Permeabilization/Wash buffer and resuspended in P2 for acquisition. For memory B cell
316 staining, PBMC were incubated for 30 min on ice with biotinylated and Alexa Fluor 647-
317 conjugated recombinant soluble S in P2, washed twice, then stained for 30 min on ice with IgG-
318 BV480 (goat polyclonal, Jackson ImmunoResearch, 1:100), IgD-Super Bright 702 (IA6-2,
319 Thermo, 1:50), IgA-FITC (M24A, Millipore, 1:500), CD45-A532 (HI30, Thermo, 1:50), CD38-
320 BB700 (HIT2, BD Horizon, 1:500), CD24-BV421 (ML5, 1:100), CD20-Pacific Blue (2H7, 1:400),

321 CD27-BV510 (O323, 1:50), CD8-BV570 (RPA-T8, 1:200), IgM-BV605 (MHM-88, 1:100), CD19-
322 BV750 (HIB19, 1:100), FcRL5-PE (509f6, 1:100), CXCR5-PE-Dazzle 594 (J252D4, 1:50),
323 CD14-PerCP (HCD14, 1:50), CD71-PE-Cy7 (CY1G4, 1:400), CD4-Spark685 (SK3, 1:200),
324 streptavidin-APC-Fire750, CD3-APC-Fire810 (SK7, 1:50) and Zombie NIR (all BioLegend)
325 diluted in Brilliant Staining buffer (BD Horizon). Cells were washed twice with P2 and
326 resuspended in P2 for acquisition. All samples were acquired on an Aurora using SpectroFlo
327 v.2.2 (Cytex). Flow cytometry data were analyzed using FlowJo v.10 (BD Biosciences).

328 For sorting PBs from peripheral blood, B cells were enriched from PBMC by first using
329 EasySep Human Pan-B cell Enrichment Kit (StemCell Technologies), and then stained with
330 CD20-PB (2H7, 1:400), CD3-FITC (HIT3a, 1:200), IgD-PerCP-Cy5.5 (IA6-2, 1:200), CD71-PE
331 (CY1G4, 1:400), CD38-PE-Cy7 (HIT2, 1:200), CD19-APC (HIB19, 1:200) and Zombie Aqua (all
332 BioLegend). For sorting GC B cells and LNPs from the lymph node, single-cell suspensions
333 were stained for 30min on ice with PD-1-BB515 (EH12.1, BD Horizon, 1:100), CD20-Pacific
334 Blue (2H7, 1:100), IgD-PerCP-Cy5.5 (IA6-2, 1:200), CD19-PE (HIB19, 1:200), CXCR5-PE-
335 Dazzle 594 (J252D4, 1:50), CD38-PE-Cy7 (HIT2, 1:200), CD4-Alexa-Fluor-700 (SK3, 1:400),
336 CD71-APC (CY1G4, 1:100), and Zombie Aqua (all BioLegend). Cells were washed twice, and
337 single PBs (live singlet CD19⁺ CD4⁻ IgD^{low} CD38⁺ CD20⁻ CD71⁺), GC B cells (live singlet CD19⁺
338 CD4⁻ IgD^{low} CD71⁺ CD38^{int} CD20⁺ CXCR5⁺), LNPs (live singlet CD19⁺ CD4⁻ IgD^{low} CD38⁺ CD20⁻
339 CD71⁺) were sorted using a FACSAria II.

340

341 **ELISA**

342 Assays were performed in MaxiSorp 96-well plates (Thermo Fisher) coated with 100 ul of
343 recombinant SARS-CoV-2 S, Donkey anti-human IgG (H+L) antibody (Jackson
344 ImmunoResearch, 709-005-149) or bovine serum albumin diluted to 1ug/ml in PBS, and plates
345 were incubated at 4 °C overnight. Plates then were blocked with 10% FBS and 0.05% Tween 20
346 in PBS. Plasma or purified monoclonal antibodies were serially diluted in blocking buffer and

347 added to the plates. Monoclonal antibodies and plasma samples were tested at 10 ug/ml and
348 1:30 starting dilution, respectively, followed by 7 additional 3-fold serial dilutions. Plates were
349 incubated for 90 min at room temperature and then washed 3 times with 0.05% Tween 20 in
350 PBS. Secondary antibodies were diluted in blocking buffer before adding to wells and incubating
351 for 60 min at room temperature. HRP-conjugated goat anti-human IgG (H+L) antibody (Jackson
352 ImmunoResearch, 109-035-088, 1:2500) was used to detect monoclonal antibodies. HRP-
353 conjugated goat anti-Human IgG Fc γ fragment (Jackson ImmunoResearch, 109-035-190,
354 1:1500), HRP-conjugated goat anti-human serum IgA α chain (Jackson ImmunoResearch, 109-
355 035-011, 1:2500), and HRP-conjugated goat anti-human IgM (Caltag, H15007, 1:4000) were
356 used to detect plasma antibodies. Plates were washed 3 times with PBST and 3 times with PBS
357 before the addition of *o*-phenylenediamine dihydrochloride peroxidase substrate
358 (MilliporeSigma). Reactions were stopped by the addition of 1M hydrochloric acid. Optical
359 density measurements were taken at 490 nm. The threshold of positivity for recombinant mAbs
360 was set as two times the optical density of background binding to BSA at the highest
361 concentration of each mAb. The area under the curve for each monoclonal antibody and half-
362 maximal binding dilution for each plasma sample were calculated using GraphPad Prism v.9.
363 Plasma antibody avidity was measured as previously described³⁶. Briefly, plasma dilutions that
364 would give an optical density reading of 2.5 were calculated from the serial dilution ELISA. S-
365 coated plates were incubated with this plasma dilution as above and then washed one time for 5
366 minutes with either PBS or 8M urea in PBS, followed by 3 washes with PBST and developed as
367 above. The avidity index was calculated for each sample as the optical density ratio of the urea-
368 washed to PBS-washed wells.

369

370 **ELISpot**

371 ELISpot plates were coated overnight at 4 °C with Flucelvax Quadrivalent 2019/2020
372 seasonal influenza virus vaccine (Seqirus), tetanus/diphtheria vaccine (Grifols), SARS-CoV-2 S,

373 anti-human Ig (Cellular Technology Limited) and bovine serum albumin. A direct ex vivo
374 ELISpot assay was performed to determine the number of total, vaccine-binding or recombinant
375 S-binding IgG- and IgA-secreting cells present in PBMCs or enriched BMPCs using Human
376 IgA/IgG double-color ELISpot kits (Cellular Technology Limited) according to the manufacturer's
377 protocol. ELISpot plates were analyzed using an ELISpot analyzer (Cellular Technology
378 Limited).

379

380 **Single-cell RNA-seq library preparation and sequencing**

381 Sorted PBs and whole FNA from each time point were processed using the following 10x
382 Genomics kits: Chromium Next GEM Single Cell 5' Kit v2 (PN-1000263); Chromium Next GEM
383 Chip K Single Cell Kit (PN-1000286); BCR Amplification Kit (PN-1000253); Dual Index Kit TT
384 Set A (PN-1000215). Chromium Single Cell 5' Gene Expression Dual Index libraries and
385 Chromium Single Cell V(D)J Dual Index libraries were prepared according to manufacturer's
386 instructions without modifications. Both gene expression and V(D)J libraries were sequenced on
387 a Novaseq S4 (Illumina), targeting a median sequencing depth of 50,000 and 5,000 read pairs
388 per cell, respectively.

389

390 **Bulk B cell receptor sequencing**

391 RNA was prepared from sorted GC B cells and LNPs from FNA or enriched BMPCs from
392 bone marrow using the RNeasy Plus Micro kit (Qiagen). Libraries were prepared using the
393 NEBNext Immune Sequencing Kit for Human (New England Biolabs) according to the
394 manufacturer's instructions without modifications. High-throughput 2×300-bp paired-end
395 sequencing was performed on the Illumina MiSeq platform with a 30% PhiX spike-in according
396 to manufacturer's recommendations, except for performing 325 cycles for read 1 and 275 cycles
397 for read 2.

398

399 **Preprocessing of bulk sequencing BCR reads**

400 Preprocessing of demultiplexed pair-end reads were performed using pRESTO v.0.6.2³⁷ as
401 previously described⁶, with the exception that sequencing errors were corrected using the UMIs
402 as they were without additional clustering (Extended Data Table 8). Previously preprocessed
403 unique consensus sequences from reported samples⁶ were included as they were. Previously
404 preprocessed unique consensus sequences from reported samples²² were subset to those with
405 at least two contributing reads and included.

406

407 **Preprocessing of 10× Genomics single-cell BCR reads**

408 Demultiplexed pair-end FASTQ reads were preprocessed using the ‘cellranger vdj’
409 command from 10× Genomics’ Cell Ranger v.6.0.1 for alignment against the GRCh38 human
410 reference v.5.0.0 (‘refdata-cellranger-vdj-GRCh38-alts-ensembl-5.0.0’). The resultant
411 ‘filtered_contig.fasta’ files were used as preprocessed single-cell BCR reads (Extended Data
412 Table 7).

413

414 **V(D)J gene annotation and genotyping**

415 Initial germline V(D)J gene annotation was performed on the preprocessed BCRs using
416 IgBLAST v.1.17.1³⁸ with IMGT/GENE-DB release 202113-2³⁹. IgBLAST output was parsed
417 using Change-O v.1.0.2⁴⁰. For the single-cell BCRs, isotype annotation was pulled from the
418 ‘c_call’ column in the ‘filtered_contig_annotations.csv’ files outputted by Cell Ranger.

419 For both bulk and single-cell BCRs, sequence-level quality control was performed, requiring
420 each sequence to have non-empty V and J gene annotations; exhibit chain consistency in all
421 annotations; bear fewer than 10 non-informative (non-A/T/G/C, such as N or -) positions; and
422 carry a non-empty CDR3 with no N and a nucleotide length that is a multiple of 3. For single-cell
423 BCRs, cell-level quality control was also performed, requiring each cell to have either exactly
424 one heavy chain and at least one light chain, or at least one heavy chain and exactly one light

425 chain. Within a cell, for the chain type with more than one sequence, the most abundant
426 sequence in terms of UMI count (when tied, the sequence that appeared earlier in the file) was
427 kept. Ultimately, exactly one heavy chain and one light chain per cell were kept. Additionally,
428 quality control against cross-sample contamination was performed by examining the extent, if
429 any, of pairwise overlapping between samples in terms of BCRs with both identical UMIs and
430 identical non-UMI nucleotide sequences.

431 Individualized genotypes were inferred based on sequences that passed all quality control
432 using TIgGER v.1.0.0⁴¹ and used to finalize V(D)J annotations. Sequences annotated as non-
433 productively rearranged by IgBLAST were removed from further analysis.

434

435 **Clonal lineage inference**

436 B cell clonal lineages were inferred on a by-individual basis based on productively
437 rearranged sequences using hierarchical clustering with single linkage⁴². When combining both
438 bulk and single-cell BCRs, heavy chain-based clonal inference was performed⁴³. First, heavy
439 chain sequences were partitioned based on common V and J gene annotations and CDR3
440 lengths. Within each partition, heavy chain sequences with CDR3s that were within 0.15
441 normalized Hamming distance from each other were clustered as clones. When using only
442 single-cell BCRs, clonal inference was performed based on paired heavy and light chains. First,
443 paired heavy and light chains were partitioned based on common V and J gene annotations and
444 CDR3 lengths. Within each partition, pairs whose heavy chain CDR3s were within 0.15
445 normalized Hamming distance from each other were clustered as clones.

446 Following clonal inference, full-length clonal consensus germline sequences were
447 reconstructed for each clone with the D-segment (for heavy chains) and the N/P regions
448 masked with Ns, resolving any ambiguous gene assignments by majority rule. Within each
449 clone, duplicate IMGT-aligned V(D) J sequences from bulk sequencing were collapsed except
450 for duplicates derived from different time points, tissues, B cell compartments, or isotypes.

451

452 **BCR analysis**

453 BCR analysis was performed in R v4.1.0 with visualization performed using base R,
454 ggplot2 v3.3.5⁴⁴, and GraphPad Prism v9.

455 For the B cell compartment label, gene expression-based cluster annotation was used for
456 single-cell BCRs; FACS-based sorting was used in general for bulk BCRs, except that PB sorts
457 from lymph nodes were labelled LNPCs, d35 IgDlo sorts from blood were labelled activated, and
458 d60 IgDlo sorts from blood were labelled memory. For the time point label, one blood PB
459 sample that pooled collections on both d28 and d35 was treated as d28.

460 A heavy chain-based B cell clone was considered a S⁺ clone if the clone contained any
461 sequence corresponding to a recombinant mAb that was synthesized based on the single-cell
462 BCRs and that tested positive for S-binding.

463 Clonal overlap between B cell compartments was visualized using circlize v.0.4.13⁴⁵.

464 Somatic hypermutation (SHM) frequency was calculated for each heavy chain sequence by
465 counting the number of nucleotide mismatches from the germline sequence in the variable
466 segment leading up to the CDR3, while excluding the first 18 positions that could be error-prone
467 due to the primers used for generating the mAb sequences. Calculation was performed using
468 the calcObservedMutations function from SHazaM v.1.0.2⁴⁰.

469 Phylogenetic trees for S⁺ clones containing BMPCs were constructed on a by-participant
470 basis using IgPhyML v1.1.3⁴⁶ with the HLP19 model⁴⁷. Only heavy chain sequences from the
471 PB, GC, LNPC, RMB, and BMPC compartments were considered. For clones with >100
472 sequences, subsampling was applied with probabilities proportional to the proportions of
473 sequences from different compartments, in addition to keeping all sequences corresponding to
474 synthesized mAbs and all BMPC sequences. Only subsampled sequences from the PB, LNPC,
475 and BMPC compartments were used for eventual tree-building. Trees were visualized using
476 ggtree v3.0.4⁴⁸.

477

478 **Human housekeeping genes**

479 A list of human housekeeping genes was compiled from the 20 most stably expressed
480 genes across 52 tissues and cell types in the Housekeeping and Reference Transcript (HRT)
481 Atlas v1.0⁴⁹; 11 highly uniform and strongly expressed genes reported⁵⁰; and some of the most
482 commonly used housekeeping genes⁵¹. The final list includes 34 genes: *ACTB*, *TLE5* (*AES*),
483 *AP2M1*, *BSG*, *C1orf43*, *CD59*, *CHMP2A*, *CSNK2B*, *EDF1*, *EEF2*, *EMC7*, *GABARAP*, *GAPDH*,
484 *GPI*, *GUSB*, *HNRNPA2B1*, *HPRT1*, *HSP90AB1*, *MLF2*, *MRFAP1*, *PCBP1*, *PFDN5*, *PSAP*,
485 *PSMB2*, *PSMB4*, *RAB11B*, *RAB1B*, *RAB7A*, *REEP5*, *RHOA*, *SNRPD3*, *UBC*, *VCP*, and *VPS29*.

486

487 **Processing of 10× Genomics single-cell 5' gene expression data**

488 Demultiplexed pair-end FASTQ reads were first preprocessed on a by-sample basis using
489 the 'cellranger count' command from 10× Genomics' Cell Ranger v.6.0.1 for alignment against
490 the GRCh38 human reference v.2020-A ('refdata-gex-GRCh38-2020-A'). To avoid a batch
491 effect introduced by sequencing depth, the 'cellranger aggr' command was used to subsample
492 from each sample so that all samples had the same effective sequencing depth, which was
493 measured in terms of the number of reads confidently mapped to the transcriptome or assigned
494 to the feature IDs per cell. Gene annotation on human reference chromosomes and scaffolds in
495 Gene Transfer Format ('gencode.v32.primary_assembly.annotation.gtf') was downloaded
496 (2021-06-02) from GENCODE v32⁵², from which a biotype ('gene_type') was extracted for each
497 feature. Quality control was performed as follows on the aggregate gene expression matrix
498 consisting of 360,803 cells and 36,601 features using SCANPY v1.7.2⁵³ and Python v3.8.8. (1)
499 To remove presumably lysed cells, cells with mitochondrial content greater than 12.5% of all
500 transcripts were removed. (2) To remove likely doublets, cells with more than 8,000 features or
501 80,000 total UMIs were removed. (3) To remove cells with no detectable expression of common
502 endogenous genes, cells with no transcript for any of the 34 housekeeping genes were

503 removed. (4) The feature matrix was subset, based on their biotypes, to protein-coding,
504 immunoglobulin, and T cell receptor genes that were expressed in at least 0.1% of the cells in
505 any sample. The resultant feature matrix contained 15,744 genes. (5) Cells with detectable
506 expression of fewer than 200 genes were removed. After quality control, there were a total of
507 318,156 cells from 47 single-cell samples (Extended Data Table 7).

508

509 **Single-cell gene expression analysis**

510 Single-cell gene expression analysis was performed in SCANPY v1.7.2⁵³. UMI counts
511 measuring gene expression were log-normalized. The top 2,500 highly variable genes (HVGs)
512 were identified using the 'scanpy.pp.highly_variable_genes' function with the 'seurat_v3'
513 method, from which immunoglobulin and T cell receptor genes were removed. The data were
514 scaled and centred, and principal component analysis (PCA) was performed based on HVG
515 expression. PCA-guided neighborhood graphs embedded in Uniform Manifold Approximation
516 and Projection (UMAP) were generated using the top 20 principal components via the
517 'scanpy.pp.neighbors' and 'scanpy.tl.umap' functions.

518 Overall clusters (Extended Data Table 5, top) were identified using Leiden graph-clustering
519 via the 'scanpy.tl.leiden' function with resolution 0.15 (Extended Data Fig. 2b). UMAPs were
520 faceted by batch, by participant, and by participant followed by sample; and inspected for
521 convergence across batches, participants, and samples within participants, to assess whether
522 there was a need for integration (Extended Data Fig. 2b). Cluster identities were assigned by
523 examining the expression of a set of marker genes for different cell types (Extended Data Fig.
524 2c): *MS4A1*, *CD19* and *CD79A* for B cells; *CD3D*, *CD3E*, *CD3G*, *IL7R* and *CD4* or *CD8A* for
525 $CD4^+$ or $CD8^+$ T cells, respectively; *GZMB*, *GNLY*, *NKG7* and *NCAM1* for natural killer (NK)
526 cells; *CD14*, *LYZ*, *CST3* and *MS4A7* for monocytes; *IL3RA* and *CLEC4C* for plasmacytoid
527 dendritic cells (pDCs); and *FDCSP*, *CXCL14*²⁰ and *FCAMR*²¹ for follicular dendritic cells (FDCs).
528 To remove potential contamination by platelets, 60 cells with a log-normalized expression value

529 of >2.5 for PPBP were removed. All 349 cells from the FDC cluster were confirmed to have
530 originated from FNA samples instead of blood.

531 Cells from the overall B cell cluster (Extended Data Table 5, bottom) were further clustered
532 to identify B cell subsets using Leiden graph-clustering via the 'scanpy.tl.leiden' function with
533 resolution 0.2 (Extended Data Fig. 2d). Cluster identities were assigned by examining the
534 expression of a set of marker genes for different B cell subsets (Extended Data Fig. 2e) along
535 with the availability of BCRs. The following marker genes were examined: *BCL6*, *RGS13*,
536 *MEF2B*, *STMN1*, *ELL3* and *SERPINA9* for GC B cells; *XBP1*, *IRF4*, *SEC11C*, *FKBP11*, *JCHAIN*
537 and *PRDM1* for PBs and LNPs; *TCL1A*, *IL4R*, *CCR7*, *IGHM*, and *IGHD* for naive B cells; and
538 *TNFRSF13B*, *CD27* and *CD24* for RMB cells. Although 2 groups clustered with B cells during
539 overall clustering, they were labelled 'B and T' as their cells tended to have both BCRs and high
540 expression levels of *CD2* and *CD3E*; and were subsequently excluded from the final B cell
541 clustering. 19 cells that were found in the GC cluster but came from blood samples were
542 labelled 'PB-like'. 198 cells that were found in the PB cluster but came from FNA samples were
543 re-assigned as LNPs. 37 cells that were found in the LNP cluster but came from blood
544 samples were re-assigned as PBs. Heavy chain SHM frequency and isotype usage of the B cell
545 subsets were assessed for consistency with expected values to further confirm their assigned
546 identities.

547

548 **Selection and curation of single-cell BCRs from GC/LNP for expression**

549 Single-cell gene expression analysis was performed using lymph node samples on a by-
550 participant basis. Clonal inference was performed based on paired heavy and light chains from
551 the same samples. From every clone with a clone size of >3 cells that contained cells from the
552 GC and/or LNP clusters, one GC or LNP cell was selected. For selection, where a clone
553 spanned both the GC and LNP compartments, and/or multiple time points, a compartment and
554 a timepoint were first randomly selected. Within that clone, the cell with the highest heavy chain

555 UMI count was then selected, breaking ties based on *IGHV* SHM frequency. In all selected
556 cells, native pairing was preserved.

557 BCRs from the selected cells were curated prior to synthesis. First, artificial gaps introduced
558 under the IMGT unique numbering system⁵⁴ were removed from the IMGT-aligned observed
559 V(D)J sequences. IMGT gaps were identified as positions containing in-frame triplet dots ('...')
560 in the IMGT-aligned germline sequences. Second, any non-informative (non-A/T/G/C, such as N
561 or -) positions in the observed sequences, with the exception of potential in-frame indels, were
562 patched by the nucleotides at their corresponding germline positions. Third, if applicable, the 3'
563 end of the observed sequences were trimmed so that the total nucleotide length would be a
564 multiple of 3. Finally, potential in-frame indels were manually reviewed. For a given potential in-
565 frame indel from a selected cell, its presence or lack thereof in the unselected cells from the
566 same clone was considered. Barring strong indications that an in-frame indel was due to
567 sequencing error rather than the incapability of the IMGT unique numbering system of capturing
568 it, the in-frame indels were generally included in the final curated sequences.

569

570 **Transfection for recombinant mAbs production**

571 Selected pairs of heavy and light chain sequences were synthesized by GenScript and
572 sequentially cloned into IgG1 and Igκ/λ expression vectors, respectively. Heavy and light chain
573 plasmids were co-transfected into Expi293F cells (Thermo Fisher Scientific) for recombinant
574 mAb production, followed by purification with protein A agarose resin (GoldBio). Expi293F cells
575 were cultured in Expi293 Expression Medium (Gibco) according to the manufacturer's protocol.

576 **Acknowledgements**

577 We thank the generous participation of the donors for providing specimens.

578 The Ellebody laboratory was supported by National Institute of Allergy and Infectious
579 Diseases (NIAID) grants U01AI141990 and 1U01AI150747, NIAID Centers of Excellence for
580 Influenza Research and Surveillance contract HHSN272201400006C and
581 HHSN272201400008C, and NIAID Collaborative Influenza Vaccine Innovation Centers contract
582 75N93019C00051. W.K. was supported by the Basic Science Research Program through the
583 National Research Foundation of Korea (NRF) funded by the Ministry of Education
584 (2021R1A6A3A03041509). J.S.T. was supported by NIAID 5T32CA009547. This study utilized
585 samples obtained from the Washington University School of Medicine's COVID-19 biorepository
586 supported by the NIH/National Center for Advancing Translational Sciences, grant number UL1
587 TR002345. The WU353 and WU368 studies were reviewed and approved by the Washington
588 University Institutional Review Board (approval no. 202003186 and 202012081, respectively).

589

590 **Author Contributions**

591 A.H.E. conceived and designed the study. A.H., M.K.K., L.P., P.A.M., I.P., J.A.O., and
592 R.M.P. wrote and maintained the IRB protocol, recruited, and phlebotomized participants,
593 coordinated sample collection, and analyzed clinical data. W.K., E.K, and J.S.T. processed
594 specimens. W.K., E.K., W.B.A., and J.S.T. performed ELISA and ELISpot. W.K., S.C.H,
595 A.J.Schmitz, T.L., M.T., and W.B.A. generated and characterized monoclonal antibodies.
596 W.K.and A.J.Sturtz prepared libraries for scRNA sequencing. A.J.Schmitz performed RNA
597 extractions and library preparation for BCR bulk sequencing. J.Q.Z. analyzed scRNA
598 sequencing and BCR bulk sequencing data. A.J.Schmitz expressed SARS-CoV-2 S and RBD
599 proteins. J.S.T. sorted cells and collected and analysed the flow cytometry data. T.S. and
600 W.D.M performed FNA. W.D.M. and S.A.T. supervised lymph node evaluation prior to FNA and
601 specimen collection and evaluated lymph node ultrasound data. W.K., J.Q.Z, J.S.T., and A.H.E.

602 analyzed the data. A.H.E. supervised experiments and obtained funding. W.K., P.A.M, J.S.T.,
603 and A.H.E. composed the manuscript. All authors reviewed the manuscript.

604

605 **Competing Interests**

606 The Ellebedy laboratory received funding under sponsored research agreements that are
607 unrelated to the data presented in the current study from Emergent BioSolutions and from
608 AbbVie. A.H.E. is a consultant for Mubadala Investment Company and the founder of
609 ImmuneBio Consulting. J.S.T. is a consultant for Gerson Lehrman Group. J.S.T., A.J.Schmitz.,
610 and A.H.E. are recipients of a licensing agreement with Abbvie that is unrelated to the data
611 presented in the current study. A patent application related to this work has been filed by
612 Washington University School of Medicine. The content of this manuscript is solely the
613 responsibility of the authors and does not necessarily represent the official view of NIAID or
614 NIH.

615

616 **Data availability statement**

617 New raw sequencing data, processed transcriptomics data, and processed BCR data will
618 be uploaded to Sequence Read Archive, GEO, and Zenodo respectively before final publication.
619 Previously reported bulk-sequenced BCR data used in this study were deposited under
620 PRJNA731610 and PRJNA741267 on SRA, and at <https://doi.org/10.5281/zenodo.5042252> and
621 <https://doi.org/10.5281/zenodo.5040099> on Zenodo.

- 623 1 Victora, G. D. & Nussenzweig, M. C. Germinal Centers. *Annual Review of Immunology*
624 **30**, 429-457, doi:10.1146/annurev-immunol-020711-075032 (2012).
- 625 2 Cyster, J. G. & Allen, C. D. C. B Cell Responses: Cell Interaction Dynamics and
626 Decisions. *Cell* **177**, 524-540, doi:10.1016/j.cell.2019.03.016 (2019).
- 627 3 Radbruch, A. *et al.* Competence and competition: the challenge of becoming a long-lived
628 plasma cell. *Nature Reviews Immunology* **6**, 741-750, doi:10.1038/nri1886 (2006).
- 629 4 Slifka, M. K., Antia, R., Whitmire, J. K. & Ahmed, R. Humoral Immunity Due to Long-
630 Lived Plasma Cells. *Immunity* **8**, 363-372, doi:10.1016/s1074-7613(00)80541-5 (1998).
- 631 5 Tarlinton, D., Radbruch, A., Hiepe, F. & Dörner, T. Plasma cell differentiation and
632 survival. *Current Opinion in Immunology* **20**, 162-169, doi:10.1016/j.coi.2008.03.016
633 (2008).
- 634 6 Turner, J. S. *et al.* SARS-CoV-2 mRNA vaccines induce persistent human germinal
635 centre responses. *Nature* **596**, 109-113, doi:10.1038/s41586-021-03738-2 (2021).
- 636 7 Mudd, P. A. *et al.* SARS-CoV-2 mRNA vaccination elicits robust and persistent T
637 follicular helper cell response in humans (Cold Spring Harbor Laboratory, 2021).
- 638 8 Lederer, K. *et al.* Germinal center responses to SARS-CoV-2 mRNA vaccines in healthy
639 and immunocompromised individuals (Cold Spring Harbor Laboratory, 2021).
- 640 9 Baden, L. R. *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New*
641 *England Journal of Medicine* **384**, 403-416, doi:10.1056/nejmoa2035389 (2021).
- 642 10 Polack, F. P. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New*
643 *England Journal of Medicine* **383**, 2603-2615, doi:10.1056/nejmoa2034577 (2020).
- 644 11 Thomas, S. J. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine
645 through 6 Months. *New England Journal of Medicine*, doi:10.1056/nejmoa2110345
646 (2021).
- 647 12 Levin, E. G. *et al.* Waning Immune Humoral Response to BNT162b2 Covid-19 Vaccine
648 over 6 Months. *New England Journal of Medicine*, doi:10.1056/nejmoa2114583 (2021).
- 649 13 Suthar, M. S. *et al.* Durability of immune responses to the BNT162b2 mRNA vaccine
650 (Cold Spring Harbor Laboratory, 2021).
- 651 14 Goel, R. R. *et al.* mRNA vaccines induce durable immune memory to SARS-CoV-2 and
652 variants of concern. *Science*, doi:10.1126/science.abm0829 (2021).
- 653 15 Cho, A. *et al.* Anti-SARS-CoV-2 receptor binding domain antibody evolution after mRNA
654 vaccination. *Nature*, doi:10.1038/s41586-021-04060-7 (2021).
- 655 16 Chen, Y. *et al.* Differential antibody dynamics to SARS-CoV-2 infection and vaccination
656 (Cold Spring Harbor Laboratory, 2021).
- 657 17 Pape, K. A. *et al.* High-affinity memory B cells induced by SARS-CoV-2 infection
658 produce more plasmablasts and atypical memory B cells than those primed by mRNA
659 vaccines. *Cell Reports* **37**, 109823, doi:10.1016/j.celrep.2021.109823 (2021).
- 660 18 Turner, J. S. *et al.* Human germinal centres engage memory and naive B cells after
661 influenza vaccination. *Nature* **586**, 127-132, doi:10.1038/s41586-020-2711-0 (2020).
- 662 19 King, H. W. *et al.* Single-cell analysis of human B cell maturation predicts how antibody
663 class switching shapes selection dynamics. *Science Immunology* **6**, eabe6291,
664 doi:doi:10.1126/sciimmunol.abe6291 (2021).
- 665 20 Haebe, S. *et al.* Single-cell analysis can define distinct evolution of tumor sites in
666 follicular lymphoma. *Blood* **137**, 2869-2880, doi:10.1182/blood.2020009855 (2021).
- 667 21 Mourcin, F. *et al.* Follicular lymphoma triggers phenotypic and functional remodeling of
668 the human lymphoid stromal cell landscape. *Immunity* **54**, 1901,
669 doi:10.1016/j.immuni.2021.07.018 (2021).

- 670 22 Schmitz, A. J. *et al.* A vaccine-induced public antibody protects against SARS-CoV-2
671 and emerging variants. *Immunity* **54**, 2159-2166.e2156,
672 doi:10.1016/j.immuni.2021.08.013 (2021).
- 673 23 Elsner, R. A. & Shlomchik, M. J. Germinal Center and Extrafollicular B Cell Responses
674 in Vaccination, Immunity, and Autoimmunity. *Immunity* **53**, 1136-1150,
675 doi:10.1016/j.immuni.2020.11.006 (2020).
- 676 24 Turner, J. S. *et al.* SARS-CoV-2 infection induces long-lived bone marrow plasma cells
677 in humans. *Nature* **595**, 421-425, doi:10.1038/s41586-021-03647-4 (2021).
- 678 25 Tong, P. *et al.* Memory B cell repertoire for recognition of evolving SARS-CoV-2 spike.
679 *Cell* **184**, 4969-4980.e4915, doi:10.1016/j.cell.2021.07.025 (2021).
- 680 26 Sokal, A. *et al.* mRNA vaccination of naive and COVID-19-recovered individuals elicits
681 potent memory B cells that recognize SARS-CoV-2 variants. *Immunity*,
682 doi:10.1016/j.immuni.2021.09.011 (2021).
- 683 27 Lucas, C. *et al.* Impact of circulating SARS-CoV-2 variants on mRNA vaccine-induced
684 immunity. *Nature*, doi:10.1038/s41586-021-04085-y (2021).
- 685 28 Shrock, E. *et al.* Viral epitope profiling of COVID-19 patients reveals cross-reactivity and
686 correlates of severity. *Science* **370**, eabd4250, doi:doi:10.1126/science.abd4250 (2020).
- 687 29 Anderson, E. M. *et al.* Seasonal human coronavirus antibodies are boosted upon SARS-
688 CoV-2 infection but not associated with protection. *Cell* **184**, 1858-1864.e1810,
689 doi:10.1016/j.cell.2021.02.010 (2021).
- 690 30 Amanat, F. *et al.* SARS-CoV-2 mRNA vaccination induces functionally diverse
691 antibodies to NTD, RBD, and S2. *Cell* **184**, 3936-3948.e3910,
692 doi:10.1016/j.cell.2021.06.005 (2021).
- 693 31 Smith, K. G. C. The extent of affinity maturation differs between the memory and
694 antibody-forming cell compartments in the primary immune response. *The EMBO*
695 *Journal* **16**, 2996-3006, doi:10.1093/emboj/16.11.2996 (1997).
- 696 32 Florian, Griselda, Chikina, M. & Mark. A Temporal Switch in the Germinal Center
697 Determines Differential Output of Memory B and Plasma Cells. *Immunity* **44**, 116-130,
698 doi:10.1016/j.immuni.2015.12.004 (2016).
- 699 33 Purtha, W. E., Tedder, T. F., Johnson, S., Bhattacharya, D. & Diamond, M. S. Memory B
700 cells, but not long-lived plasma cells, possess antigen specificities for viral escape
701 mutants. *Journal of Experimental Medicine* **208**, 2599-2606, doi:10.1084/jem.20110740
702 (2011).
- 703 34 Pape, K. A., Taylor, J. J., Maul, R. W., Gearhart, P. J. & Jenkins, M. K. Different B Cell
704 Populations Mediate Early and Late Memory During an Endogenous Immune Response.
705 *Science* **331**, 1203-1207, doi:doi:10.1126/science.1201730 (2011).
- 706 35 Stadlbauer, D. *et al.* SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a
707 Serological Assay, Antigen Production, and Test Setup. *Current Protocols in*
708 *Microbiology* **57**, doi:10.1002/cpmc.100 (2020).
- 709 36 Davis, C. W. *et al.* Longitudinal Analysis of the Human B Cell Response to Ebola Virus
710 Infection. *Cell* **177**, 1566-1582.e1517, doi:10.1016/j.cell.2019.04.036 (2019).
- 711 37 Vander Heiden, J. A. *et al.* pRESTO: a toolkit for processing high-throughput sequencing
712 raw reads of lymphocyte receptor repertoires. *Bioinformatics* **30**, 1930-1932,
713 doi:10.1093/bioinformatics/btu138 (2014).
- 714 38 Ye, J., Ma, N., Madden, T. L. & Ostell, J. M. IgBLAST: an immunoglobulin variable
715 domain sequence analysis tool. *Nucleic Acids Research* **41**, W34-W40,
716 doi:10.1093/nar/gkt382 (2013).
- 717 39 Giudicelli, V., Chaume, D. & Lefranc, M. P. IMGT/GENE-DB: a comprehensive database
718 for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res* **33**,
719 D256-261, doi:10.1093/nar/gki010 (2005).

- 720 40 Gupta, N. T. *et al.* Change-O: a toolkit for analyzing large-scale B cell immunoglobulin
721 repertoire sequencing data: Table 1. *Bioinformatics* **31**, 3356-3358,
722 doi:10.1093/bioinformatics/btv359 (2015).
- 723 41 Gadala-Maria, D., Yaari, G., Uduman, M. & Kleinstein, S. H. Automated analysis of high-
724 throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V
725 gene segment alleles. *Proceedings of the National Academy of Sciences* **112**, E862-
726 E870, doi:10.1073/pnas.1417683112 (2015).
- 727 42 Gupta, N. T. *et al.* Hierarchical Clustering Can Identify B Cell Clones with High
728 Confidence in Ig Repertoire Sequencing Data. *The Journal of Immunology* **198**, 2489-
729 2499, doi:10.4049/jimmunol.1601850 (2017).
- 730 43 Zhou, J. Q. & Kleinstein, S. H. Cutting Edge: Ig H Chains Are Sufficient to Determine
731 Most B Cell Clonal Relationships. *The Journal of Immunology* **203**, 1687-1692,
732 doi:10.4049/jimmunol.1900666 (2019).
- 733 44 Wickham, H. ggplot2: Elegant graphics for data analysis. *Springer-Verlag New York*
734 (2016).
- 735 45 Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances
736 circular visualization in R. *Bioinformatics* **30**, 2811-2812,
737 doi:10.1093/bioinformatics/btu393 (2014).
- 738 46 Hoehn, K. B., Lunter, G. & Pybus, O. G. A Phylogenetic Codon Substitution Model for
739 Antibody Lineages. *Genetics* **206**, 417-427, doi:10.1534/genetics.116.196303 (2017).
- 740 47 Hoehn, K. B. *et al.* Repertoire-wide phylogenetic models of B cell molecular evolution
741 reveal evolutionary signatures of aging and vaccination. *Proceedings of the National*
742 *Academy of Sciences* **116**, 22664-22672, doi:10.1073/pnas.1906020116 (2019).
- 743 48 Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree : an r package for
744 visualization and annotation of phylogenetic trees with their covariates and other
745 associated data. *Methods in Ecology and Evolution* **8**, 28-36, doi:10.1111/2041-
746 210x.12628 (2017).
- 747 49 Hounkpe, B. W., Chenou, F., Franciele & Erich. HRT Atlas v1.0 database: redefining
748 human and mouse housekeeping genes and candidate reference transcripts by mining
749 massive RNA-seq datasets. *Nucleic Acids Research* **49**, D947-D955,
750 doi:10.1093/nar/gkaa609 (2021).
- 751 50 Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends in*
752 *Genetics* **29**, 569-574, doi:10.1016/j.tig.2013.05.010 (2013).
- 753 51 Valente, V. *et al.* Selection of suitable housekeeping genes for expression analysis in
754 glioblastoma using quantitative RT-PCR. *BMC Molecular Biology* **10**, 17,
755 doi:10.1186/1471-2199-10-17 (2009).
- 756 52 Frankish, A. *et al.* GENCODE reference annotation for the human and mouse genomes.
757 *Nucleic Acids Research* **47**, D766-D773, doi:10.1093/nar/gky955 (2019).
- 758 53 Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression
759 data analysis. *Genome Biology* **19**, doi:10.1186/s13059-017-1382-0 (2018).
- 760 54 Lefranc, M.-P. *et al.* IMGT unique numbering for immunoglobulin and T cell receptor
761 variable domains and Ig superfamily V-like domains. *Developmental & Comparative*
762 *Immunology* **27**, 55-77, doi:10.1016/s0145-305x(02)00039-3 (2003).
- 763

764 **Main Figure legends**

765 **Figure 1. Persistence of humoral immune responses to SARS-CoV-2 mRNA**

766 **vaccination. a**, Study design. Forty-three healthy adult volunteers (13 with a history of SARS-
767 CoV-2 infection) were enrolled, followed by BNT162b2 mRNA SARS-CoV-2 vaccination. Blood
768 (n=42) was collected before immunization, and at 3, 4, 5, 7, 15, and 29 weeks after primary
769 immunization. For 15 participants without a history of SARS-CoV-2 infection, aspirates of
770 ipsilateral axillary lymph nodes were collected at 3, 4, 5, 7, 15, and 29 weeks after primary
771 vaccination. For 11 participants without a history of SARS-CoV-2 infection, aspirates of bone
772 marrow were collected at 29 weeks after primary vaccination. **b**, Representative flow cytometry
773 plots of GC B cells (CD19⁺ CD3⁻ IgD^{low} BCL6⁺ CD38^{int} live singlet lymphocytes) and SARS-CoV-2
774 S staining on GC B cells in draining lymph nodes 29 weeks post-vaccination. **c**, Kinetics of total
775 (left) and S⁺ GC B cells (right) as gated in **b**. **d**, Representative ELISpot wells coated with the
776 indicated antigens or anti-immunoglobulin and developed in blue and red for IgG and IgA,
777 respectively, after plating the indicated numbers of magnetically enriched BMPCs. **e**,
778 Frequencies of BMPCs secreting IgG antibodies specific for the indicated antigens 29 weeks
779 after vaccination. **f**, Plasma IgG titers against SARS-CoV-2 S measured by ELISA in
780 participants without (red) and with (black) a history of SARS-CoV-2 infection. Horizontal lines
781 indicate geometric means, also shown above time points. Results are from one experiment
782 performed in duplicate. **g**, Representative flow cytometry plot of SARS-CoV-2 S staining on
783 MBCs (CD20⁺ CD38⁻ IgD^{low} CD19⁺ CD3⁻ live singlet lymphocytes) in blood 29 weeks after
784 primary vaccination. **h**, Frequencies of S⁺ MBCs in participants without (red) and with (black) a
785 history of SARS-CoV-2 infection as gated in **g**. Horizontal lines indicate median values in **e** and
786 **h**. Dotted lines indicate limits of detection in **e** and **f**. Symbols at each time point represent one
787 sample in **c** (n=15), **e** (n=11), **f** (n=38), and **h** (n=42).

788

789 **Figure 2. Identification of SARS-CoV-2 S-binding B cell clones in draining lymph**
790 **nodes. a**, Uniform manifold approximation and projection (UMAP) of scRNA-seq data from PBs
791 sorted from PBMC (upper) and whole FNA of draining axillary lymph nodes (lower), and UMAP
792 of B cell scRNA-seq clusters from each compartment (right). Each dot represents a cell, colored
793 by phenotype as defined by the gene expression profile. Total numbers of cells are at the top
794 right corner. FDC, follicular dendritic cell; GC, germinal center B cell; Mo, monocyte; NK, natural
795 killer cell; LNPC, lymph node plasma cell; PB, plasmablast; pDC, plasmacytoid dendritic cell;
796 RMB, resting memory B cell. **b**, Positive binding of recombinant monoclonal antibodies (mAbs)
797 derived from GC B cells (blue) or LNPCs (green) to SARS-CoV-2 S measured by ELISA. Areas
798 under the curve were calculated by setting the mean + three times the s.d. of background
799 binding to bovine serum albumin (BSA) as a baseline. Results are from one experiment
800 performed in duplicate. **c**, SARS-CoV-2 S-binding clones visualized on UMAP of B cell clusters
801 in each participant. Percentages are of SARS-CoV-2 S-binding clones within GC B cells (blue),
802 LNPCs (green), PBs (red), RMBs (pink) or naive B cells (yellow). Total numbers of cells are at
803 the bottom right corner.

804

805 **Figure 3. Maturation of SARS-CoV-2 S-binding B cells in the lymph node. a**, Clonal
806 overlap of SARS-CoV-2 S-binding sequences from bulk and single-cell BCR repertoire analysis
807 between PBs sorted from blood 4 week post-vaccination and GC B cells at indicated time
808 points. Arc length corresponds to the number of BCR sequences and chord width corresponds
809 to clonal group size. Purple chords correspond to clones spanning both the PB and GC
810 compartments. Grey chords correspond to clones spanning only the GC compartment.
811 Percentages are of GC B cell clones related to PBs at each time point. **b**, Comparison of
812 immunoglobulin heavy chain variable (*IGHV*) region nucleotide mutation frequency of clonally
813 related, SARS-CoV-2 S-binding PBs and GC B cell clones that are clonally linked at the
814 indicated time points. Median values are presented on the top of each data set. **c**, Comparison

815 of *IGHV* nucleotide mutation frequency of SARS-CoV-2 S-binding GC B cells at the indicated
816 time points. Horizontal lines and colored numbers represent median values. *P* values were
817 determined by Kruskal-Wallis test followed by Dunn's multiple comparison test. **d**, Percentages
818 of GC B cells expressing IGHG (blue), IGHA (red), IGHM (green) or IGHD (pink) at the early (E)
819 or the late (L) time point. The early time point represents 4, 5 or 7 weeks after immunization.
820 The late time point represents 15 weeks post-immunization. Total cell numbers are on the top of
821 each column. **e**, Circos diagram showing overlap of S-binding clonal groups between GC B cells
822 and LNPCs over combined time points. Arc length corresponds to the number of BCR
823 sequences and chord width corresponds to clonal group size. Purple chords link GC B cells to
824 clonally overlapping LNPCs. Percentages are of GC B cell clones overlapping with LNPCs or
825 *vice versa* in each participant. **f**, Comparison of *IGHV* nucleotide mutation frequency of clonally
826 related, S-binding GC B cell and LNPC at the indicated time points. Median values are
827 presented on the top of each data set. Each dot represents the median SHM of a clonal group
828 within the indicated compartment in **b** and **f**. *P* values were determined by paired two-sided non-
829 parametric Mann-Whitney test corrected for multiple testing using Benjamini and Hochberg's
830 method in **b** and **f**.

831

832 **Figure 4. Evolution of B cell clones induced by SARS-CoV-2 vaccination. a,**

833 Comparison of avidity indices of plasma IgG against SARS-CoV-2 S between the indicated time
834 points in participants without (red) and with (black) a history of SARS-CoV-2 infection. Avidity
835 index was defined as the ratio of optical density values obtained in the presence and absence of
836 8M urea. Data are *P* values were determined by Wilcoxon matched-pairs signed rank test.
837 Results are from one experiment performed in duplicate. **b**, *IGHV* nucleotide mutation frequency
838 of S-binding PB, LNPC, and BMPC at the indicated time points. Horizontal lines and colored
839 numbers represent median values. *P* values were determined by Kruskal-Wallis test followed by
840 Dunn's multiple comparison test. **c**, Phylogenetic trees of each clonal group showing inferred

841 relations between PB (squares), LNPC (triangles), and BMPC (diamonds). ELISA-verified S-
842 binding mAb IDs and clonal group IDs are indicated in the diagram. Horizontal branch length
843 represents the expected number of substitutions per codon in V-region genes, corresponding to
844 the key.
845

846 **Extended Data Figure legends**

847 **Extended Data Figure 1. Persistence of humoral immune responses to SARS-CoV-2**

848 **mRNA vaccination. a**, Flow cytometry gating strategies for GC B cells (Fig. 1b) and LNPCs
849 (defined as CD19⁺ CD3⁻ IgD^{low} CD20^{low} CD38⁺ BLIMP1⁺ CD71⁺ live singlet lymphocytes) in the
850 lymph node. **b**, Kinetics of total (left) and S⁺ LNPCs (right) as gated in **a**. **c**, Frequencies of
851 BMPCs secreting IgA antibodies specific for the indicated antigens 29 weeks after
852 immunization. **d, e**, Plasma antibody titers against SARS-CoV-2 S measured by ELISA in
853 participants without (red) and with (black) a history of SARS-CoV-2 infection in SARS-CoV-2
854 vaccinated (left, center) and unvaccinated (right) participants 29 weeks after the first vaccine
855 dose or symptom onset (**d**) and in vaccinated participants over time (**e**). *P* values were
856 determined by Kruskal-Wallis test followed by Dunn's multiple comparison test between
857 unvaccinated and both vaccinated groups (**d**), and by two-sided Mann-Whitney test (**e**).
858 Horizontal lines indicate median values in **c** and **e**. **f**, Flow cytometry gating strategies for MBCs
859 (CD19⁺ CD3⁻ IgD^{low} CD20⁺ CD38⁻ live singlet lymphocytes) and S⁺ MBCs (Fig. 1g) in blood.

860

861 **Extended Data Figure 2. Identification of SARS-CoV-2 S-binding B cell clones in the**

862 **lymph node. a**, Flow cytometry gating strategies for sorting PBs (defined as CD19⁺ CD3⁻ IgD^{low}
863 CD20^{low} CD38⁺ CD71⁺ live singlet lymphocytes) from blood. **b, d** UMAPs of scRNA-seq data
864 from PBs sorted from blood and FNA of draining axillary lymph nodes (b), and UMAPs of B cell
865 scRNA-seq clusters (d). **c, e**, Dot plots for the marker genes used for identifying annotated
866 clusters. **f**, Flow cytometry gating strategies for sorting GC B cells (CD19⁺ CD4⁻ IgD^{low} CD20⁺
867 CD38^{int} CXCR5^{high} CD71⁺ live singlet lymphocytes) and LNPCs (CD19⁺ CD4⁻ IgD^{low} CD20^{low}
868 CD38⁺ CXCR5^{low} CD71⁺ live singlet lymphocytes) from FNAs. **g**, Visualization of SARS-CoV-2
869 S-binding clones from all participants on integrated UMAP of B cell clusters. Percentages are of
870 SARS-CoV-2 S-binding clones within GC B cells (blue), LNPCs (green), PBs (red), RMBs (pink)
871 or naive B cells (yellow).

872

873 **Extended Data Figure 3. Maturation of SARS-CoV-2 S-binding B cells in the lymph**

874 **node. a**, Comparison of *IGHV* nucleotide mutation frequency of SARS-CoV-2 S-binding GC B
875 cells in each participant at the indicated time points. Horizontal lines represent median values.

876

877 **Extended Data Figure 4. Evolution of B cell clones induced by SARS-CoV-2**

878 **vaccination. a**, Comparison of *IGHV* nucleotide mutation frequency of each B cell subsets.

879 Horizontal lines represent median values. *P* values are determined by Kruskal-Wallis test

880 followed by Dunn's multiple comparison test.

881	Extended Data Table legends
882	Extended Data Table 1. Demographics of participants
883	Extended Data Table 2. Vaccine side effects
884	Extended Data Table 3. Frequencies of total and SARS-CoV-2 S-binding GC B cells
885	and LNPCs in draining axillary lymph nodes
886	Extended Data Table 4. Frequencies of CD14⁺ cells in lymph node samples
887	Extended Data Table 5. Cell counts and frequencies of overall and B cell clusters in
888	scRNA-seq of PBs sorted from blood and FNA from lymph nodes
889	Extended Data Table 6. Description of recombinant mAbs derived from GC B cells
890	and LNPCs
891	Extended Data Table 7. Processing of BCR and 5' gene expression data from scRNA-
892	seq
893	Extended Data Table 8. Processing of BCR reads from bulk-seq

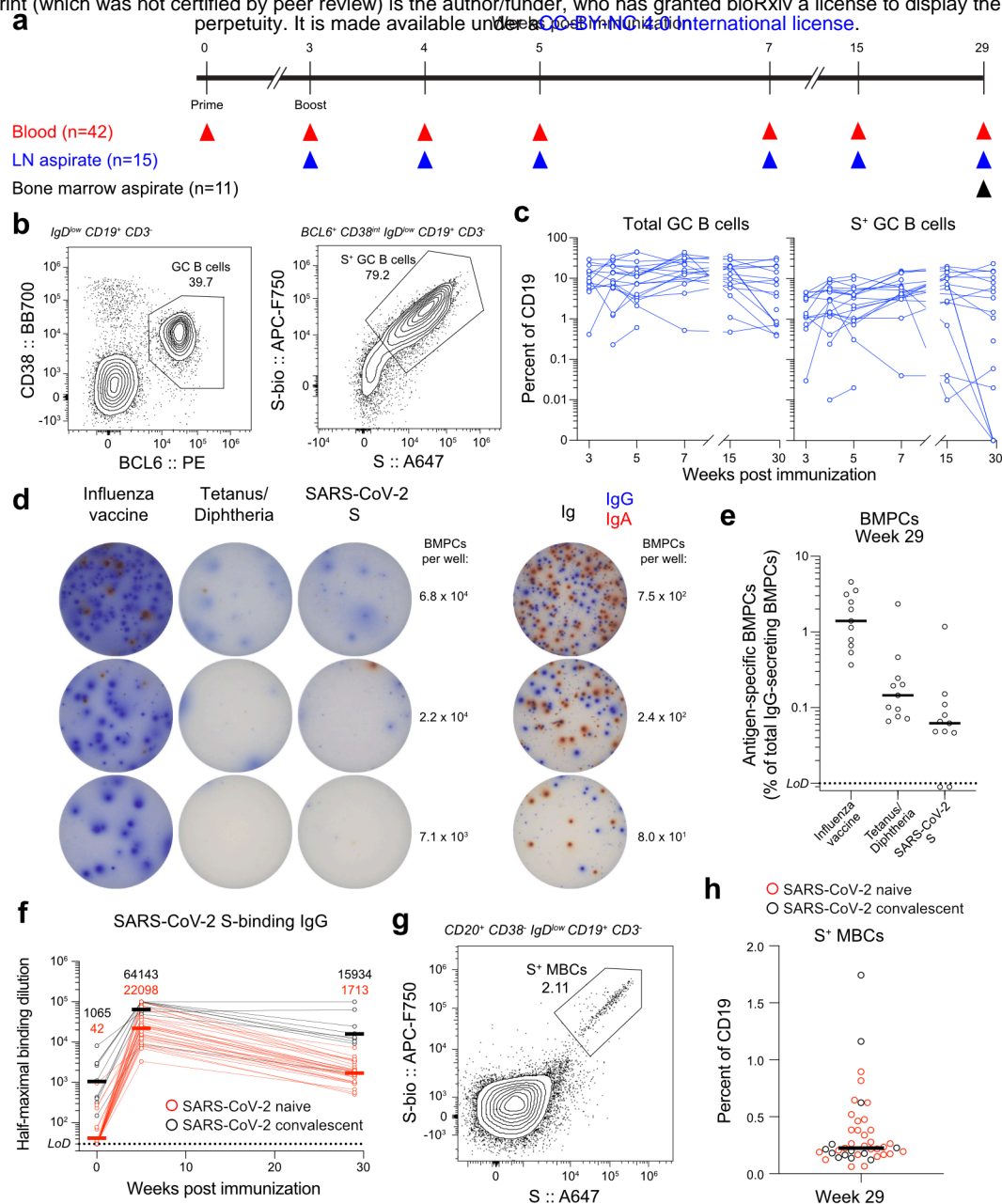


Figure 1. Persistence of humoral immune responses to SARS-CoV-2 mRNA vaccination. **a**, Study design. Forty-three healthy adult volunteers (13 with a history of SARS-CoV-2 infection) were enrolled, followed by BNT162b2 mRNA SARS-CoV-2 vaccination. Blood (n=42) was collected before immunization, and at 3, 4, 5, 7, 15, and 29 weeks after primary immunization. For 15 participants without a history of SARS-CoV-2 infection, aspirates of ipsilateral axillary lymph nodes were collected at 3, 4, 5, 7, 15, and 29 weeks after primary vaccination. For 11 participants without a history of SARS-CoV-2 infection, aspirates of bone marrow were collected at 29 weeks after primary vaccination. **b**, Representative flow cytometry plots of GC B cells (CD19⁺ CD3⁻ IgD^{low} BCL6⁺ CD38^{int} live singlet lymphocytes) and SARS-CoV-2 S staining on GC B cells in draining lymph nodes 29 weeks post-vaccination. **c**, Kinetics of total (left) and S⁺ GC B cells (right) as gated in **b**. **d**, Representative ELISpot wells coated with the indicated antigens or anti-immunoglobulin and developed in blue and red for IgG and IgA, respectively, after plating the indicated numbers of magnetically enriched BMPCs. **e**, Frequencies of BMPCs secreting IgG antibodies specific for the indicated antigens 29 weeks after vaccination. **f**, Plasma IgG titers against SARS-CoV-2 S measured by ELISA in participants without (red) and with (black) a history of SARS-CoV-2 infection. Horizontal lines indicate geometric means, also shown above time points. Results are from one experiment performed in duplicate. **g**, Representative flow cytometry plot of SARS-CoV-2 S staining on MBCs (CD20⁺ CD38⁻ IgD^{low} CD19⁺ CD3⁻ live singlet lymphocytes) in blood 29 weeks after primary vaccination. **h**, Frequencies of S⁺ MBCs in participants without (red) and with (black) a history of SARS-CoV-2 infection as gated in **g**. Horizontal lines indicate median values in **e** and **h**. Dotted lines indicate limits of detection in **e** and **f**. Symbols at each time point represent one sample in **c** (n=15), **e** (n=11), **f** (n=38), and **h** (n=42).

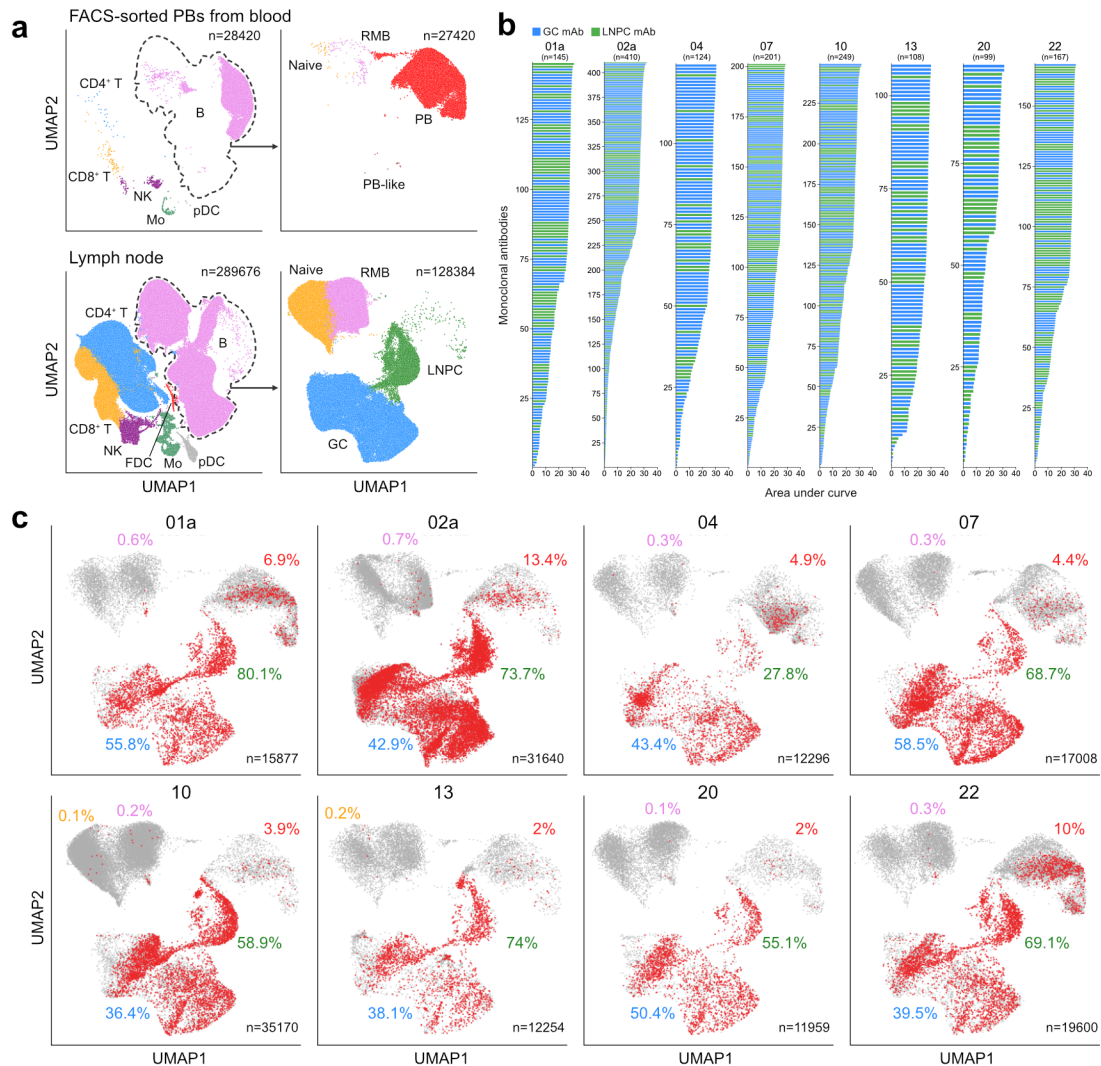


Figure 2. Identification of SARS-CoV-2 S-binding B cell clones in draining lymph nodes.

a, Uniform manifold approximation and projection (UMAP) of scRNA-seq data from PBs sorted from PBMC (upper) and whole FNA of draining axillary lymph nodes (lower), and UMAP of B cell scRNA-seq clusters from each compartment (right). Each dot represents a cell, colored by phenotype as defined by the gene expression profile. Total numbers of cells are at the top right corner. FDC, follicular dendritic cell; GC, germinal center B cell; Mo, monocyte; NK, natural killer cell; LNPC, lymph node plasma cell; PB, plasmablast; pDC, plasmacytoid dendritic cell; RMB, resting memory B cell. **b**, Positive binding of recombinant monoclonal antibodies (mAbs) derived from GC B cells (blue) or LNPCs (green) to SARS-CoV-2 S measured by ELISA. Areas under the curve were calculated by setting the mean + three times the s.d. of background binding to bovine serum albumin (BSA) as a baseline. Results are from one experiment performed in duplicate. **c**, SARS-CoV-2 S-binding clones visualized on UMAP of B cell clusters in each participant. Percentages are of SARS-CoV-2 S-binding clones within GC B cells (blue), LNPCs (green), PBs (red), RMBs (pink) or naive B cells (yellow). Total numbers of cells are at the bottom right corner.

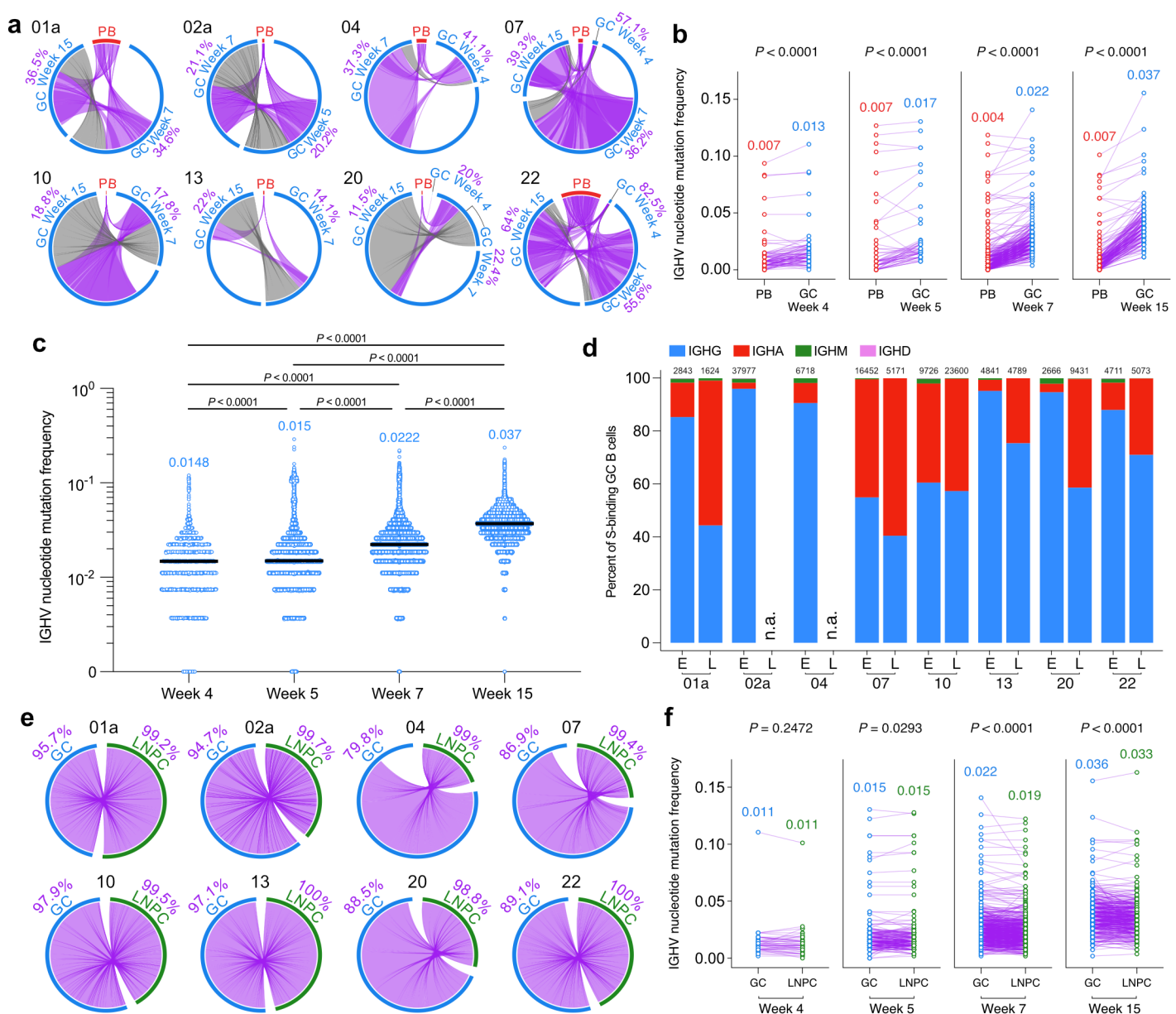


Figure 3. Maturation of SARS-CoV-2 S-binding B cells in the lymph node. a, Clonal overlap of SARS-CoV-2 S-binding sequences from bulk and single-cell BCR repertoire analysis between PBs sorted from blood 4 week post-vaccination and GC B cells at indicated time points. Arc length corresponds to the number of BCR sequences and chord width corresponds to clonal group size. Purple chords correspond to clones spanning both the PB and GC compartments. Grey chords correspond to clones spanning only the GC compartment. Percentages are of GC B cell clones related to PBs at each time point. **b**, Comparison of immunoglobulin heavy chain variable (*IGHV*) region nucleotide mutation frequency of clonally related, SARS-CoV-2 S-binding PBs and GC B cell clones that are clonally linked at the indicated time points. Median values are presented on the top of each data set. **c**, Comparison of *IGHV* nucleotide mutation frequency of SARS-CoV-2 S-binding GC B cells at the indicated time points. Horizontal lines and colored numbers represent median values. P values were determined by Kruskal-Wallis test followed by Dunn's multiple comparison test. **d**, Percentages of GC B cells expressing IGHG (blue), IGHA (red), IGHM (green) or IGHD (pink) at the early (E) or the late (L) time point. The early time point represents 4, 5 or 7 weeks after immunization. The late time point represents 15 weeks post-immunization. Total cell numbers are on the top of each column. **e**, Circos diagram showing overlap of S-binding clonal groups between GC B cells and LNPCs over combined time points. Arc length corresponds to the number of BCR sequences and chord width corresponds to clonal group size. Purple chords link GC B cells to clonally overlapping LNPCs. Percentages are of GC B cell clones overlapping with LNPCs or vice versa in each participant. **f**, Comparison of *IGHV* nucleotide mutation frequency of clonally related, S-binding GC B cell and LNPC at the indicated time points. Median values are presented on the top of each data set. Each dot represents the median SHM of a clonal group within the indicated compartment in **b** and **f**. P values were determined by paired two-sided non-parametric Mann-Whitney test corrected for multiple testing using Benjamini and Hochberg's method in **b** and **f**.

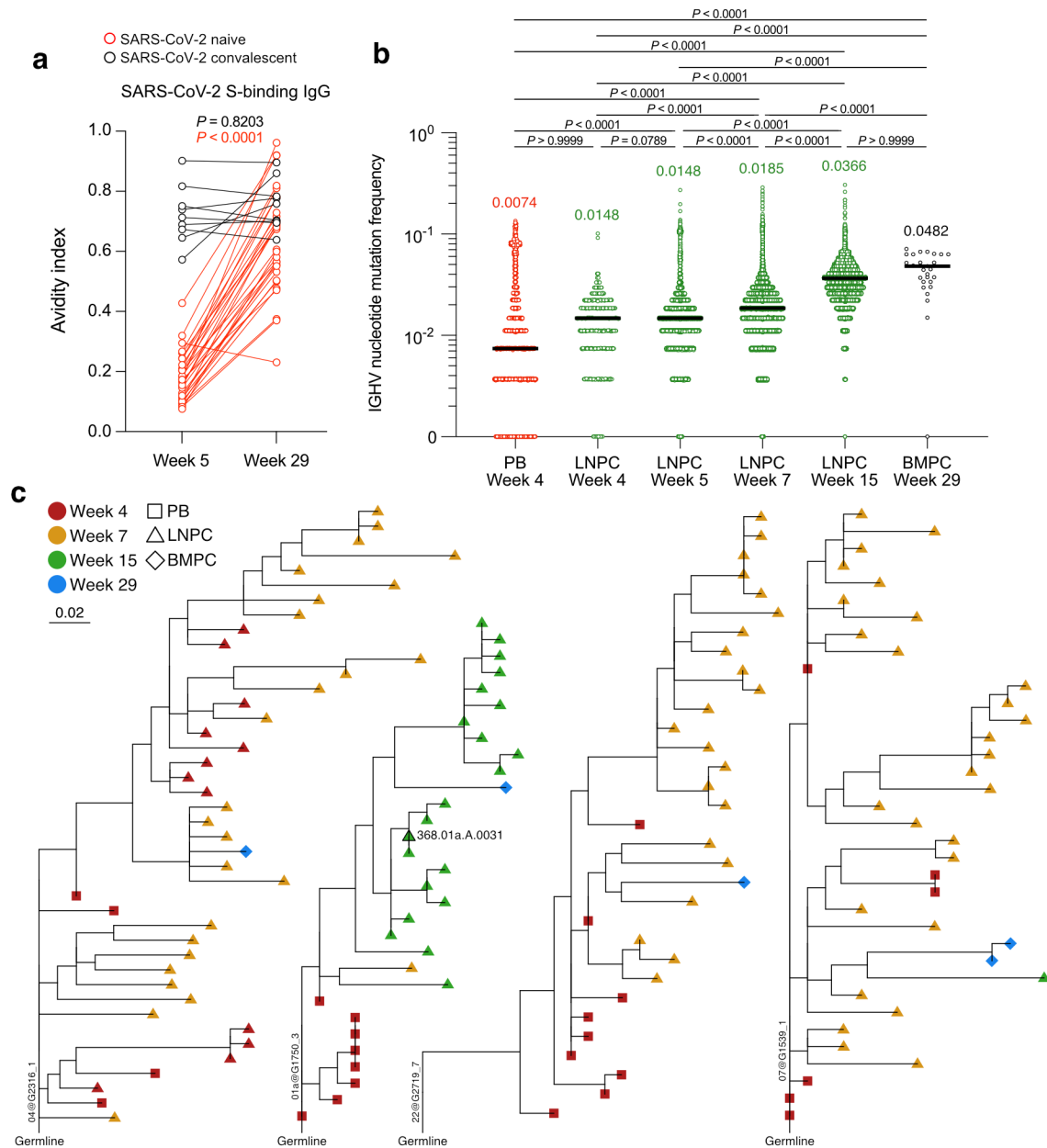


Figure 4. Evolution of B cell clones induced by SARS-CoV-2 vaccination. **a**, Comparison of avidity indices of plasma IgG against SARS-CoV-2 S between the indicated time points in participants without (red) and with (black) a history of SARS-CoV-2 infection. Avidity index was defined as the ratio of optical density values obtained in the presence and absence of 8M urea. Data are P values were determined by Wilcoxon matched-pairs signed rank test. Results are from one experiment performed in duplicate. **b**, IGHV nucleotide mutation frequency of S-binding PB, LNPC, and BMPC at the indicated time points. Horizontal lines and colored numbers represent median values. P values were determined by Kruskal-Wallis test followed by Dunn's multiple comparison test. **c**, Phylogenetic trees of each clonal group showing inferred relations between PB (squares), LNPC (triangles), and BMPC (diamonds). ELISA-verified S-binding mAb IDs and clonal group IDs are indicated in the diagram. Horizontal branch length represents the expected number of substitutions per codon in V-region genes, corresponding to the key.