

1 Pre-vaccination and early B cell signatures predict antibody response to SARS-CoV-2 mRNA
2 vaccine

3
4 Lela Kardava^{1,10}, Nicholas Rachmaninoff^{2,10}, William W. Lau^{2,10}, Clarisa M. Buckner¹, Krittin
5 Trihemasava¹, Felipe Lopes de Assis¹, Wei Wang¹, Xiaozhen Zhang¹, Yimeng Wang³, Chi-I
6 Chiang³, Sandeep Narpala⁴, Robert Reger⁵, Genevieve E. McCormack¹, Catherine A. Seamon⁶,
7 Richard W. Childs^{5,7}, Anthony F. Suffredini⁶, Jeffrey R. Strich^{6,7}, Daniel S. Chertow^{1,6,7}, Richard
8 T. Davey¹, Michael C. Sneller¹, Sarah O'Connell⁴, Yuxing Li^{3,8}, Adrian McDermott⁴, Tae-Wook
9 Chun¹, Anthony S. Fauci^{1,11}, John S. Tsang^{2,9,11*}, Susan Moir^{1,11*}

10
11 ¹Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases
12 (NIAID), National Institutes of Health (NIH), Bethesda, MD, USA; ²Multiscale Systems Biology
13 Section, Laboratory of Immune System Biology, NIAID, NIH, Bethesda, MD, USA; ³Institute
14 for Bioscience and Biotechnology Research, Rockville, MD, USA; ⁴Vaccine Research Center,
15 NIAID, NIH, Bethesda, MD, USA; ⁵Cellular and Molecular Therapeutics Branch, National
16 Heart, Lung and Blood Institute, NIH, Bethesda, MD, USA; ⁶Critical Care Medicine
17 Department, National Institutes of Health Clinical Center, NIH, Bethesda, MD, USA; ⁷United
18 States Public Health Service Commissioned Corps, Rockville, MD, USA; ⁸Department of
19 Microbiology and Immunology and Center for Biomolecular Therapeutics, University of
20 Maryland School of Medicine, Baltimore, MD, USA; ⁹NIH Center for Human Immunology,
21 NIAID, NIH, Bethesda, MD, USA

22
23

24 *Corresponding authors:

25 John Tsang and Susan Moir, National Institutes of Health, 9000 Rockville Pike, Bethesda,

26 Maryland 20892, USA,

27 e-mail: john.tsang@nih.gov and smoir@niaid.nih.gov

28

29 ¹⁰These authors contributed equally to this work.

30 ¹¹These authors contributed equally to this work.

31

32

33 **Abstract**

34 SARS-CoV-2 mRNA vaccines are highly effective, although weak antibody responses
35 are seen in some individuals with correlates of immunity that remain poorly understood. Here we
36 longitudinally dissected antibody, plasmablast, and memory B cell (MBC) responses to the two-
37 dose Moderna mRNA vaccine in SARS-CoV-2-uninfected adults. Robust, coordinated IgA and
38 IgG antibody responses were preceded by bursts of spike-specific plasmablasts after both doses,
39 but earlier and more intensely after dose two. Distinct antigen-specific MBC populations also
40 emerged post-vaccination with varying kinetics. We identified antigen non-specific pre-
41 vaccination MBC and post-vaccination plasmablasts after dose one and their spike-specific
42 counterparts early after dose two that correlated with subsequent antibody levels. These baseline
43 and response signatures can thus provide early indicators of serological efficacy and explain
44 response variability in the population.

45 The pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-
46 2) instigated rapid worldwide COVID-19 vaccine prioritization strategies. Several vaccine
47 candidates were developed, including two vaccines based on novel mRNA platforms (Moderna
48 mRNA-1273 and the Pfizer/BioNTech BNT162b2). Both mRNA vaccines encode a stabilized
49 ectodomain of the spike protein trimer (S-2P) derived from the Wuhan Hu-1 isolate¹, and are
50 given in two vaccine doses, referred to henceforth as v1 and v2. Both mRNA vaccines have been
51 shown to be highly protective and elicit strong B cell and antibody responses^{2,3}, although poorer
52 responses have also been seen in some individuals, such as the elderly⁴ and transplant recipients⁵⁻
53 ⁷, raising the question of what determines antibody response levels and whether early correlates
54 of immunity can be defined. Studies on other vaccines have shown that pre-vaccination
55 signatures and early circulating B cell responses involving plasmablasts (PB) and activated
56 memory B cells (MBC) can predict the magnitude and longevity of neutralizing antibodies
57 following vaccination⁸⁻¹¹. The Pfizer vaccine has been shown to induce robust PB and MBC
58 responses in blood and draining lymph nodes^{4,12}, but the extent by which these responses differ
59 across individuals and whether they are associated with antibody levels have not been assessed.

60 To address gaps in correlates of humoral immunity to mRNA vaccines, we evaluated
61 antibody and B cell responses following vaccination with mRNA-1273 in 21 healthy SARS-
62 CoV-2-uninfected adults (Extended Data Table 1). Blood was drawn serially over a period of
63 ~60 days (D) and paired serum and cellular assays were performed at each timepoint (Fig. 1a and
64 Extended Data Table 1). Given the fragility of PB, cellular assays were performed on freshly
65 isolated cells while sera were cryopreserved for antibody assays. Antibody binding to S-2P, its

66 receptor-binding domain (RBD) and the nucleoprotein (NP), was measured using a multiplex
67 platform¹. Strong IgG and IgA responses were induced, starting around D10, to both S-2P and
68 RBD (Fig. 1b), although the magnitude was highly variable across vaccinees at v2D28 (c.v. >
69 100%), spanning 2-3 orders of magnitude for both IgA and IgG titers (Fig. 1b, right panels). The
70 IgM response was weak across all vaccinees (Fig. 1b). This is consistent with recent reports for
71 the mRNA vaccines^{13,14}, yet in contrast to strong responses in patients who recovered from mild
72 to severe COVID-19¹⁴⁻¹⁷ (Fig. 1c). NP antibodies were also low in vaccinees (Fig. 1c), as
73 expected for SARS-CoV-2-uninfected people. Strong correlations were observed among RBD
74 and S-2P antibodies (Fig. 1d), but the correlation between IgA RBD and IgA S-2P was higher
75 than that between their IgG counterparts. The inhibition of RBD binding to the spike protein
76 receptor ACE2 by serum antibodies, a surrogate for neutralization capacity, also revealed a range
77 of responses (Fig. 1e) that correlated with RBD IgG and IgA binding antibodies (Extended Data
78 Fig. 1a).

79 In general, B cell responses to vaccination are detected in the peripheral blood in two
80 distinct phases; the first consisting of a short burst of PB, typically detected around D7, followed
81 by a slower phase that leads to the establishment of a pool of long-lived MBC¹⁸. Antigen-specific
82 B cells can be identified by flow cytometry using protein tetramers; this is an approach that has
83 been used to track SARS-CoV-2 spike-specific responses following COVID-19 infection or
84 vaccination^{15,19-21}. We used a pair of RBD and spike subunit 1 (S1) tetramers to track spike-
85 specific B cell responses of vaccinees, and as expected, dual RBD⁺S1⁺ and single S1⁺ PB
86 became detectable at v1D10 while corresponding MBC became detectable at v1D14 (Fig. 1f and

87 Extended Data Fig. 1b). S-2P tetramers also clearly detected RBD⁺ within S1⁺ and S1⁺ within S-
88 2P⁺ B cells, but they did not clearly identify S-2P⁺ PB (Extended Data Fig. 1c). We thus focused
89 on RBD and S1 tetramers to simultaneously measure spike-specific responses among all B cell
90 populations. We validated the approach for PB by showing that 1), frequencies of RBD⁺ and S1⁺
91 PB measured by flow cytometry were strongly correlated to those measured by the conventional
92 enzyme-linked immunospot (ELISpot) assay (Extended Data Fig. 1d); and 2), IgG PB could be
93 detected without the need for permeabilization, as is required by other methods²². Over 95% of
94 PB isotypes were detected in the absence of permeabilization and their distribution was largely
95 similar in the presence or absence of permeabilization (Extended Data Fig. 1e).

96 B cells that circulate in the peripheral blood are highly heterogeneous, especially MBC;
97 their phenotypes and functions vary with the source and chronicity of antigen exposure²³.
98 Accordingly, we comprehensively evaluated B cells with a flow cytometric panel of 15
99 antibodies against primarily B cell markers and two spike tetramers (Extended Data Table 2).
100 Unsupervised clustering analysis of the 15 cell-surface markers on CD19⁺ single B cells,
101 performed on all vaccinees and all time-points simultaneously, revealed 30 clusters grouped by
102 eight major B cell populations, including naïve, PB, and MBC subsets (Fig. 2a; also Extended
103 Data Fig. 2a for more granular depictions of individual clusters, and Extended Data Table 3 for
104 detailed annotations). A corresponding mean fluorescence intensity (MFI) heatmap delineated
105 the surface marker phenotype, isotype, and specificity of individual cell clusters (Fig. 2b), and
106 identified naïve B cells (C1) as the most abundant cluster, as expected for peripheral blood²².
107 Five populations of conventional (CD27⁺CD20⁺CD21⁺) MBC, each distinguished by isotype:

108 IgA⁺ (C14), IgG⁺ (C2 and C4) and IgD/M (C19), were prevalent. Both IgG (C9) and IgA (C13)
109 PB were identified, as well as eight nonconventional MBC with lower abundance (Fig 2b;
110 Extended Data Table 3 for details). IgG PB (C9) contained the highest proportion of RBD⁺S1⁺
111 cells, followed by IgA PB (C13), and the nonconventional IgG⁺ MBC C5 (Fig. 2b). Our
112 longitudinal tracking of vaccinees revealed that RBD⁺S1⁺ PB were first detected at v1D10, then
113 subsided until v2D5-7, while other RBD⁺S1⁺ B cells (namely MBC) became visible at v1D14
114 and intensified significantly following v2 (Fig. 2c, d), consistent with the analyses performed
115 with manual gating (Fig. 1f and Extended Data Fig. 1b).

116 We next used a linear model to search for cell clusters (both antigen non-specific and
117 specific) whose frequencies varied significantly as a function of time in response to vaccination
118 (Fig. 3a). Antigen non-specific clusters exhibited distinct patterns of response kinetics (Fig. 3b,
119 c). Most notable were the decreasing nonconventional CD27-IgG⁺ MBC C3 and C6 after v1 and
120 v2, while the two PB, IgG C9 and IgA C13, and the IgA pre-PB C12 were sharply increasing
121 after each dose. Other temporally changing clusters included several MBC and GC founder B
122 cells (C27) that underwent modest changes after v1 and v2 (Fig. 3b, c), possibly a reflection of
123 trafficking to and from lymphoid tissues.

124 Spike-specific B cell frequencies, measured as a fraction of RBD⁺S1⁺ cells among cells
125 within each cluster, also exhibited varying patterns of temporal responses (Fig. 3d). Two-peak
126 responses with stronger increases in v2 than v1 were observed for the two PB clusters C9 and
127 C13, nonconventional MBC C11 (CD27^{lo}IgA⁺) and C6 (CD27-IgG⁺), albeit with differences in
128 timing. For example, C9 (IgG) and C13 (IgA) PB had an initial modest burst beginning at

129 v1D10, followed by a second stronger but shorter burst at v2D5-7 (Fig. 3e). RBD⁺S1⁺ cells
130 among two pairs of conventional/nonconventional MBC, namely C2/C5 and C3/C4 respectively,
131 also underwent coordinated changes following v2 (Fig. 3d). It is notable, however, that while
132 RBD⁺S1⁺ cells in all four of these MBC clusters showed trends of declines from their peaks by
133 v2D28, those in the two nonconventional MBC (C3 and C5) appeared to drop more precipitously
134 than in the two conventional MBC (C2 and C4; Fig. 3e). These findings are consistent with
135 recent reports that spike-specific memory responses several months after SARS-CoV-2 infection
136 are enriched within CD21⁺CD27⁺ MBC^{17,21}, which have a phenotype similar to C2 and C4.

137 Despite the coherent changes observed across subjects, substantial heterogeneity in B cell
138 responses existed among vaccinees. Independent of antigen specificity, the magnitude of PB
139 increases is known to be a correlate of antibody responses for vaccines such as influenza^{24,25}. We
140 thus assessed associations between the changes in non-specific cluster frequencies over the
141 course of v1 and v2 relative to the respective baselines with IgA and IgG RBD and/or S-2P titers
142 at v2D28 by using linear models accounting for age and gender (Extended Data Fig. 3a, b). Both
143 IgA and IgG PB (C9 and C13) at v1D10 were indeed positively associated, albeit mildly, with
144 both RBD and S-2P IgA titers (but not IgG; Extended Data Fig. 3a, c, d), while several MBC
145 clusters (C3, C11, C24) at as early as v1D7 and v1D10 were correlated with S-2P IgG titers
146 (Extended Data Fig. 3a).

147 The frequency of spike-specific PB on v2D7 spanned a wide range (Fig. 3e), raising the
148 question of whether those with depressed PB responses also had lower antibody titers following
149 v2. Thus, we next used the same linear model to search for spike-specific correlates (Fig. 4a, b).

150 Indeed, at v2D7 and v2D10, IgG PB (C9) were correlated with IgG and IgA RBD antibodies,
151 while IgA PB (C13) at v2D7 were associated with IgA S-2P antibodies (Fig. 4b, c). Thus, in
152 contrast to antigen non-specific PB, spike-specific PB appeared as serological correlates only
153 after v2. However, the frequency of RBD⁺S1⁺ cells within IgA PB C13 at v1D10 predicted the
154 fraction of RBD⁺S1⁺ cells in the C2 MBC at v2D7 (Extended Data Fig. 3e), suggesting that the
155 magnitude of the spike-specific PB expansion after v1 is associated with the memory response
156 after v2. A recent preprint did not find a correlation between antibody response and
157 transcriptional modules enriched for PB at D7 following the second dose of the Pfizer vaccine²⁶,
158 likely due to differences in assessing PB responses using blood transcriptional signatures versus
159 our direct measurement of fresh, antigen-specific PB. In addition to PB, spike-specific C6 was a
160 positive correlate at v2D0 (Fig. 4a, d) when these cells reached a first peak (Fig. 3d). C6 are
161 CD27⁻IgG⁺ MBC known to have lower mutational burdens than their CD27-expressing
162 counterparts²⁷, and may as such, reflect products of early events of the antigen-driven maturation
163 process after v1. C6 are also CD21^{hi}, and similar to a stable pool of CD27⁻ MBC that are
164 generated in response to new influenza variants²⁸.

165 Most of the other antigen-specific correlates were positively associated with the titer
166 response and reflect changes after the second dose (Fig. 4b), including spike-specific MBC. At
167 v2D14 and v2D28, the frequency of RBD⁺S1⁺ cells in conventional MBC C2 was positively
168 correlated with RBD/S-2P IgA and IgG antibodies (Fig. 4b, e), consistent with their role in
169 sustaining immunological memory^{17,21}. In contrast, it is notable that C5, the nonconventional
170 MBC with features of vaccine-induced activated MBC²⁴, and that we found to contain a strong

171 RBD/S1 response following v2 (Fig. 3d, e), did not correlate with endpoint antibodies (data not
172 shown). Together, these may reflect a unique response feature of this mRNA vaccine, but timing
173 or insufficient statistical power may have played a role in the lack of detection of correlation
174 between C5 and antibodies.

175 Among the clusters that correlated with antibodies, several were not of the same isotype.
176 While it is possible that IgG B cells could give rise to IgA-secreting cells²⁹, the reverse cannot
177 occur. Thus, the most likely explanation for inter-isotype correlations is that most are not causal,
178 but reflect a coordinated immunologic response driven by shared mechanisms, as indicated by
179 the strong correlations between IgG and IgA antibodies (Fig. 1d). Indeed, when we used the
180 correlated component (the first principal component) of the v2D28 IgG and IgA RBD and S-2P
181 antibody titers as an isotype-independent endpoint, we found many of the positive antigen-
182 specific correlates highlighted above, including C6 MBC at v2D0, C9 PB at v2D7, and C2 MBC
183 at v2D14 (Extended Data Fig. 3f), indicating that these correlates reflected isotype independent
184 responses that potentially determined the magnitude of antibodies induced by the mRNA
185 vaccine.

186 Intriguingly, aside from confirming the early plasmablast correlates at v1D10 above, the
187 isotype independent analysis revealed additional non-specific correlates from as early as the
188 baseline before the first dose of vaccination (Fig. 4f): C30 MBC was a positive baseline predictor
189 independent of age and gender, suggesting that the frequency of these circulating unswitched
190 CD138⁺ MBC reflected an antigen non-specific “set point” for humoral responses to naïve
191 antigens⁸. Consistent with this notion, C30 did not appear again as a baseline (v2D0) correlate

192 for the second dose (data not shown), probably because v2 elicited a recall response.
193 Interestingly, by v1D7, this population was a negative correlate (Fig. 4f), perhaps because its
194 lower cell frequency in blood by then reflected greater ongoing B cell activation and
195 differentiation into class-switched PB and MBC. Together, our analyses revealed both antigen
196 non-specific and spike-specific predictive signatures of v2D28 antibody responses. These
197 signatures could help predict and monitor the serological efficacy of SARS-CoV-2 mRNA
198 vaccines and pave the way to a better understanding of weakened responses to this novel vaccine
199 platform, such as those associated with age and chronically compromised immune systems.
200
201

202 **Methods**

203 **Study design and participants**

204 Twenty SARS-CoV-2-uninfected NIH employees and one community member who were
205 eligible to receive an Emergency Use Authorization COVID-19 vaccine were recruited to study
206 longitudinal vaccine responses. The 21 participants received a first dose of the Moderna mRNA-
207 1273 vaccine and 25 to 34 days later, 20 participants received a second dose of the same vaccine
208 (Extended Data Table 1). One participant was lost to follow-up after testing positive for COVID-
209 19 between doses. To evaluate differences in SARS-CoV-2 antibody responses between
210 vaccination and infection, we recruited twenty-one patients with severe COVID-19 (Extended
211 Data Table 4). An additional 11 participants who received both doses of the Moderna mRNA-
212 1273 vaccine were recruited for the validation analyses of plasmablast (PB) frequencies between
213 flow cytometry and ELISpot; nine were SARS-CoV-2-uninfected and two had recovered from
214 COVID-19 infection. Research phlebotomy was performed at the NIH Clinical Research Center
215 in Bethesda, MD under protocols approved by the NIH Institutional Review Board,
216 ClinicalTrials.gov identifiers: NCT00001281, NCT04411147, NCT04280705, and
217 NCT04579393. All participants provided written informed consent.

218

219 **Blood sample collection and processing**

220 Peripheral blood mononuclear cells (PBMC) were isolated by Ficoll density gradient
221 centrifugation from whole blood collected in EDTA vacutainer tubes. Serum was isolated by
222 centrifugation of clotted whole blood collected in SST vacutainer tubes and stored at -80°C.

223

224 **SARS-CoV-2-binding antibody assay**

225 Serum samples were heat inactivated at 56°C for 60 minutes. A 4-plex antibody binding
226 assay was performed using an electrochemiluminescence immunoassay analyzer (ECLIA)
227 developed by Meso Scale Discovery (MSD). Each well of MSD SECTOR® plates was precoated
228 by manufacturer (MSD) with SARS-CoV-2 spike (S-2P), receptor binding domain (RBD)
229 protein, nucleocapsid (N) protein and a Bovine Serum Albumin (BSA) in a specific spot-
230 designation for each antigen. Plates were blocked at room temperature (RT) for 60 minutes with
231 MSD blocker A solution containing 5% BSA. Plates were washed and MSD reference standard
232 (calibrator), QC test sample (pool of COVID-19 convalescent sera) and human serum test
233 samples were added in duplicate in an 8-point dilution series and reference standards were added
234 in triplicate. MSD Control sera (low, medium and high) were added undiluted in triplicate.
235 Samples were incubated with shaking at RT for 4 hours on a Titramax Plate shaker (Heidolph).
236 Plates were washed and incubated with MSD SULFO-TAG™ anti-human IgG, IgA or IgM
237 detection antibodies at RT for 60 minutes with shaking. Plates were washed, MSD GOLD™ read
238 buffer containing electrochemiluminescence (ECL) substrate was added, and plates were read
239 with the MSD MESO Sector S 600 detection system. Analyses were performed with Excel
240 (Microsoft) and Prism 9.0 (Graphpad) software and antibody concentrations were assigned
241 arbitrary units (AU/ml) by interpolation from the standard curve. Densities of antibody
242 concentrations at endpoint (v2D28) were estimated using a gaussian kernel with bandwidth

243 automatically selected through biased cross validation using the `stat_density` function from
244 `ggplot2` (3.3.3) with `bw = "bcv"`.

245

246 **RBD-ACE2 blocking assay**

247 Samples were prepared as for the 4-plex binding assay. 384-well plates precoated with
248 RBD were supplied by the manufacturer (MSD). Plates were blocked at room temperature (RT)
249 for 30 minutes with MSD blocker A solution containing 5% BSA. Plates were washed, test
250 samples were added at dilutions of 1:10, 1:20 and 1:40, and incubated with shaking at RT for 60
251 minutes. Human ACE2 conjugated with SULFO-TAG was added and plates were further
252 incubated to allow binding to RBD. Plates were washed, ECL substrate added, and plates read as
253 in antibody binding assay. Fold reduction in ECL response for each sample was calculated
254 against based on signal emitted in wells in absence of sample (assay diluent).

255

256 **Recombinant biotinylated RBD protein**

257 A SARS-CoV-2 RBD construct containing a His-tag and Avi-tag was generated, as
258 previously described³⁰. The residues 319-541 of the S protein were codon optimized with N-
259 terminal of signal peptide (MFVFLVLLPLVSSQ) and C-terminal of 6-His tag and Avi-tag
260 (GLNDIFEAQKIEWHE). The DNA encoding sequence was cloned into the mammalian cell
261 expression vector pCAGGS and confirmed by sequencing, prior to transient transfection in
262 FreeStyle 293-F cells with 293fectin transfection reagent (ThermoFisher). Culture supernatants
263 were harvested at 5 days post transfection, filtered, and purified by in-house packed affinity

264 purification column with Complete His-tag purification resin (Roche). Elutes were buffer
265 exchanged with phosphate-buffered saline (PBS), and concentrated using an Amicon Ultra 10
266 kDa molecular weight cutoff concentrator (Millipore). Biotinylation was performed with a BirA
267 biotin-protein ligase standard reaction kit (Avidity), according to the manufacturer's instructions.
268 Excess biotin was removed by five buffer exchanges with an ultra 10K concentrator (Amicon).
269

270 **B-cell spike-specific responses and phenotyping by spectral flow cytometry**

271 A 17-color panel was developed to phenotype B-cell populations and identify SARS-
272 CoV-2-specific B cells among PBMC (Extended Data Table 2 for list and source of antibodies
273 and biotinylated spike proteins). The biotinylated spike proteins were tetramerized with
274 fluorescently labeled streptavidin (SA) as follows: S1 with SA-R-Phycoerythrin (PE), RBD with
275 SA-Allophycocyanin (APC), and S-2P with SA-Alexa Fluor 488 (Thermo Fisher Scientific). In a
276 stepwise process, 1/5 of the molar equivalent of the SA-fluorochrome reagent was added to the
277 biotinylated protein at 20-min intervals until the molar ratio of biotinylated protein and
278 streptavidin-fluorochrome reached 4:1. Incubations were carried out at 4°C with gentle rocking.
279 To titrate the labeled protein tetramers and establish background and antigen-specificity, freshly
280 isolated or cryopreserved PBMC from SARS-CoV-2-uninfected and recovered infected
281 individuals were used as negative and positive controls, respectively. For vaccinees, 10⁶ freshly
282 isolated PBMC were stained with a cocktail containing the 15 panel antibodies and 160 ng each
283 of PE-conjugated S1 and APC-conjugated RBD in staining buffer (2% FBS/PBS) supplemented
284 with Brilliant Stain Buffer Plus (BD Biosciences) at 4°C for 30 minutes. In a second 18-color

285 panel, 400 ng of Alexa Fluor 488 conjugated S-2P was added to the cocktail. The stained cells
286 were acquired on an Aurora spectral cytometer (Cytek Biosciences) and analyzed using FlowJo
287 v10.7.1 (BD Biosciences).

288

289 **Intracellular flow cytometry**

290 PBMC were first stained with antibodies against cell surface markers CD3, CD19, CD20
291 and CD27, fixed (Lysing Solution, BD Biosciences), permeabilized (Permeabilizing Solution 2;
292 BD Biosciences) and stained with antibodies against IgG, IgA, IgD and IgM (antibody details in
293 Extended Data Table 5). The stained cells were acquired on a FACS Canto II flow cytometer
294 (BD Biosciences) and analyzed using FlowJo software v9.9.6 (BD Biosciences).

295

296 **ELISpot assay to enumerate SARS-CoV-2 spike-specific PB**

297 Spike protein S1 and RBD-specific antibody-secreting PB were enumerated by
298 modifying the antigen-specific portion of a previously described ELISpot assay^{31,32}. Briefly,
299 wells of Immobilon-P polyvinylidene difluoride (PVDF) membrane 96-well plates (Millipore)
300 were coated with 5ug/ml anti-Ig light-chain antibodies (Rockland Immunochemicals) overnight
301 at 4°C. Plates were washed and wells were blocked at RT for 2 hours with RPMI containing 10%
302 FBS. Duplicate wells were plated with PBMC containing 0.01-0.003 X 10⁶ B cells for total
303 IgA/G/M-secreting PB enumeration and 0.1-0.03 X 10⁶ B cells for RBD/S1-specific PB. Plates
304 were incubated at 37°C for 5 hours, followed by overnight incubation at 4°C with biotinylated
305 antibodies (Jackson ImmunoResearch) against IgA (catalogue 109-066-011), IgG (catalogue 709-

306 066-149), IgM (catalogue 709-066-073), or biotinylated proteins S1 (Acrobiosystems, catalogue
307 S1N-C82E8) or RBD (Biolegend, catalogue 790904). Plates were washed, streptavidin-AP
308 conjugate (R&D Systems) was added, followed by incubation at RT for 2 hours. Plates were
309 washed and spots were developed with ELISpot Blue Color Module (R&D Systems).
310 Biotinylated (Biorbyt) or unlabeled (Millipore-Sigma) keyhole limpet hemocyanin (KLH) was
311 used as negative control antigen to enumerate background spots. Spots were counted using an
312 ELISPOT reader (Cellular Technology Ltd). Frequencies of S1 and RBD-specific PB were
313 calculated as the fraction of total Ig-secreting PB after subtraction of background KLH spots.

314

315 **Spectral flow cytometry data processing for FlowSOM clustering and UMAP embedding**

316 FCS files generated from spectral flow cytometry with the 17-color panel and associated
317 manual gates were read into R using FlowWorkspace (4.2.0). Live, CD19⁺ single cells were
318 selected for downstream analysis. Data from all samples were merged and transformed with the
319 arcsinh transformation with a scale factor of 1/150. Data were not z-score scaled. The following
320 markers were used to perform clustering CD20, CD138, CD38, CD10, CD11c, CD19, CD27,
321 CD21, IgD, IgM, IgG, IgA, using FlowSOM (1.22.0)³³, with the number of desired metaclusters
322 (nClus) set to 30. One small cluster, C8, determined to represent granulocytes, was removed
323 from downstream analysis. To visualize the clusters and RBD⁺S1⁺ cells in a UMAP
324 embedding³⁴, 653,683 cells were subsampled from the roughly 3.2 million CD19⁺ cells to
325 include 3,667 cells per sample and all RBD⁺S1⁺ cells from the 21 longitudinal vaccine

326 participants at all timepoints. Data were transformed as described above and the UMAP
327 embedding was fit using the uwot package (0.1.10) with default parameters.

328

329 **Analysis of temporally varying FlowSOM clusters**

330 Frequencies of cells within each sample were summarized into 1) the fraction of cells
331 within a cluster relative to the total number of CD19⁺ cells in the given sample; and, 2) the
332 fraction of cells within a cluster determined to be RBD⁺S1⁺ by manual gating, relative to the
333 number of cells in that cluster in the given sample. The extent of temporal variation was assessed
334 using a linear mixed effects model with the following formula in lme4 (1.1.26):

335 $\text{Frequency} \sim \text{timepoint} + (1|\text{Subject_ID})$

336 Timepoint is a factor variable representing the discrete timepoints (v1D0, v1D7, etc.)

337 The significance of the timepoint term was assessed with a type III ANOVA using
338 Satterthwaite's approximation using the lmerTest package (3.1.3)³⁵. P values were adjusted
339 across all comparisons using the Benjamini-Hochberg procedure³⁶. Clusters with adjusted p
340 values below 0.05 were deemed temporally fluctuating. Clusters selected by the above procedure
341 were then grouped by the similarity of their temporal patterns. Briefly, the mean frequency
342 across all subjects at a given timepoint was computed along with the 95% bootstrap confidence
343 intervals around the mean using package Hmisc (4.5.0). The means were then rescaled by
344 dividing all values by the maximum value of the 95% confidence interval throughout the time-
345 course such that all values were now in the range of [0,1]. The clusters were then grouped by
346 hierarchical clustering of the mean trends using the Euclidean distance at each timepoint and

347 using Ward’s method, as implemented in the hclust function (method = “ward.D2) in R (4.0.2).
348 After inspecting the respective dendrograms, 4 groups were determined to be appropriate, and
349 the hierarchical clustering trees were cut to produce 4 groups for both antigen non-specific and
350 specific cells.

351

352 **Modeling of association between endpoint antibody concentrations and cluster frequencies**

353 A linear model accounting for the age and gender of the longitudinal vaccine participants
354 was used to estimate whether cell cluster frequencies in response to vaccination were associated
355 with SARS-COV2 spike protein (S-2P/RBD) antibody concentration at endpoint (v2D28):

$$356 \log_2(\text{endpoint concentration}) \sim \text{cluster frequency}^{\text{timepoint}} + \text{age} + \text{sex}$$

357 Analyses were carried out on both standardized antigen non-specific and specific
358 frequencies, i.e., cluster cell counts as a fraction of total CD19⁺ cell counts and RBD⁺S1⁺ cells
359 within the clusters, respectively. For the antigen non-specific models, only clusters whose post-
360 vaccination frequencies at any timepoint changed significantly from the pre-vaccination baseline
361 (v1D0) were included. For the antigen-specific models, clusters with at least four RBD⁺S1⁺ cells
362 in any of the samples were considered. In addition, at each timepoint, a cluster was excluded if
363 there were fewer than five samples with any RBD⁺S1⁺ cells. Pre-vaccination baseline
364 frequencies (v1D0) were subtracted from frequencies post vaccine dose 1 (v1) timepoints,
365 including v1D7, v1D10, v1D14, and v2D0. Similarly, dose 2 baseline frequencies (v2D0) were
366 subtracted from frequencies of post-vaccine dose 2 timepoints, including v2D7, v1D10, v1D14,

367 and v2D28. P values were adjusted with the Benjamini-Hochberg method within each
368 combination of timepoint and antibody endpoint³⁶. R version 3.6.3 was used for this analysis.

369

370 **Correlation between principal component of endpoint antibody concentrations and cluster** 371 **frequencies**

372 In addition to modeling the association of the cluster frequencies to individual antibody
373 concentrations, their relationship to the primary correlated component of the two antibodies to
374 both S-2P and RBD proteins was also assessed. PC1 from principal component analysis of the
375 four endpoints (in log₂ scale), i.e., S-2P IgA/G and RBD IgA/IgG explained 83.6% of variance
376 across subjects. Associations between PC1 and RBD⁺S1⁺ cluster frequencies at each timepoint
377 were calculated using the same linear models and inclusion criteria described above. The same
378 analysis was carried out for all antigen non-specific clusters (cell counts as a fraction of total
379 CD19⁺ B cells) with the baseline pre-vaccination timepoint (v1D0) included.

380

381 **Acknowledgements**

382 The authors thank all the participants for their willingness to take part in our study;
383 Catherine Rehm and Ulisses Santamaria for coordinating the collection and distribution of
384 clinical samples for this study; NIAID OCICB for high-performance computing (HPC) support.
385 This work was funded by the Intramural Research Program of the NIAID of the NIH and NIH
386 extramural grant R01AI102766 (YL). YL was also supported by the University of Maryland
387 Strategic Partnership (MPower).

388

389 **Author contributions**

390 J.S.T., A.S.F., L.K., N.R., W.W.L., A.M. and S.M. designed research and methodology;
391 L.K., N.R., W.W.L., C.M.B., S.A., S.O., S.N., K.T., F.L.D.A., W.W., X.Z. and T.W.C.
392 performed experiments; Y.W., C.I.C., Y.L. and A.M. provided reagents and resources; L.K.,
393 N.R., W.W.L., S.O., C.M.B., T.W.C., J.S.T. and S.M. analyzed the data; N.R. and W.W.L. wrote
394 the software; R.R., G.E.M., C.A.S., R.W.C., A.F.S., J.R.S., D.S.C., R.T.D. and M.C.S.
395 contributed to recruitment of study participants; L.K., N.R., W.W.L., C.M.B., T.W.C., A.S.F.,
396 J.S.T. and S.M. wrote the manuscript.

397

398 **Competing Interests Statement**

399 YM is founder/CEO of ReVacc, Inc.

400

401 **Data availability**

402 All data are available in the main text, extended data and excel tables.

403

404 **Code availability**

405 All codes will be made available upon publication without restrictions.

406

407 **References**

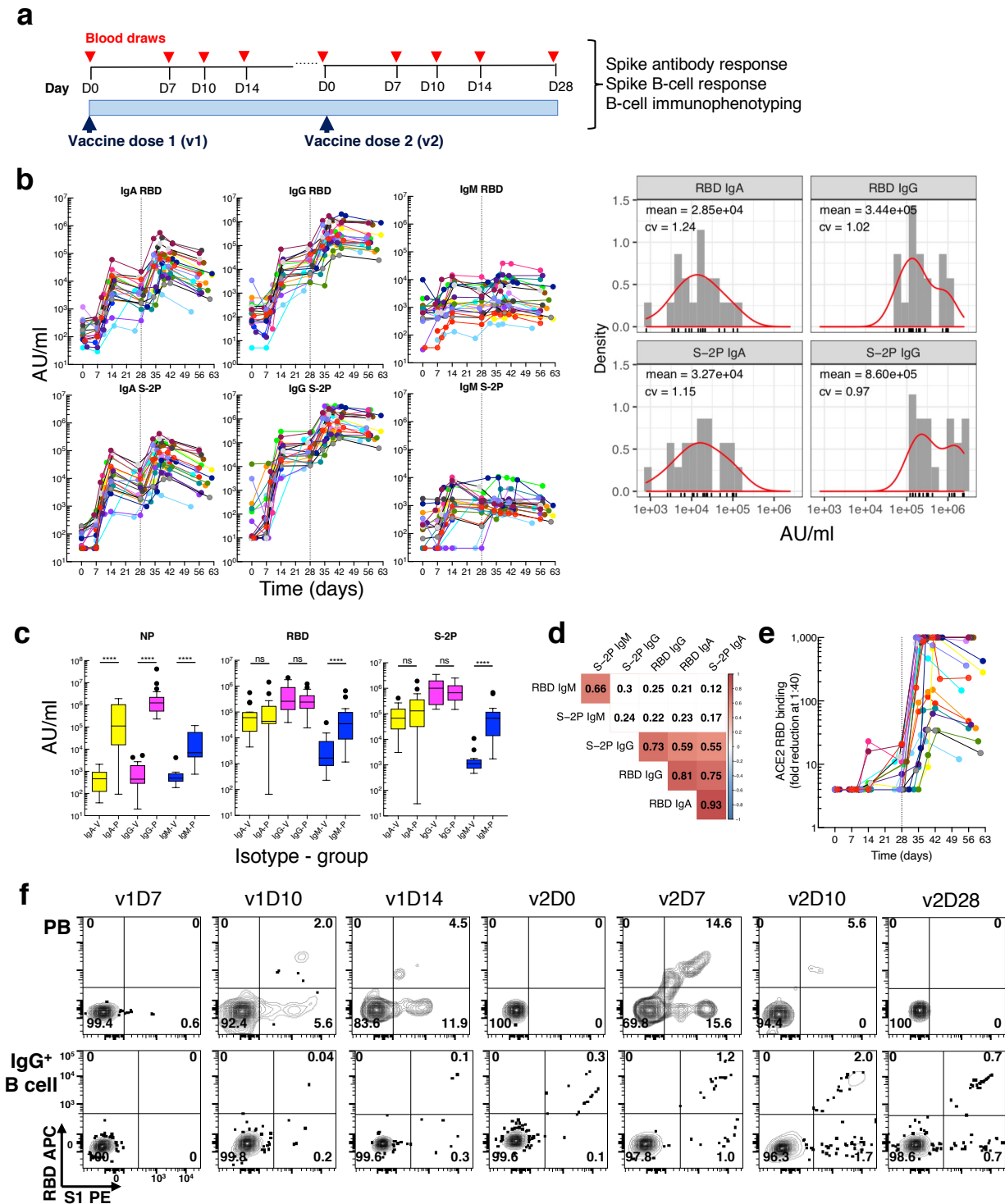
- 408 1. Wrapp, D., *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion
409 conformation. *Science* **367**, 1260-1263 (2020).
- 410 2. Polack, F.P., *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N*
411 *Engl J Med* **383**, 2603-2615 (2020).
- 412 3. Baden, L.R., *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N*
413 *Engl J Med* **384**, 403-416 (2021).
- 414 4. Goel, R.R., *et al.* Distinct antibody and memory B cell responses in SARS-CoV-2 naive
415 and recovered individuals following mRNA vaccination. *Sci Immunol* **6** (2021).
- 416 5. Tsapepas, D., Paget, K., Mohan, S., Cohen, D.J. & Husain, S.A. Clinically Significant
417 COVID-19 Following SARS-CoV-2 Vaccination in Kidney Transplant Recipients. *Am J*
418 *Kidney Dis* (2021).
- 419 6. Boyarsky, B.J., *et al.* Antibody Response to 2-Dose SARS-CoV-2 mRNA Vaccine Series
420 in Solid Organ Transplant Recipients. *JAMA* **325**, 2204-2206 (2021).
- 421 7. Husain, S.A., *et al.* Post-vaccine anti-SARS-CoV-2 spike protein antibody development
422 in kidney transplants recipients. *Kidney Int Rep* **6**(2021).
- 423 8. Kotliarov, Y., *et al.* Broad immune activation underlies shared set point signatures for
424 vaccine responsiveness in healthy individuals and disease activity in patients with lupus.
425 *Nat Med* **26**, 618-629 (2020).

- 426 9. Li, G.M., *et al.* Pandemic H1N1 influenza vaccine induces a recall response in humans
427 that favors broadly cross-reactive memory B cells. *Proc Natl Acad Sci U S A* **109**, 9047-
428 9052 (2012).
- 429 10. Wrammert, J., *et al.* Rapid cloning of high-affinity human monoclonal antibodies against
430 influenza virus. *Nature* **453**, 667-671 (2008).
- 431 11. Tsang, J.S. Utilizing population variation, vaccination, and systems biology to study
432 human immunology. *Trends Immunol* **36**, 479-493 (2015).
- 433 12. Turner, J., *et al.* SARS-CoV-2 mRNA vaccines induce robust plasmablast and germinal
434 centre responses in humans. *Nature Portfolio* (2021).
- 435 13. Danese, E., *et al.* Comprehensive assessment of humoral response after Pfizer BNT162b2
436 mRNA Covid-19 vaccination: a three-case series. *Clin Chem Lab Med* (2021).
- 437 14. Roltgen, K., *et al.* mRNA vaccination compared to infection elicits an IgG-predominant
438 response with greater SARS-CoV-2 specificity and similar decrease in variant spike
439 recognition. *medRxiv* (2021).
- 440 15. Gaebler, C., *et al.* Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639-644
441 (2021).
- 442 16. Sette, A. & Crotty, S. Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell* **184**,
443 861-880 (2021).
- 444 17. Rodda, L.B., *et al.* Functional SARS-CoV-2-Specific Immune Memory Persists after
445 Mild COVID-19. *Cell* **184**, 169-183 e117 (2021).

- 446 18. Elsner, R.A. & Shlomchik, M.J. Germinal Center and Extrafollicular B Cell Responses in
447 Vaccination, Immunity, and Autoimmunity. *Immunity* **53**, 1136-1150 (2020).
- 448 19. Stamatatos, L., *et al.* mRNA vaccination boosts cross-variant neutralizing antibodies
449 elicited by SARS-CoV-2 infection. *Science* (2021).
- 450 20. Dan, J.M., *et al.* Immunological memory to SARS-CoV-2 assessed for up to 8 months
451 after infection. *Science* **371** (2021).
- 452 21. Sokal, A., *et al.* Maturation and persistence of the anti-SARS-CoV-2 memory B cell
453 response. *Cell* **184**, 1201-1213 e1214 (2021).
- 454 22. Glass, D.R., *et al.* An Integrated Multi-omic Single-Cell Atlas of Human B Cell Identity.
455 *Immunity* **53**, 217-232 e215 (2020).
- 456 23. Baumgarth, N. The Shaping of a B Cell Pool Maximally Responsive to Infections. *Annu*
457 *Rev Immunol* **39**, 103-129 (2021).
- 458 24. Ellebedy, A.H., *et al.* Defining antigen-specific plasmablast and memory B cell subsets in
459 human blood after viral infection or vaccination. *Nat Immunol* **17**, 1226-1234 (2016).
- 460 25. Tsang, J.S., *et al.* Global analyses of human immune variation reveal baseline predictors
461 of postvaccination responses. *Cell* **157**, 499-513 (2014).
- 462 26. Arunachalam, P.S., *et al.* Systems biological assessment of human immunity to
463 BNT162b2 mRNA vaccination. *Nature Portfolio* (2021).
- 464 27. Berkowska, M.A., *et al.* Human memory B cells originate from three distinct germinal
465 center-dependent and -independent maturation pathways. *Blood* **118**, 2150-2158 (2011).

- 466 28. Andrews, S.F., *et al.* Activation Dynamics and Immunoglobulin Evolution of Pre-
467 existing and Newly Generated Human Memory B cell Responses to Influenza
468 Hemagglutinin. *Immunity* **51**, 398-410 e395 (2019).
- 469 29. Horns, F., *et al.* Lineage tracing of human B cells reveals the in vivo landscape of human
470 antibody class switching. *Elife* **5**(2016).
- 471 30. Gu, M., *et al.* One dose of COVID-19 nanoparticle vaccine REVC-128 provides
472 protection against SARS-CoV-2 challenge at two weeks post immunization. *bioRxiv*,
473 2021.2004.2002.438218 (2021).
- 474 31. Buckner, C.M., Kardava, L. & Moir, S. Evaluation of B cell function in patients with
475 HIV. *Curr Protoc Immunol* **Chapter 12**, Unit 12.13 (2013).
- 476 32. Buckner, C.M., *et al.* Maintenance of HIV-Specific Memory B-Cell Responses in Elite
477 Controllers Despite Low Viral Burdens. *J Infect Dis* **214**, 390-398 (2016).
- 478 33. Van Gassen, S., *et al.* FlowSOM: Using self-organizing maps for visualization and
479 interpretation of cytometry data. *Cytometry A* **87**, 636-645 (2015).
- 480 34. Becht, E., *et al.* Dimensionality reduction for visualizing single-cell data using UMAP.
481 *Nat Biotechnol* (2018).
- 482 35. Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. lmerTest Package: Tests in
483 Linear Mixed Effects Models. *2017* **82**, 26 (2017).
- 484 36. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate - a Practical and
485 Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* **57**, 289-300 (1995).

- 486 37. Moir, S. & Fauci, A.S. B-cell responses to HIV infection. *Immunol Rev* **275**, 33-48
487 (2017).
- 488 38. Portugal, S., Obeng-Adjei, N., Moir, S., Crompton, P.D. & Pierce, S.K. Atypical memory
489 B cells in human chronic infectious diseases: An interim report. *Cell Immunol* **321**, 18-25
490 (2017).
- 491 39. Jenks, S.A., Cashman, K.S., Woodruff, M.C., Lee, F.E. & Sanz, I. Extrafollicular
492 responses in humans and SLE. *Immunol Rev* **288**, 136-148 (2019).
- 493

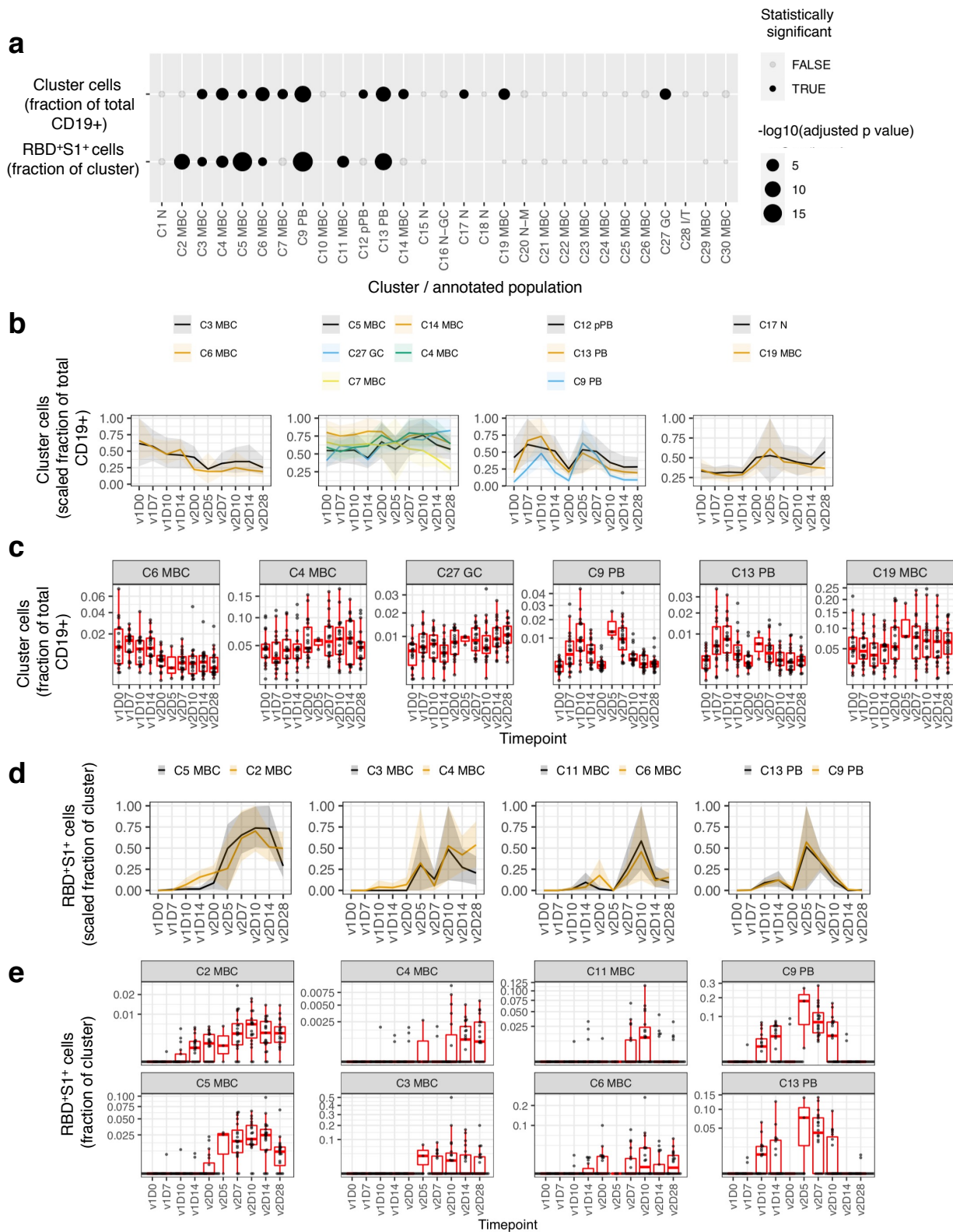


495 **Fig. 1. Longitudinal blood sampling and analysis shows robust antibody and early B cell**
496 **response to mRNA-1273 vaccine. a,** Study design with serial blood draws and assays
497 performed at all timepoints on SARS-CoV-2-uninfected vaccinees ($n = 21$; missed visits in
498 Extended Data Table 1) receiving two doses of the mRNA-1273 vaccine. **b,** Serum IgG, IgA and
499 IgM binding to S-2P and RBD proteins measured by electrochemiluminescence (ECLIA)
500 longitudinally (left panels), and corresponding histogram and distribution (based on kernel
501 density estimates) at the last timepoint (v2D28) (right panels). **c,** Peak serum IgG, IgA and IgM
502 binding to S-2P, RBD and N proteins measured by ECLIA in vaccinees (V; $n = 21$) and COVID-
503 19 patients (P; $n = 21$), shown as boxplots. **d,** Triangular heatmap of correlation between serum
504 antibodies at last measured timepoint (v2D28) in **(b)**. Numbers represent r values. Statistically
505 insignificant correlations ($p > 0.05$) shown in white. **e,** Longitudinal inhibition of RBD binding
506 to ACE2 by serum (1:40 dilution) of vaccinees ($n = 21$). **f,** Longitudinal binding of S1 and RBD
507 tetramers to PB and IgG⁺ B cells by flow cytometry shown for a high responder (VAC-611;
508 Extended Data Table 1). Numbers in each quadrant are percentages. Each vaccinee is color-
509 coded and second vaccine dose indicated by vertical dotted line **(b,e)**. Mann-Whitney test; ****,
510 $p < 0.0001$ **(c)**. Spearman's rank correlation **(d)**. AU, arbitrary units; D, day; N, nucleocapsid; ns,
511 not significant; P, patients with severe COVID-19; PB, plasmablasts; RBD, receptor binding
512 domain; S1, spike subunit 1; S-2P, stabilized spike trimer; v, vaccine dose; V, vaccinees.
513

515 **Fig. 2. Unsupervised clustering analysis identifies major B cell populations and SARS-CoV-**
516 **2-specific B cells. a**, UMAP projection of combined B cells ($n = 653,683$ cells), subsampled
517 from 3.2 million CD19⁺ cells to include 3,667 cells per sample and all RBD⁺S1⁺ cells from all
518 study participants ($n = 21$) at all timepoints with annotated major B cell populations identified by
519 FlowSOM clustering. **b**, MFI-based heatmap of FlowSOM clusters as indicated by cluster
520 number and marker. Rows ordered by hierarchical clustering. Summary of fraction of cells
521 binding both RBD and S1 within each cluster and cell counts per cluster (right). **c**, UMAP plots
522 with overlays of RBD⁺S1⁺ B cells (blue points with white center) at each timepoint. **d**, RBD⁺S1⁺
523 cells within each cluster expressed as a fraction of total CD19⁺ B cells across all subjects at each
524 timepoint (n at each timepoint shown in Extended Data Table 1). D, day; GC, germinal center;
525 I/T, immature transitional; MBC, memory B cell; MFI, mean fluorescence intensity; N, naïve;
526 PB, plasmablast; pPB, pre-plasmablast; v, vaccine dose.

527

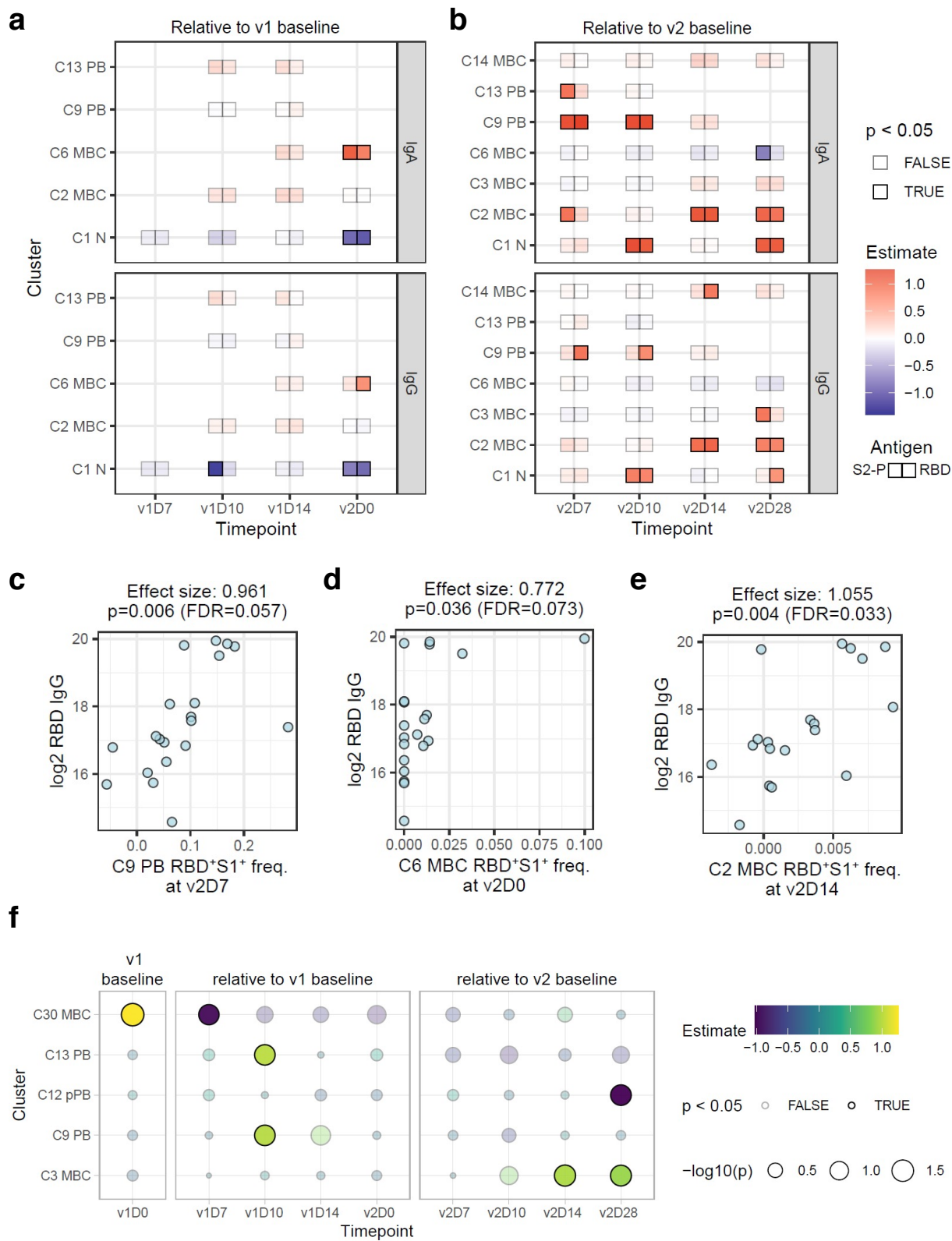
528



530 **Fig. 3. Antigen non-specific and spike-specific cells exhibit temporal change in response to**
531 **the mRNA-1273 vaccine. a,** Clusters showing significant temporal variation over course of v1
532 and v2 in the frequency of non-specific cells as a fraction of total (CD19⁺) B cells (first row) and
533 RBD⁺S1⁺ cells as a fraction within each cluster (second row). **b,** Longitudinal display of non-
534 specific cells per cluster as a fraction of total (CD19⁺) B cells, shown for clusters with
535 statistically significant temporal variations, as shown in **(a)**. Clusters were grouped by temporal
536 patterns (see methods). Lines denote the mean and shading denotes 95% bootstrap confidence
537 interval per timepoint. Values rescaled as fraction of maximum 95% confidence interval estimate
538 over the entire time course. **c,** Similar to **(b)** but showing boxplots for selected clusters. **d,** similar
539 to **(b)** but displaying of RBD⁺S1⁺ cells as a fraction within each cluster. **e,** Similar to **(d)** but
540 showing boxplots. Type III ANOVA test using Satterthwaite's approximation **(a)**. N at each
541 timepoint shown in Extended Data Table 1 **(b-e)**. D, day; RBD, receptor binding domain; S1,
542 spike subunit 1; v, vaccine dose.

543

544



546 **Fig. 4. Correlates of SARS-CoV-2 antibody titers 28 days after second dose of vaccine. a,b,**
547 Linear model effect size estimates indicate strength of association between spike-specific
548 (RBD⁺S1⁺) cell frequency in the cell clusters (rows) with antibody endpoints (IgA and IgG titers
549 for S-2P and RBD) relative to pre-vaccination baseline level (v1D0) at four timepoints between
550 v1 and v2 **(a)**, and relative to v2 baseline (v2D0) at four timepoints after v2 **(b)**. Only clusters
551 with at least one significant (unadjusted $p \leq 0.05$) association at any timepoint are shown. At
552 each timepoint, clusters that had fewer than five samples with any RBD⁺S1⁺ cells were excluded
553 from analysis (missing boxes). **c-e**, Scatter plots illustrating correlations between endpoint
554 (v2D28) RBD IgG titers and RBD⁺S1⁺ cell frequencies in C9 on v2D7, C6 on v2D0, and C2 on
555 v2D14, respectively. Effect sizes and p values were estimated by the linear models above. FDR
556 estimate of the statistical significance was calculated within each antibody endpoint and
557 timepoint combination. **f**, Effect size estimates of association between first principal component
558 (PC1) of endpoint SARS-CoV-2 antibody titers and frequencies of each cell cluster as a fraction
559 of total CD19⁺ cells. PC1 was derived from IgA and IgG titers against S-2P and RBD proteins at
560 v2D28. Only cell clusters with at least one significant (unadjusted $p \leq 0.05$) association at any
561 timepoint are shown. D, day; FDR, false discovery rate; RBD, receptor binding domain; rho,
562 Spearman's rank correlation; S1, spike subunit 1; S-2P, stabilized spike trimer; v, vaccine dose.