nature immunology

LETTERS

https://doi.org/10.1038/s41590-021-01088-9

Check for updates

BNT162b2 vaccine induces divergent B cell responses to SARS-CoV-2 S1 and S2

R. Camille Brewer[1,2], Nitya S. Ramadoss[1,2], Lauren J. Lahey[3], Shaghayegh Jahanbani[1,2], William H. Robinson[1,2,5] ✉ and Tobias V. Lanz[1,2,4,5] ✉

The first ever US Food and Drug Administration-approved messenger RNA vaccines are highly protective against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)[1–3]. However, the contribution of each dose to the generation of antibodies against SARS-CoV-2 spike (S) protein and the degree of protection against novel variants warrant further study. Here, we investigated the B cell response to the BNT162b2 vaccine by integrating B cell repertoire analysis with single-cell transcriptomics pre- and post-vaccination. The first vaccine dose elicits a recall response of IgA$^+$ plasmablasts targeting the S subunit S2. Three weeks after the first dose, we observed an influx of minimally mutated IgG$^+$ memory B cells that targeted the receptor binding domain on the S subunit S1 and likely developed from the naive B cell pool. This response was strongly boosted by the second dose and delivers potently neutralizing antibodies against SARS-CoV-2 and several of its variants.

BNT162b2 is one of the two first vaccines that are based on lipid nanoparticle delivery of modified mRNA and are dependent on the host cells for translation and expression of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike (S) protein, which consists of subunits S1 and S2 (ref. [4]). S1 contains the receptor binding domain (RBD), which binds the host entry receptor angiotensin-converting enzyme 2 (ACE2) and initiates viral cell entry, whereas S2 mediates virus–cell membrane fusion[5,6]. RBD is the target of most neutralizing antibodies found in patients with coronavirus disease 2019 (COVID-19)[7]. The cellular processes that generate potent neutralizing antibodies in response to mRNA vaccines are not fully characterized and to what degree these antibodies protect against new variants remains unclear.

Nine healthy individuals without previous SARS-CoV-2 infection were included in this study (Supplementary Table 1). One individual contracted COVID-19 8 weeks after the second dose. Peripheral blood B cells were investigated by droplet-based single-cell sequencing before vaccination (day 0), as well as 7–9 days (day 7), 21–23 days (day 21) and 28 days (day 28) after the first dose (Fig. 1a). The second dose was given on day 21. Additionally, SARS-CoV-2 S-specific B cells were labeled with S1, S2 and RBD tetramers conjugated to fluorochromes and DNA barcodes and sorted by fluorescence-activated cell sorting (FACS) before sequencing (Fig. 1a). A total of 131,138 B cells were included in the global transcriptomic analysis. Dimensionality reduction by uniform manifold approximation and projection (UMAP)[8] and graph-based clustering distinguished nine B cell populations present in all individuals at all time points, including naive B cells (B_{naive} cells), unswitched and switched memory B cells (U-B_{mem} cells and S-B_{mem} cells, respectively) and plasmablast clusters (Fig. 1b, Extended Data Fig. 1a–c and Supplementary Table 2). B_{naive} cells were further sub-categorized into *CCR7*^{lo}*PTRCAP*^{hi}*JUNB*^{hi} B_{naive} cells (B_{naive}1 cells) and *CCR7*^{hi}*PTRCAP*^{lo}*JUNB*^{lo} B_{naive} cells (B_{naive}2 cells) (Extended Data Fig. 1b). S-B_{mem} cells were divided into *PTPN6*^{lo}*BLK*^{lo}*CD86*^{lo} resting switched B_{mem} cells (S_R-B_{mem} cells) and *PTPN6*^{hi}*BLK*^{hi}*CD86*^{hi} activated switched B_{mem} cells (S_A-B_{mem} cells) (Extended Data Fig. 1b)[9]. Fitting to the cluster assignments, B_{naive} cells mainly expressed low-mutation IgM, whereas B_{mem} cells and plasmablasts expressed other isotypes with higher frequencies of somatic hypermutation (SHM) (Extended Data Fig. 1d,e). Single-cell transcriptome sequencing on days 0, 7, 21 and 28 showed decreased frequencies of U-B_{mem} cells (day 0, 35.46%; day 7, 13.88%) and increased frequencies of S_R-B_{mem} cells (day 0, 9.38%; day 7, 23.42%) on day 7, which was sustained until day 28 (Fig. 1c,d). To understand the fate of U-B_{mem} cells, we tracked them from day 0 to day 28 using B cell receptor (BCR) sequencing. Clonally related BCR sequences were defined by shared heavy chain and light-chain variable genes and >70% overlap in both CDR3 regions. B cells clonally related to day 0 U-B_{mem} cells were identified in day 7, 21 and 28 datasets (Fig. 2a). The majority of clonal day 0 U-B_{mem} cells differentiated into S_R-B_{mem} cells on days 7–28 (day 7, 46.91%; day 21, 63.21%; day 28, 66.34%) (Fig. 2a,b and Extended Data Fig. 2a,b). SHM frequencies of day 0 U-B_{mem} cells did not differ from clonally related days 7–28 S_R-B_{mem} cells (Extended Data Fig. 2c), suggesting differentiation without germinal center (GC) maturation. Only 1.98% of B cells clonally related to day 0 U-B_{mem} cells developed into plasmablasts on day 28 (Fig. 2a,b and Extended Data Fig. 2a,b), suggesting that the differentiation of day 0 U-B_{mem} cells was separate from the plasmablast response to SARS-CoV-2. Pathway enrichment analysis revealed increased expression of B cell activation genes in U-B_{mem} cells between day 0 and days 7–28 (Fig. 2b and Extended Data Fig. 2f,j). In contrast, the same genes were downregulated in S_R-B_{mem} cells between day 0 and days 7–28 (Fig. 2c and Extended Data Fig. 2g), although upregulation of CD83 and CD69 in S_R-B_{mem} cells indicated that they were recently activated (Extended Data Fig. 2k)[10,11]. Together, these observations suggest that BNT162b2 activated U-B_{mem} cells and induced class-switching and differentiation into S_R-B_{mem} cells independent of the GC.

U-B_{mem} cells are thought to develop independently of the GC and possess a polyreactive repertoire for rapid B cell responses[12]. The observed separation of clonal U-B_{mem} cells from the plasmablast

[1]Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA. [2]VA Palo Alto Health Care System, Palo Alto, CA, USA. [3]Biophysics Program, Stanford University, Stanford, CA, USA. [4]Department of Neurology, Mannheim Center for Translational Neurosciences (MCTN), Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany. [5]These authors contributed equally: William H. Robinson, Tobias V. Lanz. ✉e-mail: w.robinson@stanford.edu; tlanz@stanford.edu

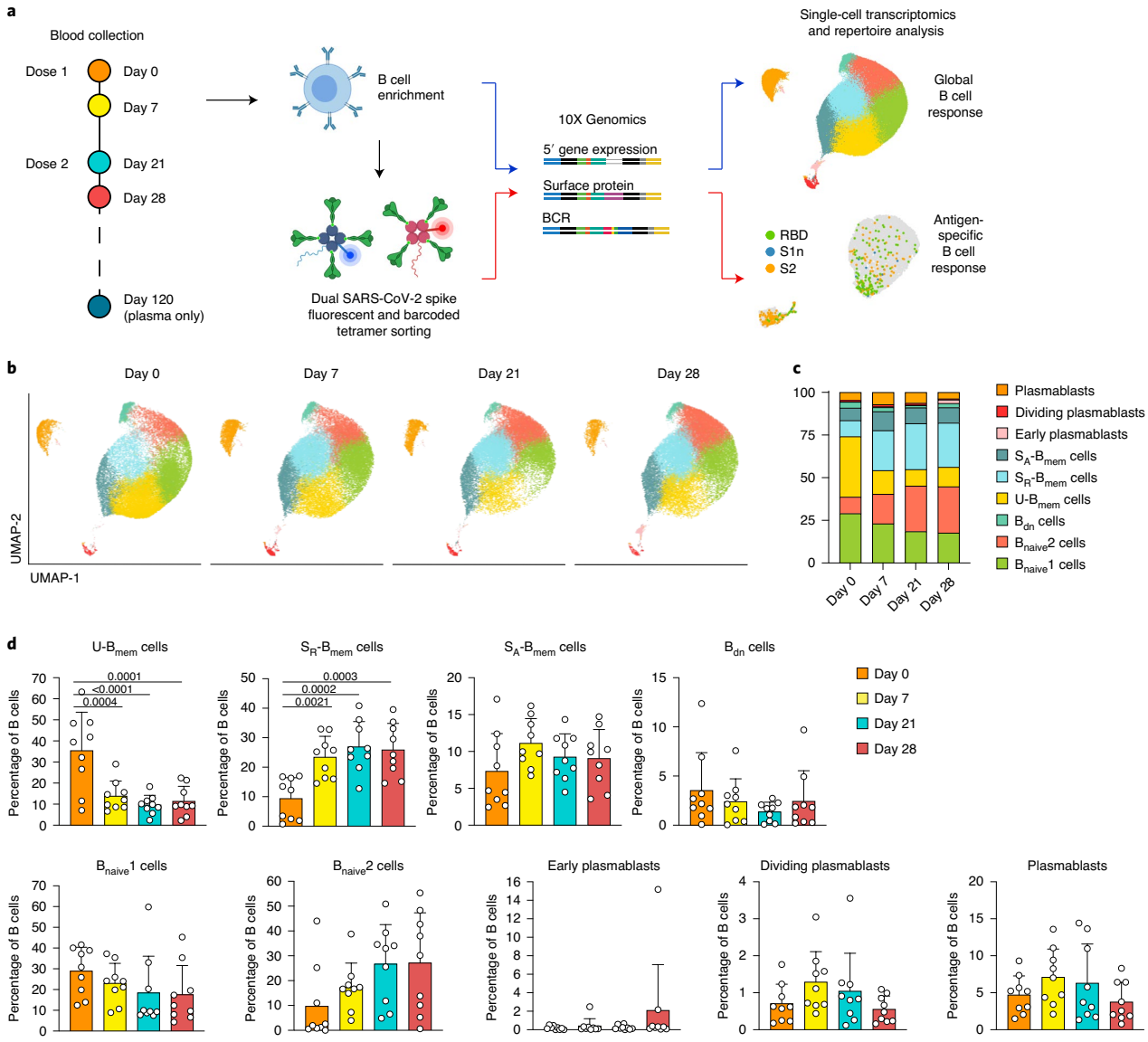


Fig. 1 | Single-cell transcriptomic analysis of B cell response to BNT162b2 vaccination. a, Outline of experimental approach. **b**–**d**, Single-cell transcriptome analysis. UMAP visualization of B cell clusters for all individuals at four time points (**b**). Cluster assignments are based on gene expression and cell surface expression (CITE-seq) of canonical B cell markers. Mean percentages of B cell clusters shown in **b** (**c**). Individual and mean percentages of B cell clusters in **b** at four time points (**d**). $n = 9$ individuals at all time points. Individual values, means and s.d. are shown. Exact P values were obtained by two-tailed one-way analysis of variance (ANOVA), followed by Dunnett's multiple comparison test. B_{dn} cells, double-negative B cells.

lineages indicates that U-B_{mem} cells were likely not a major source of S-antigen-specific antibodies. Plasmablasts are short-lived activated B cells that expand in response to antigen stimulation and secrete large amounts of antibodies[13]. To identify B cell populations directly involved in the antigen-specific response, we focused on clonally related B cell expansions that included plasmablasts at day 28. Based on shared BCR sequences, the majority of B cells clonally related to day 28 plasmablasts were found in the plasmablast compartment at days 0–21 (day 0, 53.57%; day 7, 52.75%; day 21, 66.67%) (Fig. 2d and Extended Data Fig. 2d,e). Considering the short life span of plasmablasts[14], this indicates continued recruitment from either B_{naive} cell or B_{mem} cell pools. Genes associated with leukocyte activation and protein processing were enriched in plasmablasts on days 7–28 over day 0 (Fig. 2e and Extended Data Fig. 2h), consistent with their function as antibody producers. Genes associated with SARS-CoV-2 infection were enriched in plasmablasts on day 21 over day 0 (Extended Data Fig. 2h), suggesting that recognition of viral proteins triggers similar pathways in vaccination and infection.

IgA^+ plasmablasts are prevalent in peripheral blood during health, while IgG^+ plasmablasts increase during systemic infection and vaccination[15,16]. Frequencies of IgG^+ plasmablast and in particular $IgG1^+$ plasmablast increased between days 0 and 7 and between days 21 and 28 in response to each dose (Fig. 2f and Extended Data Fig. 3a–c). The average SHM frequency in IgG^+ plasmablasts decreased from day 0 to day 28 (Fig. 2g and Extended Data Fig. 3d–k). These observations show an influx of minimally mutated IgG^+ plasmablasts developed in response to vaccination, likely originating from the B_{naive} cell pool.

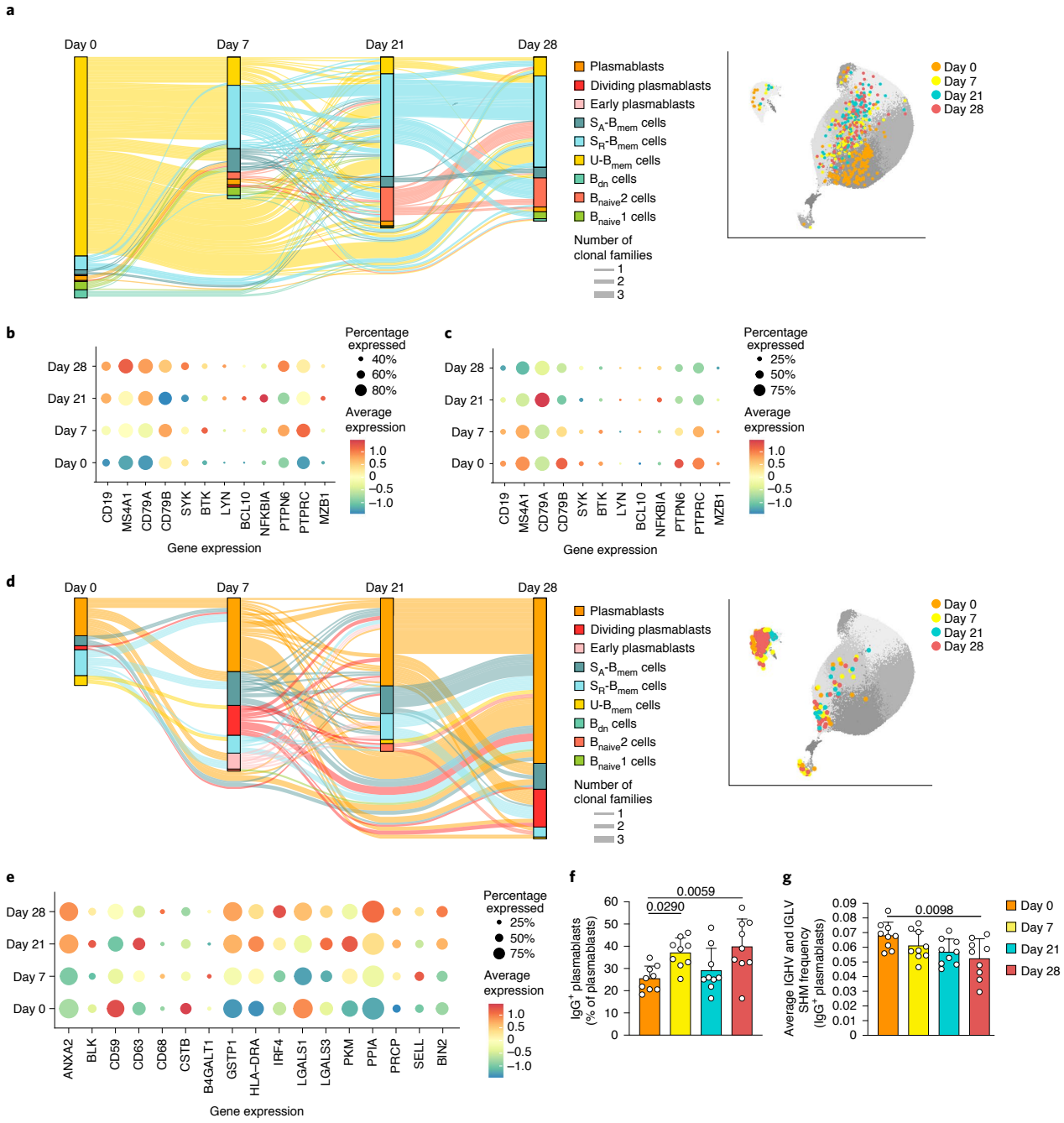


Fig. 2 | Unswitched memory B cell differentiation and expansion of IgG⁺ plasmablasts in response to vaccination. a, Single-cell BCR repertoire data, showing trajectories of B cell clonal families across time points as alluvial plot (left) and UMAP projection of clonal B cells (right) that are related to U-B_{mem} cells at day 0 of $n = 9$ participants. **b**,**c**, Single-cell transcriptomic sequencing data showing expression of genes involved in B cell activation and BCR signaling in U-B_{mem} cells (**b**) and S_R-B_{mem} cells (**c**). **d**, Single-cell BCR repertoire data, showing trajectories of B cell clonal families across time points as alluvial plot (left) and UMAP projection of clonal B cells (right) that are related to plasmablasts at day 28 of $n = 9$ participants. **e**–**g**, Single-cell transcriptomic sequencing data showing expression of genes involved in B cell activation in plasmablasts (**e**), percentage of IgG⁺ plasmablasts in all plasmablasts per individual ($n = 9$) at four time points (**f**) and average SHM frequency of IgG⁺ plasmablasts per individual ($n = 9$) at four time points (**g**). Clonal families included have ≥3 members at ≥2 time points (**a**,**d**). Line thickness represents the number of clonal families. Individual values, means and s.d. are shown (**f**,**g**). Exact *P* values obtained using two-tailed one-way ANOVA, followed by Dunnett's multiple comparison test. IGHV, immunoglobulin heavy V gene; IGLV, immunoglobulin light V gene.

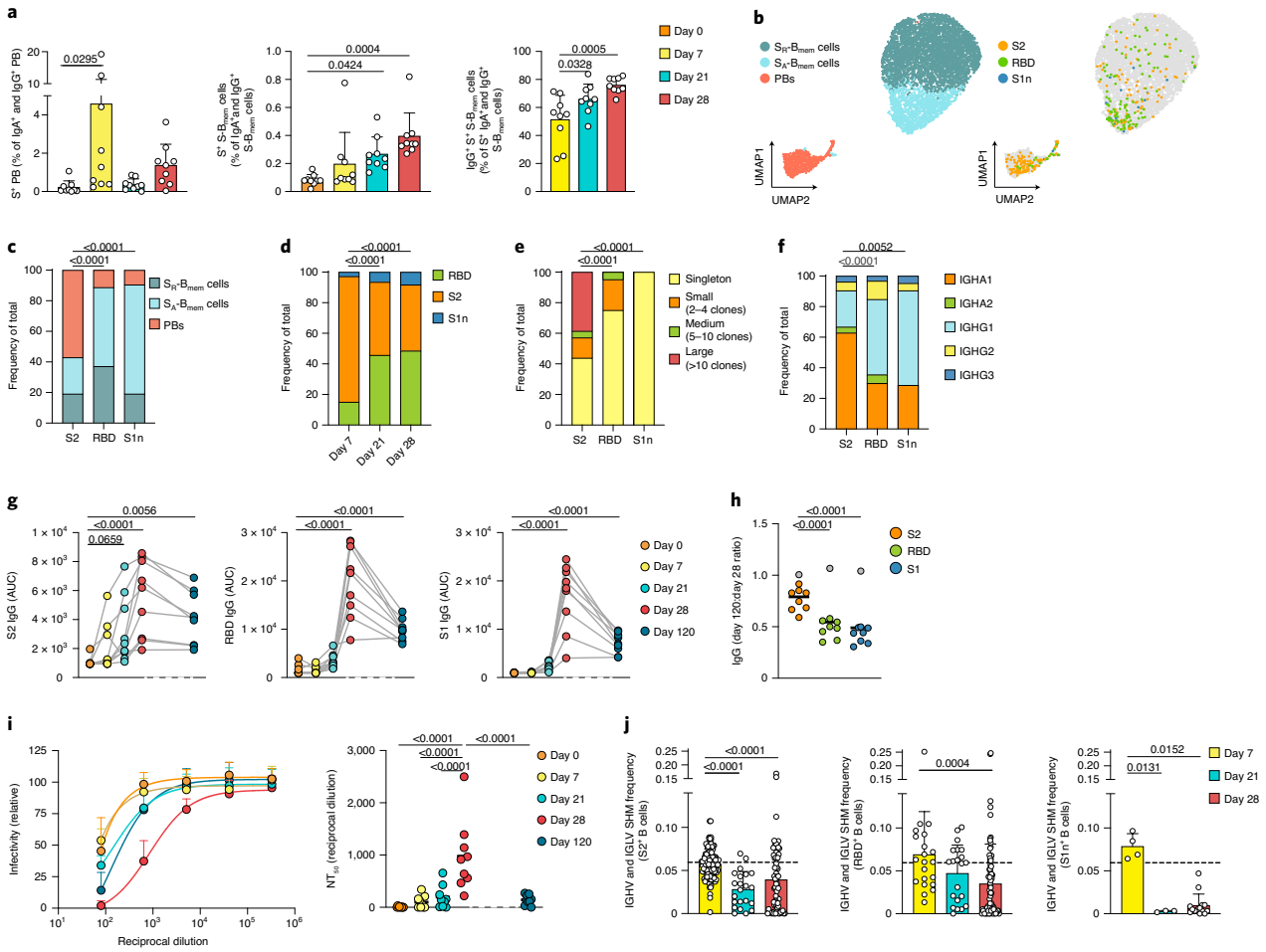


Fig. 3 | Vaccination induces an IgA$^+$ anti-S2 response on day 7 followed by an IgG$^+$ anti-RBD response on days 21 and 28. a, Flow cytometry data, showing percentages of S-specific plasmablasts in all IgG$^+$ and IgA$^+$ plasmablasts (left), S-specific S-B$_{mem}$ cells in all IgG$^+$ and IgA$^+$ S-B$_{mem}$ cells (center) and IgG$^+$ S-specific S-B$_{mem}$ cells in all S-specific IgG$^+$ and IgA$^+$ S-B$_{mem}$ cells (right). Individual data points are averages of two independent experiments, including $n = 9$ participants at four time points. **b**–**f**, Single-cell transcriptome and BCR repertoire sequencing data of S-specific and nonspecific sorted B cells showing UMAP visualization with cluster assignments of S$_R$-B$_{mem}$ cells, S$_A$-B$_{mem}$ cells and plasmablasts (**b**, left) and antigen-specificity to S2, RBD and S1n (**b**, right). Proportions of cells shown in **b**, separated by antigen and cluster (**c**), day and antigen (**d**), antigen and clonality (**e**) and antigen and immunoglobulin class (**f**). **g**, Plasma IgG levels against S2 (left), RBD (center) and S1 (right) for $n = 9$ individuals at five time points. Area under the curve (AUC) for plasma dilutions are shown as individual data points. **h**, Ratio of IgG levels of day 120 AUC to day 28 AUC for antigens S2, RBD and S1. Gray dots represents the vaccinated participant who contracted COVID-19 2 weeks before day 120 and is excluded from statistics. Individual data points and medians for $n = 9$ participants are shown. **i**, Plasma neutralization of SARS-CoV-2 pseudovirus shown as dilution curves (left) and quantification (NT50) of neutralization titers (right) of $n = 9$ individuals at five time points. **j**, Single-cell BCR repertoire sequencing data showing IGHV and IGLV gene SHM frequencies of B cells specific for S2 ($n = 210$ cells, left), RBD ($n = 124$ cells, center) and S1n ($n = 24$ cells, right) at three time points after vaccination. Dashed line indicates the average mutation frequency of sorted nonspecific B cells. Individual data points represent single B cells from $n = 9$ individuals. Individual values, means and s.d. are shown (**a**,**j**). Exact *P* values according to two-tailed one-way ANOVA, followed by Dunnett's multiple comparison test (**a**,**g**–**i**), chi-squared test (**c**–**f**), two-tailed Kruskal–Wallis test followed by Dunnett's multiple comparison test (**j**). PB, plasmablast.

To further characterize the S-specific B cell response to vaccination, we labeled B cells with fluorescent S1, S2 and RBD tetramers, each with a unique barcode and FACS-sorted antigen-tetramer-specific and nonspecific IgA$^+$, IgG$^+$ and IgM$^-$ plasmablasts and S-B$_{mem}$ cells (Extended Data Fig. 4a). Antigen specificities were determined by their barcode after demultiplexing (Extended Data Fig. 5a–e). S-specific IgA$^+$ and IgG$^+$ plasmablasts expanded at days 7 and 28 (day 0, 0.22%; day 7, 4.58%; day 28, 1.37%) (Fig. 3a and Extended Data Fig. 4b,c). S-specific IgA$^+$ and IgG$^+$ S-B$_{mem}$ cell response increased from day 0 to day 28 (day 0, 0.08%; day 28, 0.395%) (Fig. 3a and Extended Data Fig. 4b,e). After day 7, the S-specific B cell response shifted from an IgA$^+$ to an IgG$^+$ response (Fig. 3a and Extended Data Fig. 4d,f).

Antibody responses against distinct epitopes contribute differently to immune protection[17]. We investigated how the antibody response to each S subunit varies within B cell subsets and over time. UMAP analysis indicated that S2-specific B cells were predominantly plasmablasts and accounted for 82.14% of all S-specific B cells at day 7 (Fig. 3b–d). In contrast, the majority of B cells specific to RBD and specific to S1, but not RBD (S1n), were

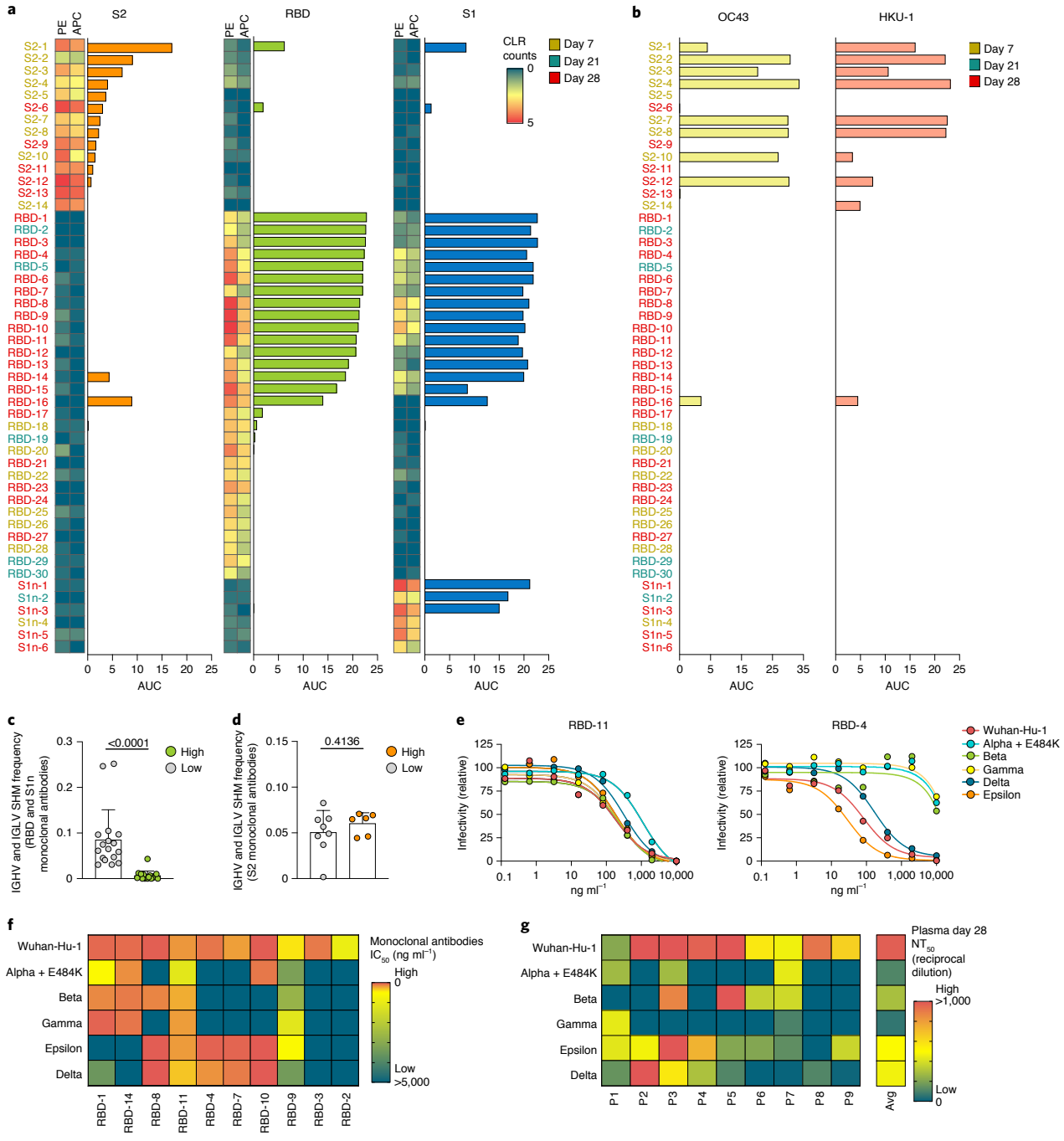


Fig. 4 | Neutralization of SARS-CoV-2 pseudovirus and variants by BNT162b2-induced antibodies. a,**b**, Single-cell sequencing data indicating abundance of barcoded and fluorochrome-tagged antigen tetramers for each sorted B cell (two-column heat maps, one barcode corresponding to PE and APC, respectively), paired with ELISA data of the corresponding expressed monoclonal antibodies (bar graphs). Reactivity against S2, RBD and S1 is shown. Barcoded tetramer binding data are represented as centered log ratio-transformed (CLR) counts. ELISA data showing reactivity of the same monoclonal antibodies as **a** against OC43 and HKU1 spike (**b**). ELISA data are shown as AUC of plasma dilutions. Threshold at 0 was set to the average binding to bovine serum albumin (BSA) plus its threefold s.d. **c**,**d**, Single-cell BCR repertoire sequencing data showing IGHV and IGLV SHM frequencies of RBD- and S1n-specific B cells (high, $n = 17$ cells; low, $n = 18$ cells) (**c**) and S2-specific B cells (high, $n = 8$ cells; low, $n = 6$ cells) (**d**). High-binding monoclonal antibodies are defined by data shown in **a** as positive for both PE and APC barcode tetramer binding and ELISA AUC > 3. Individual values, means and s.d. are shown. exact *P* values were obtained using unpaired two-tailed Student's *t*-test. **e**–**g**, Neutralization of SARS-CoV-2 pseudovirus Wuhan-Hu-1 and variants Alpha to Epsilon are shown as dilution curves for antibodies RBD-11 (**e**, left) and RBD-4 (**e**, right) and as heat maps indicating half-maximum inhibitory concentration (IC_{50}) (ng ml^{-1}) for each monoclonal antibody (**f**) and 50% neutralization titer (NT_{50}) (reciprocal dilution) of day 28 plasma for each of the nine participants (**g**).

S_A-B_{mem} cells (Fig. 3b,c and Extended Data Fig. 4g). RBD-specific and S1n-specific B cells increased from <20% on day 7 to >50% on day 28 of all S-specific B cells (Fig. 3b–d). The S2-specific response was highly clonal and dominated by IgA1$^+$ plasmablasts, whereas RBD-specific and S1n-specific S_A-B_{mem} cells used predominantly IgG1 and were less clonally expanded than S2 plasmablasts (Fig. 3e,f). The rapid onset of the S2-specific response and the delayed S1n- and RBD-specific response indicated that the S2-specific B cell response is a secondary response, whereas the S1n and RBD-specific B cell response is a primary response.

The early S2-specific plasmablast response corresponded with S2-specific plasma antibody titers, which rose quickly from day 0 to day 7 and plateaued around day 21. RBD-specific and S1-specific IgG and IgA levels were low until day 21 and increased at day 28 in response to the second dose (Fig. 3g and Extended Data Fig. 6a,b). Accordingly, day 28 plasma neutralized the SARS-CoV-2 Wuhan-Hu-1 pseudovirus more potently than day 0–21 plasma (Fig. 3i). Notably, by day 120, RBD-specific and S1-specific IgG levels decreased by 52%, whereas S2-specific IgG levels decreased only by 23.61% (excluding the participant who contracted COVID-19 2 weeks before day 120) (Fig. 3g,h). This suggests less-effective induction of RBD-specific long-lived plasma cells upon recruitment from B_{naive} cells.

The participant with the lowest RBD-specific IgG and IgA levels contracted COVID-19 8 weeks after the second vaccination (Extended Data Fig. 6a–d). While this participant produced high-affinity neutralizing B cells (Fig. 4f,g), antibody levels were not protective. In addition, the participant's antibody levels did not increase by day 120, 2 weeks after COVID-19 infection (Extended Data Fig. 6a–c), suggesting that the low antibody levels were not due to the effectiveness of the mRNA vaccine.

Neutralizing antibodies from patients with COVID-19 show characteristically low mutation rates[18], indicating recruitment of B_{naive} cells to the GC in response to a novel antigen. We found that on day 7, when the majority of antigen-specific B_{mem} cells and plasmablasts was S2-specific (Fig. 3d), mutation frequencies of antigen-specific B_{mem} cells and plasmablasts were similar to control-sorted B_{mem} cells and plasmablasts without specificity to S-antigens (Fig. 3j and Extended Data Fig. 7a–c). In contrast, mutation frequencies of antigen-specific B_{mem} cells and plasmablasts on days 21 and 28, when the majority of B cells reacted to S1 and RBD, were significantly lower than on day 7 (Fig. 3d,j and Extended Data Fig. 7a–c). Thus, the SHM analysis indicated that the S2-specific response was a recall response from B_{mem}, whereas the delayed S1n- and RBD-specific response was a primary response from B_{naive} cells.

To further characterize the primary and secondary B cell response, we expressed 50 representative BCR sequences from S-antigen-specific B cells; 14 S2-specific, 30 S1- and RBD-specific and 6 S1n-specific, as recombinant monoclonal antibodies (Supplementary Table 3). ELISA and biolayer interferometry indicated that of the selected monoclonal antibodies, 8 potently bound S2, 15 bound RBD and 3 bound S1n (Fig. 4a and Extended Data Fig. 8). When tested against S-protein of four other human-pathogenic coronaviruses (229E, NL63, OC43 and HKU1), 9 and 8 of the 12 S2-reactive monoclonal antibodies cross-reacted with betacoronaviruses OC43 and HKU1 S-protein, respectively, but none cross-reacted with the alphacoronaviruses 229E and NL63 (Fig. 4b). Of the RBD-specific monoclonal antibodies, only the polyreactive monoclonal antibody RBD-16 (Extended Data Fig. 9), showed cross-reactivity to OC43 and HKU1 S-protein (Fig. 4b). These observations are in line with a higher conservation of S2 than S1 across coronavirus species[19,20].

Monoclonal antibodies that bound strongly to S1 by ELISA and biolayer interferometry (high binders) harbored low SHM frequencies in comparison to monoclonal antibodies that bound S1-specific antigen tetramers but showed little or no reactivity to S-antigens by ELISA (low binders) (Fig. 4c and Extended Data Fig. 7d). In contrast, SHM frequencies in S2-specific monoclonal antibodies did not differ significantly between high and low binders (Fig. 4d and Extended Data Fig. 7e). Monoclonal antibodies with the highest affinity for S2 were derived from day 7 plasmablasts, whereas monoclonal antibodies with the highest affinity for S1 (including RBD) originated from S-B_{mem} cells from days 21 and 28 (Fig. 4a). While a correlation between increased affinity and low SHM frequency is counter intuitive, it is characteristic for the B cell repertoire in patients with COVID-19 (refs. [21,22]) and likely caused by the reduced structural overlap of the S1 protein to other pathogens, requiring recruitment from the B_{naive} cell pool.

The emergence of novel SARS-CoV-2 variants could jeopardize vaccine efficacy. RBD mutations contribute significantly to immune escape[23]. After testing neutralization efficacy of all RBD-specific monoclonal antibodies to the Wuhan-Hu-1 pseudovirus (Extended Data Fig. 10), we selected the ten best Wuhan-Hu-1-neutralizing RBD-specific monoclonal antibodies and tested their neutralization of five SARS-CoV-2 variants, Alpha+E484K (B.1.1.7+E484K; N501Y, E484K), Beta (B.1.351; N501Y, K417N), Gamma (P.1; N501Y, K417T), Delta (B.1.617.2; L452R, T478K) and Epsilon (B.1.429; L452R) (Supplementary Table 4). Each variant was neutralized by at least four of the ten monoclonal antibodies (Fig. 4e,f and Extended Data Fig. 10), indicating at least partial protection. Most antibodies neutralized either Alpha, Beta and Gamma (N501Y) or Delta and Epsilon (L452R) (Fig. 4f), indicating that neutralization was dependent on RBD mutations shared between the variants. Two antibodies were broadly neutralizing (Fig. 4f and Extended Data Fig. 10). Additionally, we tested neutralization efficacy in plasma of the nine participants (Fig. 4g). Overall, plasma neutralization potency was highest against Wuhan-Hu-1, followed by Epsilon and Delta and lowest against Alpha, Beta and Gamma. Together, despite low plasma neutralization potency against variants in some participants, we detected B cells with potently neutralizing BCRs against variants in the same participants.

Here we provided a detailed characterization of the B cell response to the BNT162b2 mRNA vaccine at a single-cell level. Parsing the S1-specific and S2-specific responses provides important insights into why a second dose is vital for protection. Our results indicate that the first dose activated a non-neutralizing recall response that initially targets epitopes in the S2-protein subunit, conserved across human-pathogenic betacoronaviruses[19,20], while the second dose boosted the neutralizing B cell responses to S1 including RBD.

The first dose provides an IgA-dominant plasmablast response against S2 with high SHM, which is cross-reactive to the betacoronaviruses OC43 and HKU1. This is consistent with a recall response of mucosal B_{mem} cells to previous pulmonary coronavirus infections. The first dose conveys a degree of protection against COVID-19 (ref. [3]) and this initial S2-specific response likely contributes to it. S2-specific antibodies can neutralize SARS-CoV-2 by inhibiting virus–cell membrane fusion and boosting antiviral T cell immunity[24–26]

After the S2-specific response on day 7, the frequency of minimally mutated S_A-B_{mem} cells increased on days 21 and 28. The S_A-B_{mem} cells at these later time points mostly targeted S1, including RBD and their low SHM frequency indicates a primary B cell response. As shown for B cell responses in influenza vaccine settings, blocking of a dominant epitope by early antibodies (anti-S2) can facilitate the development of later antibodies against another epitope (anti-S1)[27]. We observed that low mutation frequency corresponded to high affinity against RBD. High BCR affinity and a naive phenotype foster preferential recruitment into GC[28–30] and high affinity also promotes release from the GC as plasmablasts, plasma cells or B_{mem} cells[31]. High affinity in minimally mutated BCRs could therefore limit GC maturation to a relatively short time frame. mRNA vaccines against SARS-CoV-2 induce robust and prolonged GC

reactions, with plasmablasts and S-B_{mem} cells persisting in GC for over 3 months[32,33]. However, despite the prolonged GC reaction, the more rapid decrease of antibody titers against RBD and S1 than against S2 suggests less-efficient development of long-lived plasma cells upon recruitment from B_{naive} cells[34].

Plasma neutralization assays indicated a marked degree of immune escape of SARS-CoV-2 variants. However, several selected monoclonal antibodies from the same participants were neutralizing, indicating that even individuals with low neutralizing titers can raise a recall memory response after infection with SARS-CoV-2 variant. Together, our study provides a detailed characterization of the blood B cell response to the BNT162b2 mRNA vaccine. Our data emphasize the importance of the second dose in inducing generation of RBD-specific antibodies that contribute to neutralization of SARS-CoV-2 variants.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-021-01088-9.

Received: 26 August 2021; Accepted: 2 November 2021;
Published online: 30 November 2021

References

1. Amanat, F. et al. SARS-CoV-2 mRNA vaccination induces functionally diverse antibodies to NTD, RBD, and S2. *Cell* https://doi.org/10.1016/j.cell.2021.06.005 (2021).
2. Röltgen, K. et al. mRNA vaccination compared to infection elicits an IgG-predominant response with greater SARS-CoV-2 specificity and similar decrease in variant spike recognition. Preprint at https://doi.org/10.1101/2021.04.05.21254952 (2021).
3. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
4. Vogel, A. B. et al. BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature* **592**, 283–289 (2021).
5. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
6. Huang, Y., Yang, C., Xu, X.-F., Xu, W. & Liu, S.-W. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacol. Sin.* **41**, 1141–1149 (2020).
7. Piccoli, L. et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* **183**, 1024–1042 (2020).
8. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* https://doi.org/10.1038/nbt.4314 (2018).
9. Sanz, I. et al. Challenges and opportunities for consistent classification of human B cell and plasma cell populations. *Front. Immunol.* **10**, 2458 (2019).
10. Krzyzak, L. et al. CD83 modulates B cell activation and germinal center responses. *J. Immunol.* **196**, 3581–3594 (2016).
11. Maddur, M. S. et al. Human B cells induce dendritic cell maturation and favour Th2 polarization by inducing OX-40 ligand. *Nat. Commun.* **5**, 4092 (2014).
12. Palm, A.-K. E. & Henry, C. Remembrance of things past: long-term B Cell memory after infection and vaccination. *Front. Immunol.* **10**, 1787 (2019).
13. Bortnick, A. & Allman, D. What is and what should always have been: long-lived plasma cells induced by T cell-independent antigens. *J. Immunol.* **190**, 5913–5918 (2013).
14. Khodadadi, L., Cheng, Q., Radbruch, A. & Hiepe, F. The maintenance of memory plasma cells. *Front. Immunol.* **10**, 721 (2019).
15. Fink, K. Origin and function of circulating plasmablasts during acute viral infections. *Front. Immunol.* **3**, 78 (2012).
16. Mei, H. E. et al. Blood-borne human plasma cells in steady state are derived from mucosal immune responses. *Blood* **113**, 2461–2469 (2009).
17. Krammer, F. The human antibody response to influenza A virus infection and vaccination. *Nat. Rev. Immunol.* **19**, 383–397 (2019).
18. Dugan, H. L. et al. Preexisting immunity shapes distinct antibody landscapes after influenza virus infection and vaccination in humans. *Sci. Transl. Med.* https://doi.org/10.1126/scitranslmed.abd3601 (2020).
19. Ladner, J. T. et al. Epitope-resolved profiling of the SARS-CoV-2 antibody response identifies cross-reactivity with endemic human coronaviruses. *Cell Rep. Med.* **2**, 100189 (2021).
20. Premkumar, L. et al. The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Sci. Immunol.* https://doi.org/10.1126/sciimmunol.abc8413 (2020).
21. Kreer, C. et al. Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell* **182**, 843–854 (2020).
22. Gaebler, C. et al. Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639–644 (2021).
23. Greaney, A. J. et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe* **29**, 44–57(2021).
24. Chi, X. et al. A neutralizing human antibody binds to the N-terminal domain of the spike protein of SARS-CoV-2. *Science* **369**, 650–655 (2020).
25. Shah, P., Canziani, G. A., Carter, E. P. & Chaiken, I. The case for S2: the potential benefits of the S2 subunit of the SARS-CoV-2 spike protein as an immunogen in fighting the COVID-19 pandemic. *Front. Immunol.* **12**, 637651 (2021).
26. Winkler, E. S. et al. Human neutralizing antibodies against SARS-CoV-2 require intact Fc effector functions for optimal therapeutic protection. *Cell* **184**, 1804–1820 (2021).
27. Ellebedy, A. H. et al. Adjuvanted H5N1 influenza vaccine enhances both cross-reactive memory B cell and strain-specific naive B cell responses in humans. *Proc. Natl Acad. Sci. USA* **117**, 17957–17964 (2020).
28. Zuccarino-Catania, G. V. et al. CD80 and PD-L2 define functionally distinct memory B cell subsets that are independent of antibody isotype. *Nat. Immunol.* **15**, 631–637 (2014).
29. Abbott, R. K. et al. Precursor frequency and affinity determine B cell competitive fitness in germinal centers, tested with germline-targeting HIV vaccine immunogens. *Immunity* **48**, 133–146 (2018).
30. Turner, J. S. et al. Human germinal centres engage memory and naive B cells after influenza vaccination. *Nature* **586**, 127–132 (2020).
31. Li, X. et al. Cbl ubiquitin ligases control B cell exit from the germinal-center reaction. *Immunity* **48**, 530–541 (2018).
32. Lederer, K. et al. SARS-CoV-2 mRNA vaccines foster potent antigen-specific germinal center responses associated with neutralizing antibody generation. *Immunity* **53**, 1281–1295 (2020).
33. Turner, J. S. et al. SARS-CoV-2 mRNA vaccines induce persistent human germinal centre responses. *Nature* https://doi.org/10.1038/s41586-021-03738-2 (2021).
34. Weisel, F. J., Zuccarino-Catania, G. V., Chikina, M. & Shlomchik, M. J. A temporal switch in the germinal center determines differential output of memory B and plasma cells. *Immunity* **44**, 116–130 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2021

Methods

Study design, sample collection and storage. All studies were approved by the Institutional Review Board of Stanford University (IRB-3780) and the studies complied with relevant ethical regulations. All participants provided written informed consent before participating in the study. Nine healthy individuals were enrolled in the study (Supplementary Table 1). All individuals had undergone routine PCR with reverse transcription (RT–PCR) testing for SARS-CoV-2 infection before study. None of the participants had been previously diagnosed with SARS-CoV-2 infection. No statistical methods were used to predetermine sample sizes but our sample sizes are similar to those reported in previous publications[35–37]. Blood samples were collected in heparin tubes (BD) at four different time points including pre-vaccination (day 0), 7–9 d after initial vaccination (day 7), on the day of and before the second dose 21–23 d after initial vaccination (day 21) and 28–30 d after initial and 7–9 d after second vaccination (day 28). Plasma samples were obtained after centrifugation and stored at −80 °C. Peripheral blood mononuclear cells (PBMCs) were obtained by density gradient centrifugation over Ficoll PLUS medium (Cytiva) and stored in cell-freezing medium (Thermo Fisher Scientific). PBMCs were aliquoted and stored until use at −80 °C.

Generation of barcoded fluorescent antigen tetramers. Recombinant Avi-tag biotinylated SARS-CoV-2 S2 protein (Acro Biosystems, S2N-C52E8-25ug), SARS-CoV-2 RBD (Acro, SPD-C82E9-25ug) and SARS-CoV-2 S1 (Acro, S1N-C82E8-25ug) were mixed with barcoded, fluorescently labeled streptavidin (BioLegend) at a 4:1 molar ratio for 45 min while rotating. Excess biotin was added to saturate all streptavidin binding sites.

Flow cytometry, cell sorting and 10X sample preparation. PBMCs were thawed at 37 °C, treated for 15 min with DNase and washed in complete RPMI. PBMCs were enriched for B cells using the EasySep Human Pan-B Cell Enrichment kit (Stem Cell Technologies) according to the manufacturer instructions. B cell samples without antigen enrichment were stained with CD19, IgD, CD27 and CD38 TotalSeq-C antibodies (0.5 µg per 1,000,000 cells) (all BioLegend). For antigen-sorted B cell samples, cells were stained with the following fluorescently labeled antibodies according to standard protocols: CD19 (1:100 dilution), CD20 (1:300 dilution), CD38 (1:100 dilution) (all BD Biosciences), CD3 (1:60 dilution), CD27 (1:100 dilution), IgM (1:100 dilution), IgD (1:100 dilution), (all BioLegend), IgA (1:100 dilution) (Miltenyi Biotec), Sytox blue (1:1,000 dilution) (Thermo Fisher Scientific) and S-antigen tetramers. Additionally, sorted samples were labeled with TotalSeq-C hashtag 1–9 antibodies (0.5 µg per 1,000,000 cells) (BioLegend) for demultiplexing individual samples in downstream analysis. Single cells were sorted with a FACSAria II cell sorter (BD Biosciences) into cooled 1.5-ml tubes (BioRad). FACS data were collected with the BD FACSDiva (v.8.0) software. FlowJo v.10.7.1 (BD Biosciences) was used for flow cytometry data analysis. Flow cytometry experiments were performed with $n = 9$ biological replicates (study participants) and the experiment was performed twice.

Droplet-based single-cell sequencing. Using a Single-Cell 5′ Library and Gel Bead kit v.1.1 (10X Genomics, 1000165) and Next GEM Chip G Single-Cell kit (10X Genomics, 1000120), the cell suspension was loaded onto a Chromium single-cell controller (10X Genomics) to generate single-cell gel beads in the emulsion (GEMs) according to the manufacturer's protocol. Briefly, approximately 8,000 cells were added to each channel and approximately 4,000 target cells were recovered. Captured cells were lysed and released RNA was barcoded through reverse transcription in individual GEMs. The 5′ gene expression (GEX) libraries, single-cell V(D)J libraries (1000016) and cell surface protein libraries were constructed according to manufacturer protocols. Library quality was assessed using a 2200 TapeStation (Agilent). The libraries were sequenced using an Illumina Novaseq6000 sequencer with a paired-end 150-bp (PE150) reading strategy (Novogene).

Single-cell RNA-seq data processing. Raw gene expression and cell surface matrices were generated for each sample by the Cell Ranger Pipeline (v.6.0.1) coupled with human reference version GRCh38. Briefly, gene expression analyses of single cells were conducted using the R package Seurat (v.4.0.2) to perform data scaling, transformation, clustering, dimensionality reduction, differential expression analyses and most visualization[38]. The count matrix was filtered to remove cells with >10% of mitochondrial genes or low gene counts (<600 for enriched B cells and <200 for sorted B cells). The normalized data were integrated into one Seurat data file using the IntegrateData function. Principal component analysis was performed using variable genes. We compared the ranking of principal components (PCs) with the percentage of variance to determine the number of first-ranked PCs to use for UMAP[8] to reduce the integrated dataset into two dimensions. Afterwards, the same number of first-ranked PCs were used to construct a shared nearest-neighbor graph, which was used to cluster the cells. For sorted cells, all sorted cells (antigen-specific and nonspecific) were used in the UMAP. Contaminant cells (non-B cells) were removed and the data were normalized, integrated and clustered with only B cells. Specific B cell clusters were identified using canonical B cell markers[9] (Extended Data Fig. 1).

Identification of differentially expressed genes and functional enrichment. We performed differential gene expression testing using the FindMarkers function in Seurat with Wilcoxon rank-sum test and the Benjamini–Hochberg method was used to adjust the *P* values for multiple hypothesis testing. Differentially expressed genes were filtered using a minimum $\log_2$(fold change) of 0.25 and a maximum false discovery rate value of 0.05. Pathway analysis for differentially expressed genes was conducted using Metascape[39].

VDJ sequence analysis. BCR VDJ regions were generated for each sample using the Cell Ranger Pipeline (v.6.0.1). BCR sequences were then filtered to include cells that have one light and one heavy chain per cell. Consensus sequences were aligned to germline variable-chain immunoglobulin sequences with IMGT HighV-QUEST v.1.8.3 (ref. [40]). Clonal families were defined based on sharing the same heavy and light-chain V and J genes with >70% amino acid identity in heavy and light-chain CDR3s. Mutations were identified by aligning the nucleotide sequence to germline variable-chain immunoglobulin sequences with IMGT HighV-QUEST. To calculate the mutation frequency, we divided the number of mutations (silent and nonsilent) by the length of the V gene.

Recombinant monoclonal antibody production. Heavy chain and light-chain variable sequences were codon optimized and cloned into in-house vectors, containing human IgG1 constant region or κ or λ constant regions, respectively. Expi293F cells were transfected with heavy chain and light-chain plasmids using FectoPro (Polyplus transfection). Medium was collected after 7 days and monoclonal antibodies were purified with AmMag Protein A magnetic beads (Genscript). Antibody concentrations were measured with a nanodrop spectrophotometer (Thermo Fisher Scientific) and human IgG quantitation ELISAs (Bethyl Laboratories).

ELISA. For protein ELISAs, MaxiSorp 384-well plates (Thermo Fisher Scientific) were coated with 1 µg ml^{-1} recombinant SARS-CoV-2 S2 protein (Acro, S2N-C52H5), SARS-CoV-2 RBD (Acro, SPD-C52H3) or SARS-CoV-2 S1 (Acro, S1N-C52H3), HCoV-OC43 spike protein (Sino Biological, 40607-V08B), HCoV-HKU1 spike protein (Sino Biological, 40606-V08B), HCoV-229E spike protein (Sino Biological, 40605-V08B), HCoV-NL63 spike protein (Sino Biological, 40604-V08B), 10 µg ml^{-1} LPS (Sigma), 10 µg ml^{-1} calf thymus DNA (Invitrogen), 5 µg ml^{-1} of insulin (Sigma) or 2 µg ml^{-1} of flagellin (Invivogen) in carbonate-bicarbonate buffer at 4 °C overnight. Plates were washed six times with PBST (PBS + 0.1% Tween20) after each step. The plates were blocked with blocking buffer (PBS + 1% BSA) for 1 h at room temperature. Human plasma was serially diluted and added for 1 h at room temperature. Human monoclonal antibodies were added at concentrations of 10 µg ml^{-1} and three tenfold serial dilutions and incubated overnight at 4 °C. Secondary HRP-conjugated antibodies goat anti-human IgG (Bethyl Laboratories) or HRP-conjugated goat anti-human IgA (Bethyl Laboratories) were applied for 1 h at room temperature and plates were developed with TMB substrate (Thermo Fisher Scientific) and stopped with 2 N sulfuric acid. On each plate, four dilutions of positive-control plasma and secondary only controls were run. Additionally, a BSA-only plate was run in parallel to each antigen. Plates were read on a GloMax Explorer Microplate Reader (Promega). ELISA assays were performed at least twice, in duplicate or triplicate.

Biolayer interferometry. Monoclonal antibody interactions with S2, RBD and S1 protein were measured on an Octet Red96e (Fortebio/Sartorius). Association and dissociation curves were measured with monoclonal antibodies bound to anti-human IgG Fc Capture (AHC) sensors at 20 nM and antigens in solution at 0, 16.7, 50, 150 and 450 nM in 1× kinetic buffer (Fortebio/Sartorius). BLI analysis software (Fortebio/Sartorius, v.7.1) was used for data processing and analysis. Buffer controls were subtracted from antigen values and curves were fitted globally for each group, consisting of all concentrations of the same ligand. Association and dissociation curves and constants as well as K_D values for each antibody were reported and graphed with GraphPad Prism. Biolayer interferometry assays were performed 1–2 times, in triplicate.

Cell culture. Expi293F cells were cultured in 33% Expi293 Expression Medium (Gibco) and 67% Freestyle293 Expression Medium (Gibco). HeLa-ACE2 were kindly provided by D. Burton[41] and were cultured in Eagle's minimum essential medium (ATCC, 30-2003) with 10% heat-inactivated FBS (Corning) and 100 U ml^{-1} penicillin/streptomycin (Gibco). LentiX 293T cells (Takara Bio) were cultured in DMEM (ATCC, 30-2002) with 10% heat-inactivated FBS (Corning) and 100 U ml^{-1} penicillin/streptomycin (Gibco).

Generation SARS-CoV-2 spike pseudotyped lentiviral particles. Pseudotyped lentiviral particles were generated as previously described[42]. Briefly, LentiX 293T cells were seeded in 10-cm plates. After 24 h, cells were transfected using Fugene transfection reagent (Promega) with pHAGE-CMV-Luc2-IRES-ZsGreen-W, lentiviral helper plasmids (HDM-Hgpm2, HDM-tat1b and pRC-CMV-Rev1b) and wild-type or variant SARS-CoV-2 spike plasmids (parent plasmids publicly available from J. Bloom laboratory). After 48–60 h, viral supernatants were collected and spun at 1,000*g* for 10 min to remove cell debris.

Lentiviral supernatants were concentrated using LentiX concentrator (Takara) according to the manufacturer's instructions. Lentiviral pellets were resuspended at a 20-fold viral increase in EMEM and stored at −80 °C. Virus was titrated on HeLa-ACE2 cells.

Viral inhibition assays. Neutralization assays were performed as previously described[42]. Briefly, eightfold serially diluted plasma starting at 1:80 from vaccinated individuals or fivefold serially diluted monoclonal antibodies starting at a concentration of 10 μg ml^{-1} were incubated with SARS-CoV-2 pseudotyped virus for 1 h at 37 °C. The mixture was added to HeLa-ACE2 cells plated the previous day. After ~50 h post-infection, luciferase activity was measured on a GloMax Explorer Microplate Reader (Promega). Neutralization assays were performed 1–2 times in triplicate.

Statistics and software. GraphPad Prism v.9.1.0 and R v.4.0.3 were used for statistical analyses. Statistical tests used and significance levels are indicated in the respective methods section or in the figure legends. Normal distribution was assumed for the nine biological replicates where Student's *t*-test and ANOVA was performed. Graphical illustrations were created with BioRender.

Materials availability. Materials generated in this study will be made available on request and may require a material transfer agreement.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data have been deposited in the Sequence Read Archive linked to BioProject accession number PRJNA775994. All other data needed to evaluate the conclusions in the paper are present in the paper or the Source Data files. Source data are provided with this paper.

References

35. Wilk, A. J. et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat. Med.* **26**, 1070–1076 (2020).
36. Yang, A. C. et al. Publisher correction: dysregulation of brain and choroid plexus cell types in severe COVID-19. *Nature* https://doi.org/10.1038/s41586-021-04080-3 (2021).
37. Combes, A. J. et al. Publisher correction: global absence and targeting of protective immune states in severe COVID-19. *Nature* **596**, E8 (2021).
38. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
39. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
40. Alamyar, E., Duroux, P., Lefranc, M.-P. & Giudicelli, V. IMGT(®) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol. Biol.* **882**, 569–604 (2012).
41. Rogers, T. F. et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* **369**, 956–963 (2020).
42. Crawford, K. H. D. et al. Protocol and reagents for pseudotyping lentiviral particles with SARS-CoV-2 spike protein for neutralization assays. *Viruses* **12**, 513 (2020).

Acknowledgements

We thank all study participants who devoted time to our research. We thank D. Burton at Scripps Research Institute for providing ACE2-HeLa cells. We thank M. Baker at Stanford University for key discussions. This work was supported by a Foundation for the National Institutes of Health training grant (T32 AI007290-35) and National Science Foundation Graduate Research Fellowship (DGE-1656518) to R.C.B.; the German Research Foundation (LA3657/1) to T.V.L.; and Foundation for the National Institutes of Health R01 AR078268, U19 AI057229 and U01 AI101981 to W.H.R.

Author contributions

R.C.B., T.V.L. and W.H.R. were responsible for conceptualization, resources, project administration and funding acquisition. R.C.B., T.V.L., N.S.R., L.J.L. and S.J. were responsible for the methodology. R.C.B. and T.V.L. were responsible for software, validation, formal analysis, investigation, data curation, writing of the original draft, visualization and supervision. R.C.B., T.V.L., N.S.R., L.J.L. and W.H.R. were responsible for review and editing of the manuscript.

Competing interests

W.H.R. is a founder, member of the Board of Directors and consultant to Atreca, Inc. The remaining authors declare no competing interests.

Additional information

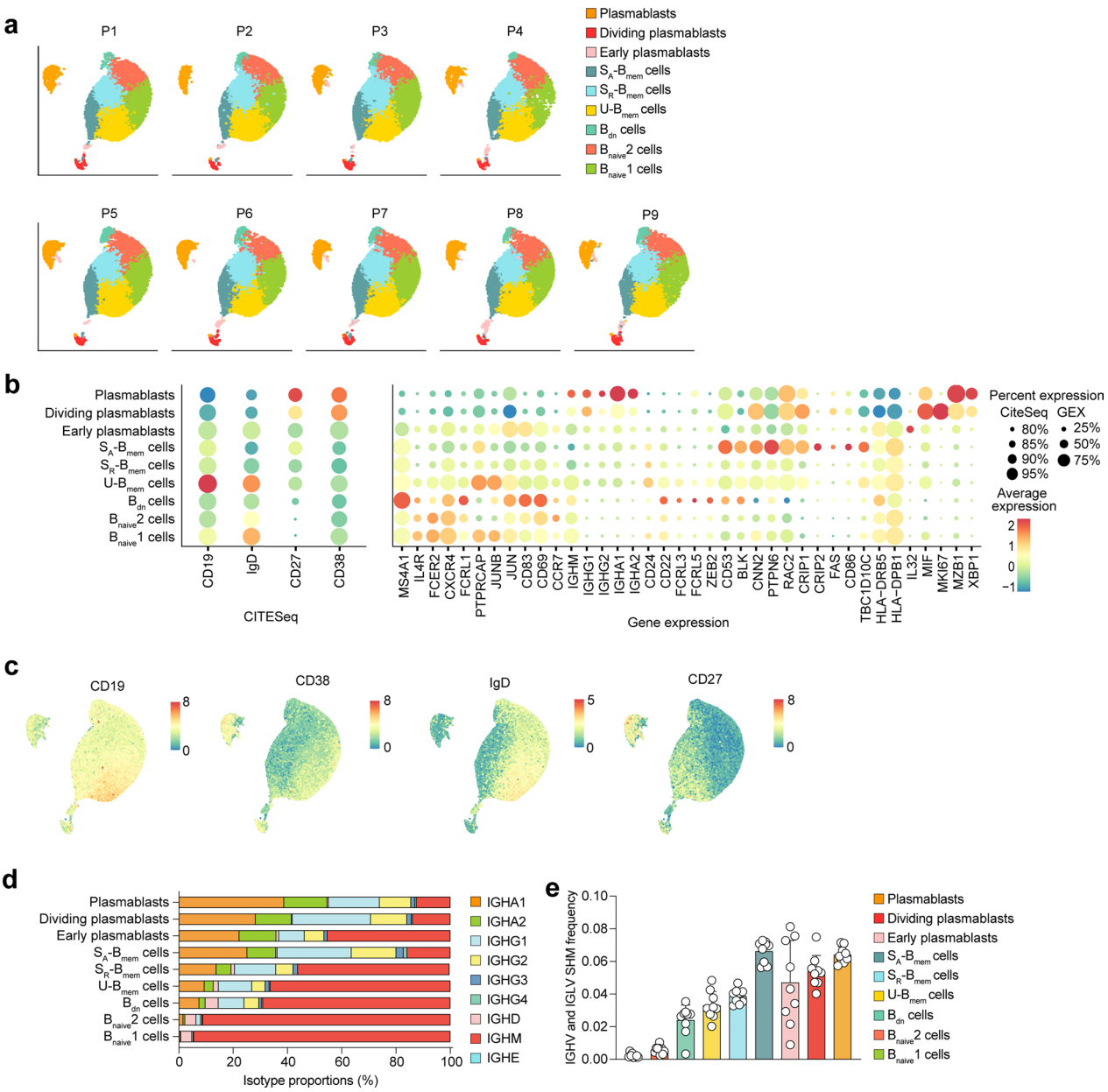
Extended data is available for this paper at https://doi.org/10.1038/s41590-021-01088-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-021-01088-9.

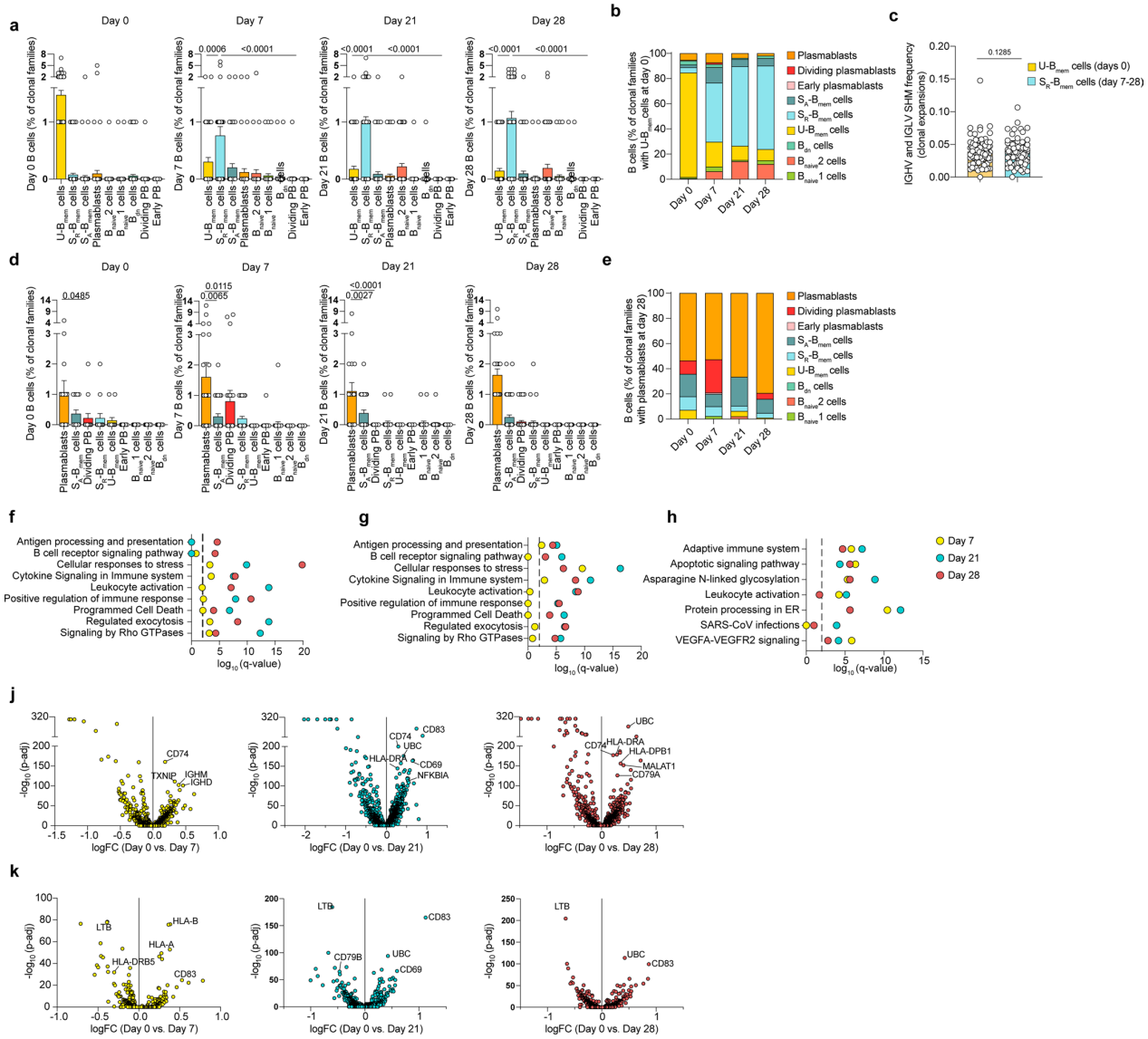
Correspondence and requests for materials should be addressed to William H. Robinson or Tobias V. Lanz.

Peer review information *Nature Immunology* thanks the anonymous reviewers for their contribution to the peer review of this work. Ioana Visan was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

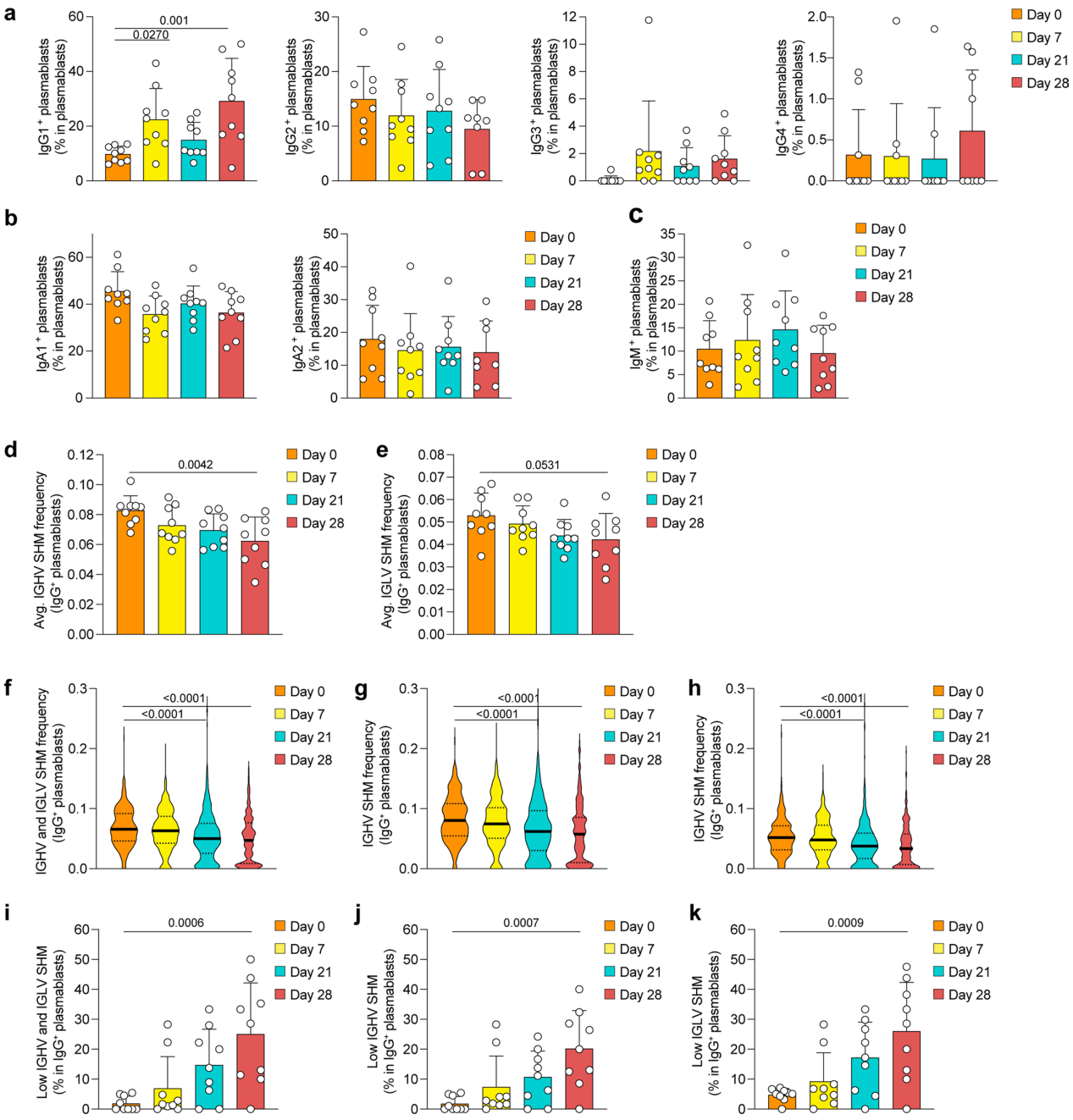
Reprints and permissions information is available at www.nature.com/reprints.



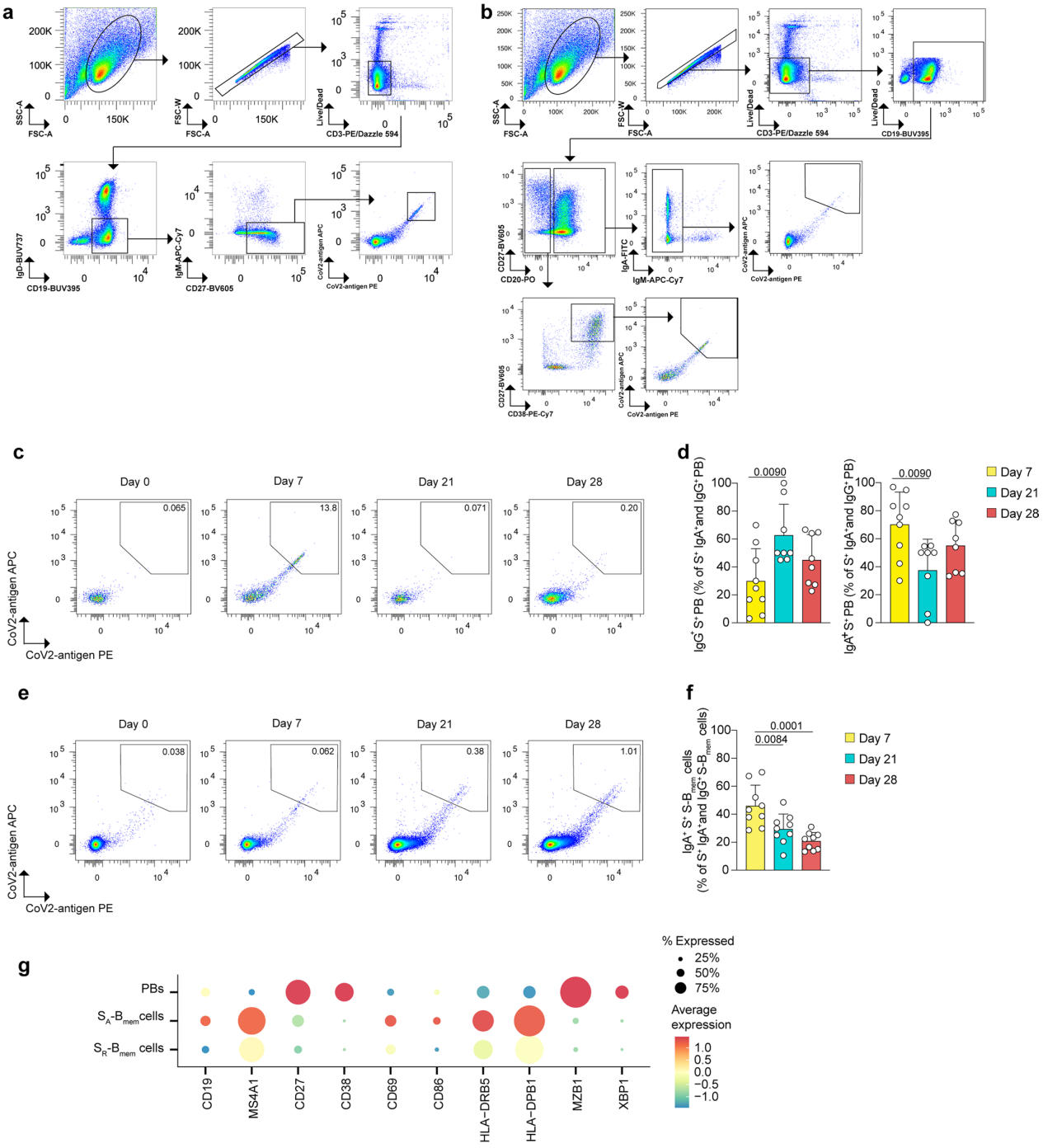
Extended Data Fig. 1 | Classification of UMAP B cell clusters. a-c, Single-cell transcriptome and CITESeq sequencing data showing **a**, the distribution of B cell clusters for each participant included in the study ($n = 9$). **b**, gene and cell surface expression of CITESeq markers (left) and sets of differentially expressed genes (right) for each B cell cluster, represented as heatmaps (average expression levels) and dot plots (percent expression). **c**, UMAP visualization with overlaid CITESeq data showing expression of cell surface markers CD19, CD38, IgD and CD27. **d**, **e**, Single-cell BCR repertoire sequencing data showing **d**, average percentages of immunoglobulin classes in each B cell cluster in $n = 9$ participants across all time points and **e**, average IGHV and IGLV SHM frequency in $n = 9$ individuals in the respective nine different B cell clusters. Individual values, means and standard deviations are shown. Memory B cells (B_{mem} cells), double-negative B cells (B_{dn} cells), naive B cells (B_{naive} cells), immunoglobulin heavy-V gene (IGHV), immunoglobulin light-V gene (IGLV), somatic hypermutation (SHM).



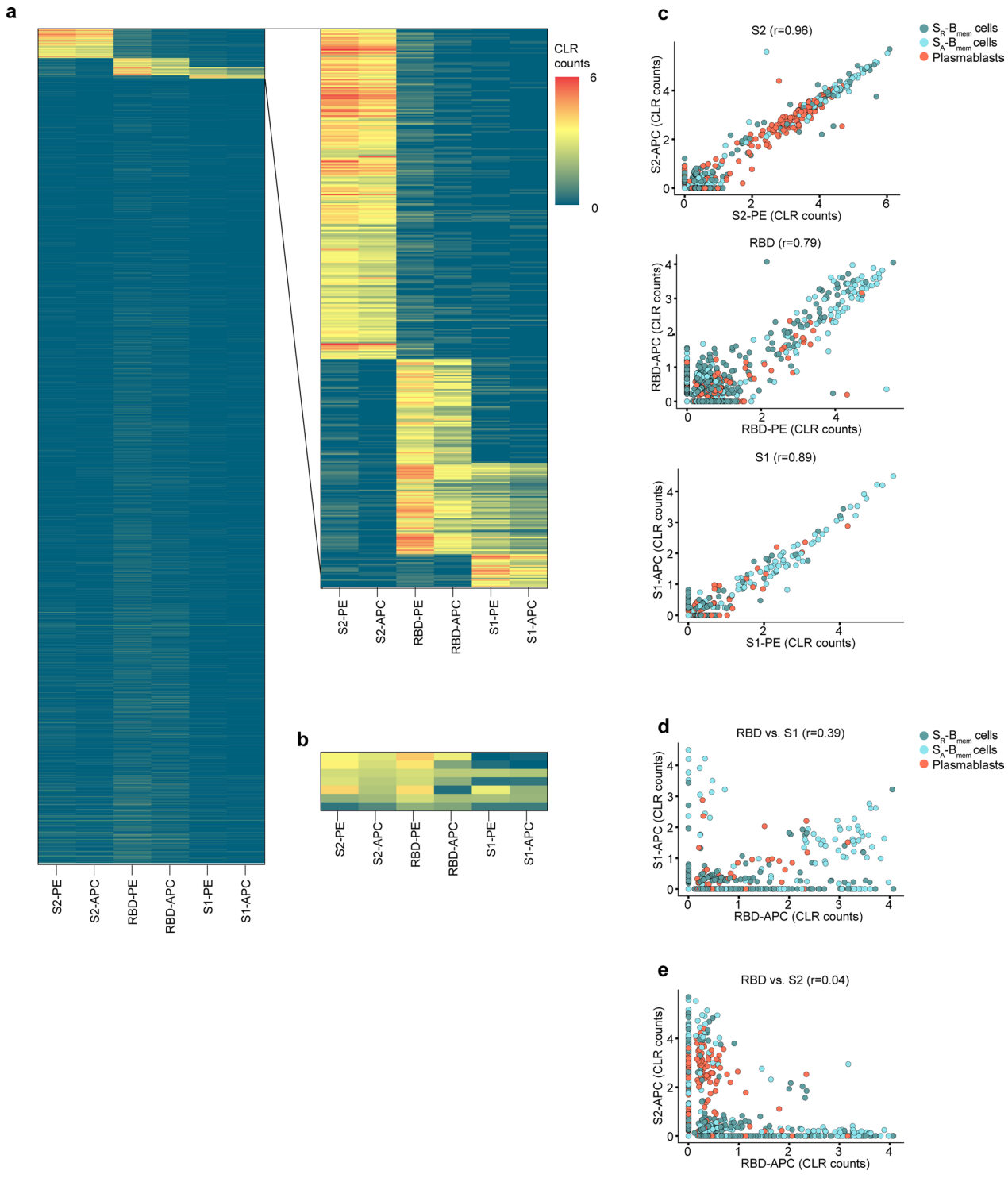
Extended Data Fig. 2 | Clonal and transcriptomic B cell responses to mRNA vaccination. a-k, Single-cell transcriptome and BCR repertoire sequencing data showing **a**, cells corresponding to clonal families shown in (Fig. 2a). Number of cells in each B cell cluster for each clonal family ($n = 106$ clonal families) at all four timepoints. Statistics are not calculated for day 0 because clonal families are selected to contain at least one clone at day 0 in the U-B_{mem} cell cluster. **b**, Proportions of cells ($n = 476$ cells) in each cluster from clonal families containing U-B_{mem} cells at day 0 for all four timepoints. **c**, Average SHM frequency of clonal U-B_{mem} cell families at day 0 and clonal S_R-B_{mem} cell families at days 7-28. Each dot represents the average of each mutation frequency for one clonal family. **d**, Cells corresponding to the clonal families shown in (Fig. 2d). Number of cells in each clonal family ($n = 52$ clonal families) in each B cell cluster at all four timepoints. Statistics are not calculated for D28 because clonal families are selected to contain at least one clone at D28 in the plasmablast cluster. **e**, Proportions of cells ($n = 274$ cells) in each cluster from clonal families with a plasmablasts at D28 for all four timepoints. **f-h**, Gene enrichment signatures of differentially expressed genes via Metascape in **f**, U-B_{mem} cells, **g**, S_R-B_{mem} cells and **h**, plasmablasts for day 7, day 21 and day 28 versus day 0. Dashed line represents $q < 0.01$. **j,k**, Volcano plots of differentially expressed genes in **j**, U-B_{mem} cells and **k**, S_R-B_{mem} cells at timepoints day 7 (left), day 21 (center), and day 28 (right) relative to day 0. **a**, **d**, Individual values with mean and standard errors of the mean are shown. **c**, Individual values with mean and standard deviations are shown. Exact p-values according to **a,d**, two-tailed Kruskal-Wallis test, followed by Dunnett's multiple comparison test and **c**, paired two-tailed t test. Memory B cells (B_{mem} cells), resting switched memory B cells (S_R-B_{mem} cells), activated switched memory B cells (S_A-B_{mem} cells), double-negative B cells (B_{dn} cells), naive B cells (B_{naive} cells), immunoglobulin heavy-V gene (IGHV), immunoglobulin light-V gene (IGLV), somatic hypermutation (SHM).



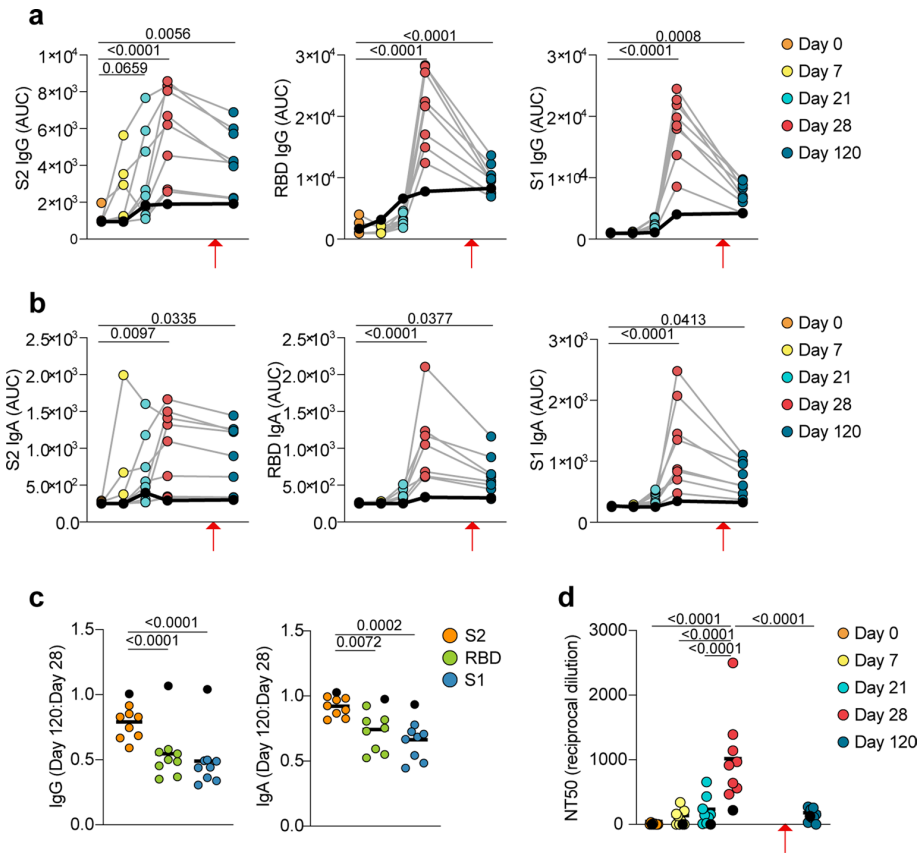
Extended Data Fig. 3 | Plasmablast characteristics in response to mRNA vaccination. a-k, Single-cell transcriptome and BCR repertoire sequencing data of the global (non-sorted) B cell dataset showing proportions of **a**, IgG subclasses, **b**, IgA subclasses and **c**, IgM in total plasmablasts in $n = 9$ participants. **d,e**, Average SHM frequency in IgG$^+$ plasmablasts for **d**, IGHV and **e**, IGLV in $n = 9$ participants. **f-h**, Distribution of SHM frequencies in single plasmablasts from all $n = 9$ individuals, separated by **f**, IGHV and IGLV, **g**, IGHV, and **h**, IGLV. **i-k**, Proportion of low SHM frequency (≤ 0.01) IgG$^+$ plasmablasts in all plasmablasts, separated by **i**, IGHV and IGLV, **j**, IGHV and **k**, IGLV for $n = 9$ individuals. **a-e**, **i-k**, Individual values, means, and standard deviations are shown. **f-h**, Violin plots with medians and interquartile ranges are shown. Exact p-values according to **a-e, i-k**, two-tailed one-way ANOVA, followed by Dunnett's multiple comparison test, and **f-h**, Kruskal-Wallis test followed by Dunnett's multiple comparison test. Immunoglobulin heavy-V gene (IGHV), immunoglobulin light-V gene (IGLV), somatic hypermutation (SHM).



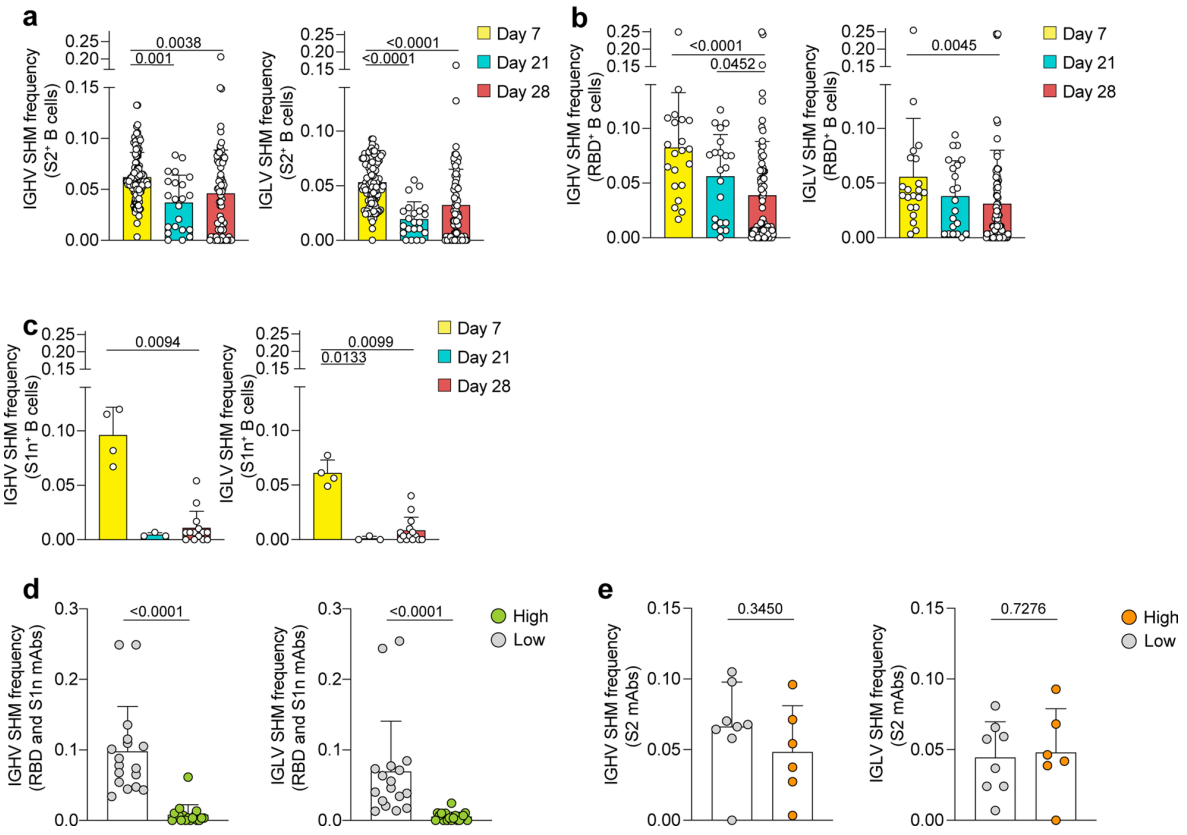
Extended Data Fig. 4 | Characterization of tetramer-sorted S-antigen-specific switched B_{Mem} and plasmablasts. a-f, Flow cytometry data showing **a**, the gating strategy for cell sorting of live singlet S-B_{mem} cells and plasmablasts based on marker expression: $CD3^-$, $CD19^+$, IgD^-, IgM^-, $CD27^+$. Cells were stained with two antigen-tetramers (PE labeled and APC labeled) for each subunit of S (S1, S2, and RBD), respectively. Tetramer-double-positive cells were sorted and tetramer-double-negative cells were sorted as controls. **b**, Gating strategy for flow cytometry analysis of S-B_{mem} cells and plasmablasts. **c**, Representative flow cytometry analyses showing two-dimensional antigen-tetramer staining of S-specific plasmablasts. **d**, Proportion of IgG^+ (left) and IgA^+ (right) S-specific plasmablasts in all IgG^+ and IgA^+ S-specific plasmablasts at the indicated timepoints post vaccination. **e**, Representative flow cytometry analyses showing two-dimensional antigen-tetramer staining of S-specific S-B_{mem} cells. **f**, Proportion of IgA^+ S-specific S-B_{mem} cells in IgG^+ and IgA^+ S-specific S-B_{mem} cells at the indicated timepoints post vaccination. **g**, Single cell transcriptome sequencing data showing marker genes characteristic for the respective clusters shown in (Fig. 3d). **d**, **f**, individual values, means, and standard deviations are shown for $n = 9$ participants. Exact p-values according to two-tailed one-way ANOVA, followed by Dunnett's multiple comparison test. Memory B cells (B_{mem} cells), switched memory B cells (S-B_{mem} cells), plasmablasts (PB), S-specific (S^+).



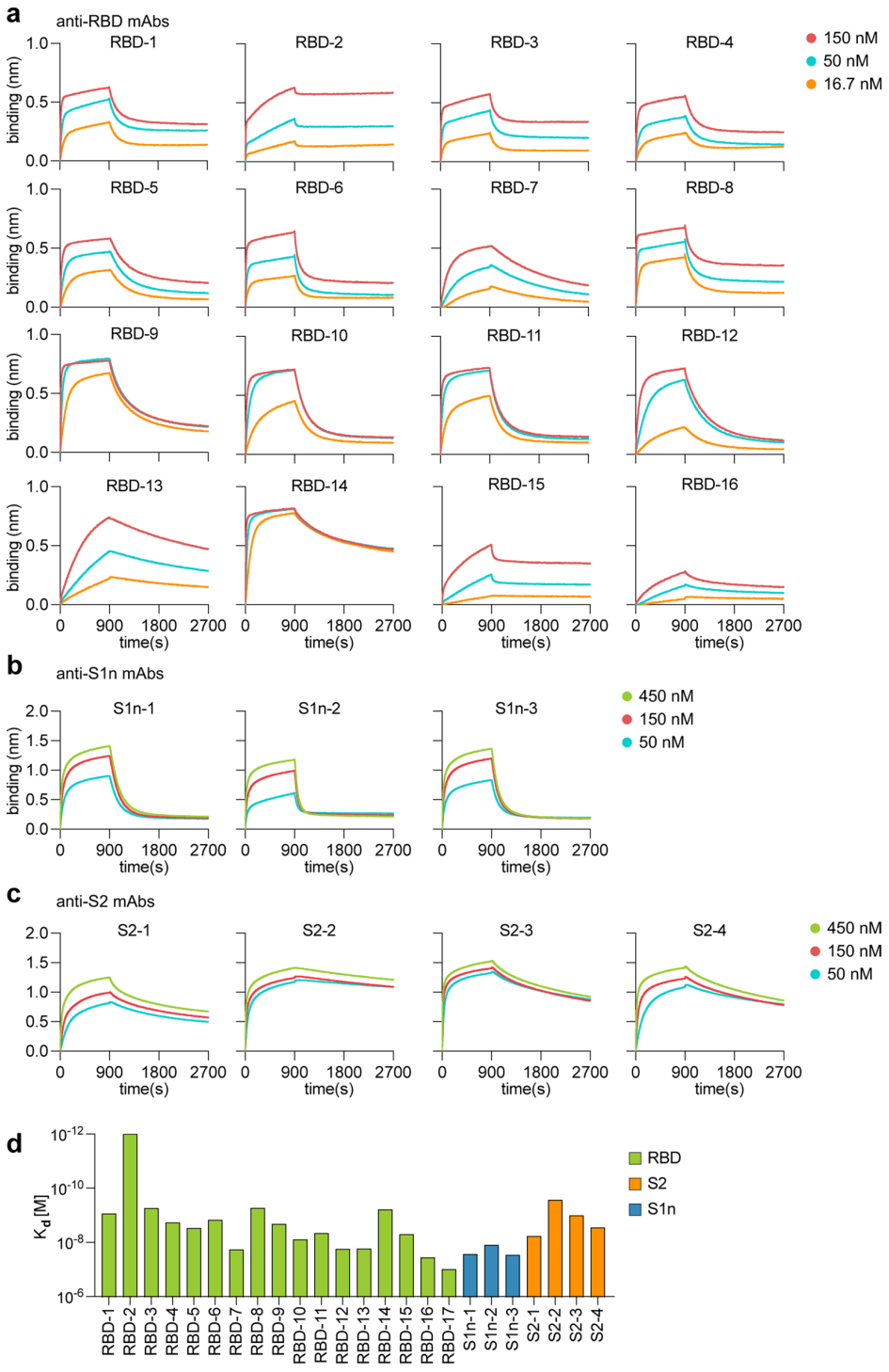
Extended Data Fig. 5 | Specific binding of S-antigen-tetramers to sorted B cells. a-e, Single-cell CITESeq sequencing data of antigen-tetramer barcodes identified in the de-multiplexed data set. Antigen-tetramer positive and negative CD3$^-$, CD19$^+$, IgD$^-$, IgM$^-$, CD27$^+$ B cells were mixed during the sort and distinguished during analysis. **a**, Heatmap of centered log ratio transformed counts (CLR) of barcoded PE and APC tetramers with S2, RBD, and S1 for all sorted cells. Inset magnifies the antigen-specific cells used for downstream analyses. **b**, Heatmap of CLR counts of polyreactive cells that were removed from downstream analyses. **c**, Correlation of barcoded PE and APC tetramers for S2 (top), RBD (middle) and S1 (bottom). **d**, Correlation of barcoded APC tetramers of S1 and RBD. **e**, Correlation of barcoded APC tetramers of S2 and RBD. **c-e**, r: Pearson's correlation coefficient.



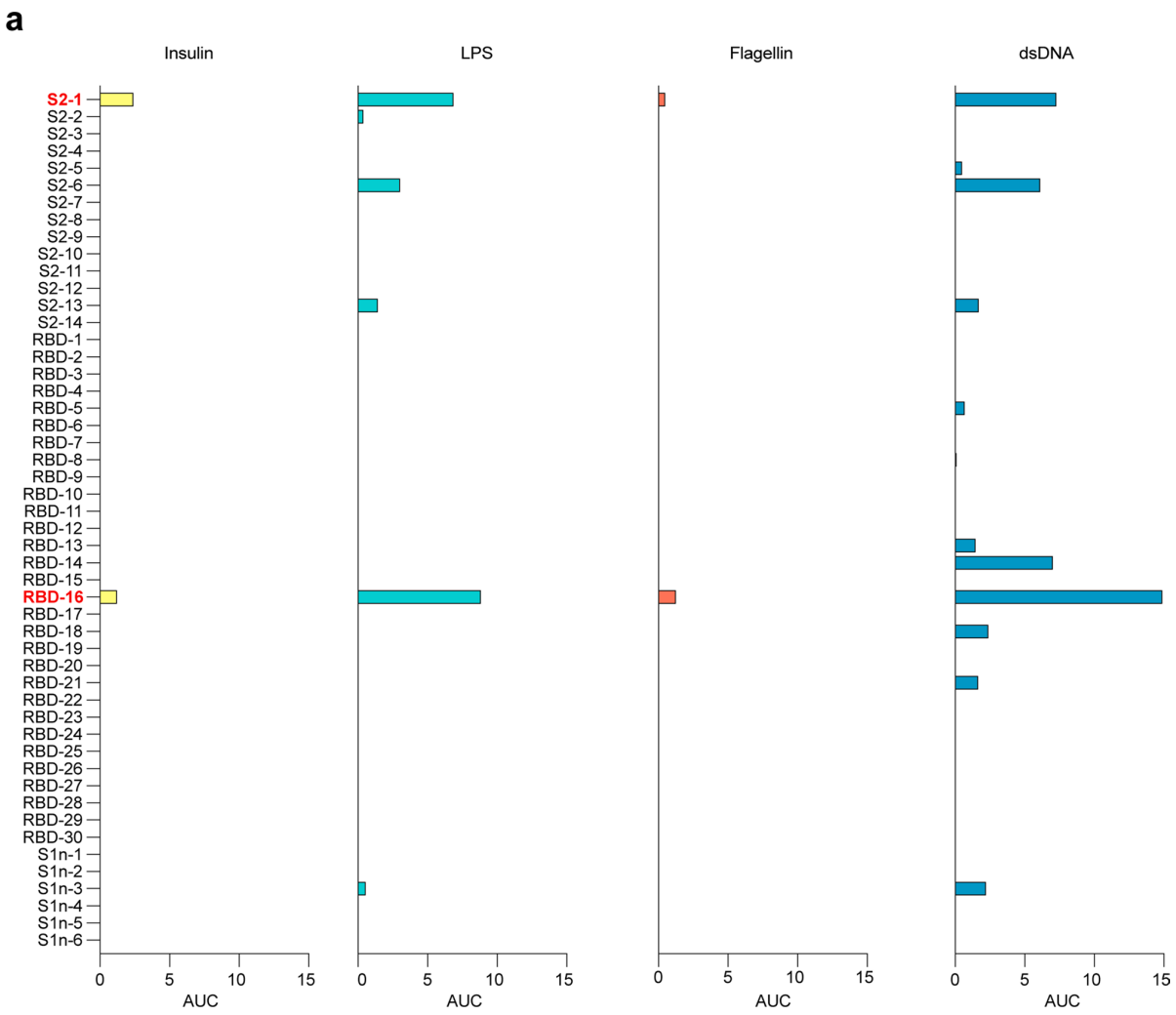
Extended Data Fig. 6 | Plasma reactivity to S protein subunits. a-c, ELISA data showing plasma reactivity against S-antigen subunits. **a**, IgG and **b**, IgA levels against S2 (left), RBD (center) and S1 (right). Area under the curve (AUC) of plasma dilutions is shown for $n = 9$ participants at five timepoints. **c**, Ratio of IgG titers (left) and IgA titers (right) of day 120 AUC to day 28 AUC for S2, RBD, and S1. **d**, Plasma neutralization of SARS-CoV-2 pseudovirus. Quantification (NT50) of neutralization titers is shown for $n = 9$ participants at five timepoints. Individual data points represent single B cells from $n = 9$ individuals. **a**, **b**, individual values are shown. **c**,**d**, Individual values and medians are shown. **a-d**, Exact p-values according to two-tailed one-way ANOVA, followed by Dunnett's multiple comparison test. P1 contracted COVID-19 at day 105, eight weeks after the second vaccination dose and two weeks before day 120 (red arrows). P1 is denoted by black dots and black lines. P1 is excluded from statistics in (**c**).



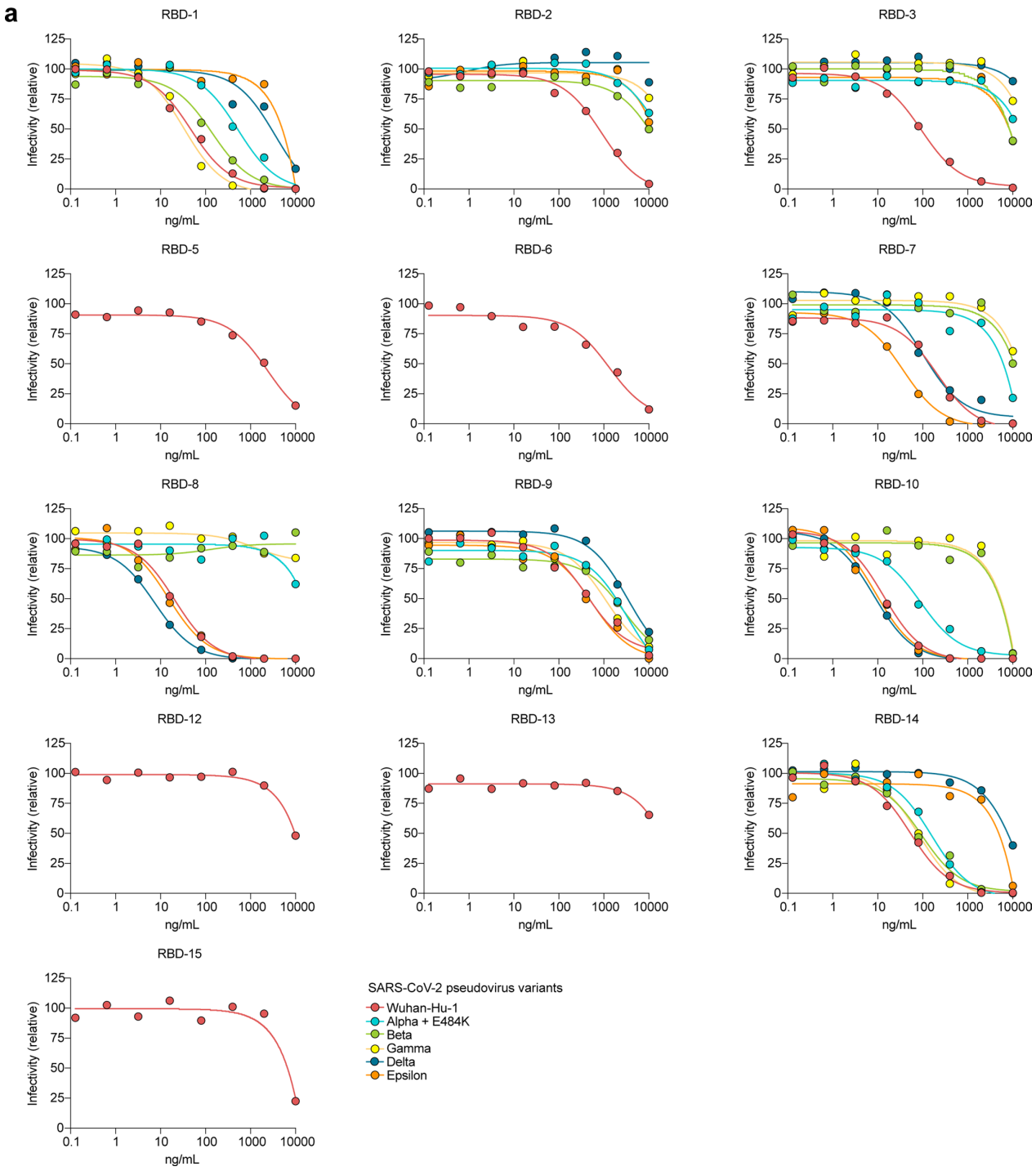
Extended Data Fig. 7 | Somatic hypermutation frequencies of SARS-CoV-2 S-specific B cells. a-e, Single-cell BCR repertoire sequencing data showing SHM frequencies for IGHV (left) and IGLV (right) of B cells specific for **a**, S2 ($n=210$ cells), **b**, RBD ($n=124$ cells) and **c**, S1n ($n=21$ cells), as determined by barcoded S-antigen-tetramers at three timepoints post vaccination. **d-e**, SHM frequencies of B cells that were recombinantly expressed as monoclonal antibodies specific for **d**, S1 (high, $n=17$ cells; low, $n=18$ cells) and **e**, S2 (high, $n=8$ cells; low, $n=6$ cells), corresponding to (Fig. 4a). High binding monoclonal antibodies are defined by the data shown in (Fig. 4a) as positive for both PE and APC barcode tetramer binding and ELISA AUC > 3. Individual values, means and standard deviations are shown. Exact p-values according to two-tailed one-way Kruskal Wallis test followed by Dunnett's multiple comparison test (**a**-**c**), and unpaired two-tailed Student's t-test (**d**, **e**). Immunoglobulin heavy-V gene (IGHV), immunoglobulin light-V gene (IGLV), somatic hypermutation (SHM), monoclonal antibodies (mAbs).



Extended Data Fig. 8 | Binding affinities of RBD-, S1n-, and S2-pecific monoclonal antibodies. Extended Data Figure 8: Binding affinities of RBD-, S1n-, and S2-specific monoclonal antibodies. **a–d**, Bio-layer interferometry data showing affinities of monoclonal antibodies for **a**, RBD, **b**, S1n, and **c**, S2. Binding curves are shown for the indicated concentrations of antigen. **d**, Kd values of antibodies calculated from the binding curves shown in (**a-c**). Monoclonal antibodies (mAbs).



Extended Data Fig. 9 | Polyreactivity of recombinantly expressed antibodies. a, ELISA data showing reactivity of recombinant monoclonal antibodies corresponding to (Fig. 4a,b) against human insulin, LPS, Flagellin and dsDNA, represented as area under the curve (AUC) of serial dilutions. Threshold at 0 was set to the average binding to BSA plus its three-fold standard deviation. Lipopolysaccharide (LPS), double-stranded DNA (dsDNA), bovine serum albumin (BSA).



Extended Data Fig. 10 | Neutralization of SARS-CoV-2 Wuhan-Hu-1 pseudovirus and variants by monoclonal antibodies from vaccinated participants. **a**, Neutralization of SARS-CoV-2 pseudovirus Wuhan-Hu-1 and variants alpha to epsilon are shown as dilution curves for monoclonal antibodies corresponding to the indicated monoclonal antibodies in (Fig. 4a).

nature portfolio

Corresponding author(s): Tobias V. Lanz, William H. Robinson

Last updated by author(s): 10/26/2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BD FACSDiva (v.8.0) for collection of flow cytometry and FACS data. BLI analysis software (Fortebio/Sartorius, version 7.1) was used for collection of biolayer-interferometry data.
Data analysis	For single cell analysis, we used the cell ranger pipeline (v.6.0.1) from 10X genomics, and R package Seurat (v.4.0.2). FlowJo was used for data analysis of flow cytometry and FACS data. R version 4.0.3 and GraphPad Prism 9.1.0 was used for statistical analyses. BLI analysis software (Fortebio/Sartorius, version 7.1) was used for collection of biolayer-interferometry data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw sequencing data are available on NCBI sequence read archive (SRA) under BioProject accession number PRJNA775994. Additionally, processed data will be made available upon request. Sequencing data was aligned to GRCh38 and IMGT (05/2021).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Our single-cell transcriptomic analysis includes four serial timepoint measurements of n=9 individuals (36 samples), yielding a total of 131,138 single cells transcriptomes. The study size is in line with prior single-cell RNA-sequencing studies in COVID-19 patients (Wilk AJ, et al. Nat Med. 2020;26(7):1070-1076; Yang AC, et al. Nature. 2021;595(7868):565-571; Combes AJ, et al., Nature. 2021;591(7848):124-130). In addition, vaccine study samples tend to be more uniform than samples from individuals with infectious diseases, due to the defined dose of the immunogen and its application at a specific timepoint. Quality control during analysis assured that all identified B cell populations were present in all individuals at all four timepoints, as shown in Extended Data Fig. 1a.
Data exclusions	No data were excluded from the study.
Replication	All regimens of biological and technical replication are annotated in the methods section. Briefly, ELISA and neutralization experiments were performed at 1-2 times in at least two technical replicates. Flow cytometry experiments were performed with n=9 biological replicates (study participants), and the experiment was performed twice. Bio-layer interferometry experiments were performed 1-2 times, each in triplicate dilutions, and cross-checked with ELISA-results. Single-cell transcriptome and repertoire sequencing was performed in n=9 individual samples at 4 timepoints. For droplet-based single-cell sequencing, n=9 biological replicates for each timepoint were performed, and all nine samples were used in the downstream analysis. All attempts at replication were successful.
Randomization	As this study does not explore group differences, no specific randomization strategy has been applied for the selection of study participants. However, we aimed to capture a cross-section of the population with regard to gender and age (Extended Data table 1: female=4, male=5; age: 23-52). For antigen specific B cell analysis, the antigen group was defined by CLR counts of the two barcoded, florescent tetramers for each antigen. For analysis of mAbs used in the manuscript, the groups were defined by the reactivity of the mAbs to specific antigens by ELISA.
Blinding	Participants were de-identified. As this study does not explore group differences, investigators were not blinded to the de-identified sample names. The study does not rely on subjective measures but analyzes quantitative sequencing data. Differences across timepoints were assessed using a computational pipeline that was applied equally to all samples. Investigators were not blinded during the additional experiments, including FACS, ELISAs, and plasma neutralization. The additional experiments involving samples from participants all have quantitative readouts, and methods were applied equally to all samples.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

1. IgA-FITC Antibody, anti-human, Miltenyi, 130-113-475
2. APC/Cy7 anti-human IgM Antibody, Biolegend, 314520
3. PE/Dazzle 594 anti-human CD3 Antibody, BioLegend, 317346
4. BUV395 Mouse Anti-Human CD19, BD Biosciences, B563549
5. CD20 Monoclonal Antibody (HI47), Pacific Orange, Thermo Fisher Scientific, MHCD2030
6. BV605 Mouse Anti-Human CD27, BD Biosciences, 562656
7. PE-Cy7 Mouse Anti-Human CD38, BD Biosciences, 335790
8. BUV737 Mouse Anti-Human IgD, BD Biosciences, 612798

9. TotalSeq-C0050 anti-human CD19 Antibody, Biolegend, 302265
10. TotalSeq-C0571 anti-mouse IgD Antibody, Biolegend, 405747
11. TotalSeq-C0154 anti-human CD27 Antibody, Biolegend, 302853
12. TotalSeq-C0389, anti-human CD38 Antibody, Biolegend, 303543
13. TotalSeq-C0951 PE Streptavidin, Biolegend, 405261
14. TotalSeq-C0952 PE Streptavidin, Biolegend, 405263
15. TotalSeq-C0953 PE Streptavidin, Biolegend, 405265
16. TotalSeq-C0956 APC Streptavidin, Biolegend, 405283
17. TotalSeq-C0957 APC Streptavidin, Biolegend, 405285
18. TotalSeq-C0958 APC Streptavidin
19. TotalSeq-C0251 anti-human Hashtag 1 Antibody, Biolegend, 394661
20. TotalSeq-C0252 anti-human Hashtag 2 Antibody, Biolegend, 394663
21. TotalSeq-C0253 anti-human Hashtag 3 Antibody, Biolegend, 394665
22. TotalSeq-C0254 anti-human Hashtag 4 Antibody, Biolegend, 394667
23. TotalSeq-C0255 anti-human Hashtag 5 Antibody, Biolegend, 394669
24. TotalSeq-C0256 anti-human Hashtag 6 Antibody, Biolegend, 394671
25. TotalSeq-C0257 anti-human Hashtag 7 Antibody, Biolegend, 394673
26. TotalSeq-C0258 anti-human Hashtag 8 Antibody, Biolegend, 394675
27. TotalSeq-C0259 anti-human Hashtag 9 Antibody, Biolegend, 394677
28. Antibodies derived from participants were made in-house, see methods and Extended Data Table 3
29. HRP-conjugated goat anti-human IgG, Bethyl, A80-104P
30. HRP-conjugated goat anti-human IgA, Bethyl, A80-102P

Validation

1. https://www.miltenyibiotec.com/US-en/products/iga-antibody-anti-human-is11-8e10.html#fitc:100-tests-in-200-ul
2. https://www.biolegend.com/en-us/products/apc-cyanine7-anti-human-igm-antibody-7403?GroupID=BLG4120
3. https://www.biolegend.com/en-us/search-results/pe-dazzle-594-anti-human-cd3-antibody-11986
4. https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/buv395-mouse-anti-human-cd19.563549
5. https://www.thermofisher.com/antibody/product/CD20-Antibody-clone-HI47-Monoclonal/MHCD2030
6. https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/bv605-mouse-anti-human-cd27.562656
7. https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/clinical-discovery-research/single-color-antibodies-ruo-gmp/pe-cy-7-mouse-anti-human-cd38.335790
8. https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/buv737-mouse-anti-human-igd.612798
9. https://www.biolegend.com/en-us/punchout/search-results/totalseq-c0050-anti-human-cd19-antibody-16832
10. https://www.biolegend.com/en-us/punchout/search-results/totalseq-c0571-anti-mouse-igd-antibody-19765
11. https://www.biolegend.com/en-us/punchout/search-results/totalseq-c0154-anti-human-cd27-antibody-16840
12. https://www.biolegend.com/en-us/products/totalseq-c0389-anti-human-cd38-antibody-17311
13. https://www.biolegend.com/en-us/products/totalseq-c0951-pe-streptavidin-18176
14. https://www.biolegend.com/en-us/products/totalseq-c0952-pe-streptavidin-18177
15. https://www.biolegend.com/en-us/products/totalseq-c0953-pe-streptavidin-18178
16. https://www.biolegend.com/en-us/products/totalseq-c0956-apc-streptavidin-19607
17. https://www.biolegend.com/en-us/products/totalseq-c0957-apc-streptavidin-19606
18. https://www.biolegend.com/en-us/products/totalseq-c0958-apc-streptavidin-19883
19. https://www.biolegend.com/en-us/search-results/totalseq-c0251-anti-human-hashtag-1-antibody-17162
20. https://www.biolegend.com/en-us/products/totalseq-c0252-anti-human-hashtag-2-antibody-17163?GroupID=GROUP28
21. https://www.biolegend.com/en-us/search-results/totalseq-c0253-anti-human-hashtag-3-antibody-17164?GroupID=GROUP28
22. https://www.biolegend.com/en-us/search-results/totalseq-c0254-anti-human-hashtag-4-antibody-17165?GroupID=GROUP28
23. https://www.biolegend.com/en-us/search-results/totalseq-c0255-anti-human-hashtag-5-antibody-17166
24. https://www.biolegend.com/en-us/search-results/totalseq-c0256-anti-human-hashtag-6-antibody-18373
25. https://www.biolegend.com/en-us/products/totalseq-c0257-anti-human-hashtag-7-antibody-18374?GroupID=GROUP28
26. https://www.biolegend.com/en-us/clone-search/totalseq-c0258-anti-human-hashtag-8-antibody-18375
27. https://www.biolegend.com/en-gb/search-results/totalseq-c0259-anti-human-hashtag-9-antibody-18376
28. Antibodies derived from participants were validated by ELISA for their specific antigens and for polyreactive antigens, by Octet for their specific antigens, and in neutralization experiments with SARS-CoV-2 pseudovirus, as described in this manuscript (see Fig. 4a,e,f; Extended Data Fig. 8; Extended Data Fig. 9, Extended Data Fig. 10)

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Lenti-X 293T Cell Line (human, Takara, 632180), Expi293F (human, Thermo Fisher Scientific, A14527), HeLa-ACE2 (human, gift from Dennis Burton, see ref. 44). HeLa-hACE2 were generated in the lab of Dennis Burton. Briefly, HeLa-hACE2 were generated through transduction using a lentivirus carrying the human ACE2 gene. The pBOB-hACE2 construct was co-transfected into HEK293T cells along with lentiviral packaging plasmids pMDL, pREV, and pVSV-G (Addgene #12251, #12253, #8454) by Lipofectamine 2000 (Thermo Fisher Scientific, 11668019) according to the manufacturer's instructions. Supernatants were collected 48 h after transfection, then were transduced to pre-seeded HeLa or A549 cells. 12 h after transduction, stable cell lines were collected, scaled up and stored for neutralization assay (ref. 44).

Authentication

Cells from commercial sources were confirmed by short tandem repeat by the commercial sources. HeLa-ACE2 were not authenticated. However, we tested infectivity of both HeLa and HeLa-ACE2, and determined that only HeLa-ACE2 were able to be infected by SARS-CoV-2 pseudovirus.

Mycoplasma contamination	Expi293F were tested and were found negative for mycoplasma contamination. Other cell lines were not tested for mycoplasma contamination but no indication of contamination was observed.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Human participants were 5 males (ages 23 - 38) and 4 females (ages 26 - 52). All participants were naive to prior SARS-CoV-2 infection. Blood samples were collected before BNT162b2 first vaccination, as well as 7-9 days, 21-23 days, and 28 days after the first vaccination.
Recruitment	Participants were recruited as a part of elective vaccination campaigns for VA and Stanford with the BNT162b2 mRNA vaccine (Pfizer/BioNTech) at the VA Hospital Palo Alto and the Stanford Hospital. As the BNT162b2 vaccination was not widely available to the general population at the time, participants were personnel of the VA and Stanford hospitals. As mentioned above, it was still possible to select a relatively diverse group considering gender and age. As our study does not compare several groups and as all investigated parameters are quantitatively measurable parameters, the danger of introducing selection biases is low. Major factors known to alter the vaccine response are comorbidities and therapeutics that lead to immunosuppression. All participants were questioned on these matters to avoid immunosuppression. Comorbidities and immunosuppressive therapeutics were denied by all participants. We have no reason to assume that our study group diverges from the general population in significant ways.
Ethics oversight	Stanford University Institutional Review Board (IRB-3780)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	n/a
Study protocol	n/a
Data collection	n/a
Outcomes	n/a

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	PBMCs were thawed at 37C°, treated for 15min with DNAse and washed in complete RPMI. PBMCs were enriched for B cells using the EasySep Human Pan-B Cell Enrichment Kit (Stem Cell Technologies) according to the manufacturer instructions. B cell samples without antigen enrichment were stained with CD19, IgD, CD27, CD38 TotalSeq-C antibodies (all Biolegend). For antigen-sorted B cell samples, cells were stained with the following fluorescently labeled antibodies according to standard protocols: CD19, CD20, CD38 (all BD Biosciences), CD3, CD27, IgM, IgD, (all BioLegend), IgA (Miltenyi Biotec), Sytox blue (Thermo Fisher Scientific), and S-antigen tetramers (Extended Data Fig. 6a,b,d). Additionally, samples were labeled with TotalSeq-C hashtag 1-9 antibodies (Biolegend) for demultiplexing individual samples in downstream analysis. Single cells were sorted with a FACSAria II cell sorter (BD Biosciences) into cooled 1.5 ml tubes (BioRad).
Instrument	FACSAria II cell sorter (BD Biosciences)
Software	FACS Diva
Cell population abundance	IgA+/IgG+ switched memory B cells and plasmablasts (~20%) were sorted from enriched B cell blood samples. Purity was >95% when analyzed downstream in Seurat (v.4.0.2).

Gating strategy

Details of the gating strategy are annotated in Extended Data Fig. 5 a and b. Briefly, we applied forward and side scatter parameters (FSC-A, FSC-W, SSC) to select for live lymphocytes. We gated on SytoxBlue-/CD3-/CD19+/IgD-/CD27+/IgM- for the sorted cells. We gated on SytoxBlue-/CD3-/CD19+/CD20+/CD27+/IgM- for analysis of SARS-CoV-2 antigen staining of switched memory B cells. We gated on SytoxBlue-/CD3-/CD19+/CD20-/CD27+/CD38+ for analysis of SARS-CoV-2 antigen staining of plasmablasts.

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