nature

https://doi.org/10.1038/s41586-022-04527-1

Accelerated Article Preview

Germinal centre-driven maturation of B cell response to mRNA vaccination

Received: 31 October 2021

Accepted: 4 February 2022

Accelerated Article Preview
Published online: 15 February 2022

Cite this article as: Kim, W. et al. Germinal centre-driven maturation of B cell response to mRNA vaccination. *Nature* https://doi.org/10.1038/s41586-022-04527-1 (2022).

Wooseob Kim, Julian Q. Zhou, Stephen C. Horvath, Aaron J. Schmitz, Alexandria J. Sturtz, Tingting Lei, Zhuoming Liu, Elizaveta Kalaidina, Mahima Thapa, Wafaa B. Alsoussi, Alem Haile, Michael K. Klebert, Teresa Suessen, Luis Parra-Rodriguez, Philip A. Mudd, Sean P. J. Whelan, William D. Middleton, Sharlene A. Teefey, Iskra Pusic, Jane A. O'Halloran, Rachel M. Presti, Jackson S. Turner & Ali H. Ellebedy

Article

Germinal centre-driven maturation of B cell response to mRNA vaccination

https://doi.org/10.1038/s41586-022-04527-1

Received: 31 October 2021

Accepted: 4 February 2022

Published online: 15 February 2022

Wooseob Kim[1,11], Julian Q. Zhou[1,11], Stephen C. Horvath[1], Aaron J. Schmitz[1], Alexandria J. Sturtz[1], Tingting Lei[1], Zhuoming Liu[2], Elizaveta Kalaidina[3], Mahima Thapa[1], Wafaa B. Alsoussi[1], Alem Haile[4], Michael K. Klebert[4], Teresa Suessen[5], Luis Parra-Rodriguez[6], Philip A. Mudd[7,8], Sean P. J. Whelan[2], William D. Middleton[5], Sharlene A. Teefey[5], Iskra Pusic[9], Jane A. O'Halloran[6], Rachel M. Presti[6,8], Jackson S. Turner[1] & Ali H. Ellebedy[1,8,10] ✉

Germinal centres (GC) are lymphoid structures where B cells acquire affinity-enhancing somatic hypermutations (SHM), with surviving clones differentiating into memory B cells (MBCs) and long-lived bone marrow plasma cells (BMPCs)[1–5]. SARS-CoV-2 mRNA vaccination induces a persistent GC response that lasts for at least six months in humans[6–8]. The fate of responding GC B cells as well as the functional consequences of such persistence have not been elucidated. We detected SARS-CoV-2 spike (S)-specific MBCs in 42 individuals who had received two doses of BNT162b2, a SARS-CoV-2 mRNA vaccine six months earlier. S-specific IgG-secreting BMPCs were detected in 9 out of 11 participants. Using a combined approach of sequencing the B cell receptors of responding blood plasmablasts and MBCs, lymph node GC and plasma cells and BMPCs from eight individuals and expression of the corresponding monoclonal antibodies (mAbs), we tracked the evolution of 1540 S-specific B cell clones. We show that early blood S-specific plasmablasts – on average – exhibited the lowest SHM frequencies. In comparison, SHM frequencies of S-specific GC B cells increased by 3.5-fold within six months after vaccination. S-specific MBCs and BMPCs accumulated high levels of SHM, which corresponded with enhanced anti-S antibody avidity in blood and affinity as well as neutralization capacity of BMPC-derived mAbs. This study documents how the striking persistence of SARS-CoV-2 vaccination-induced GC reaction in humans culminates in affinity-matured long-term antibody responses that potently neutralize the virus.

B cell response to mRNA vaccination

We have previously shown that vaccination of humans with The Pfizer-BioNTech SARS-CoV-2 mRNA vaccine, BNT162b2 induces a robust but transient circulating plasmablast (PB) response and a persistent germinal centre (GC) reaction in the draining lymph nodes[6]. Whether these persistent GC responses lead to the generation of affinity-matured memory B cells (MBCs) and long-lived bone marrow-resident plasma cells (BMPCs) remains unclear. To address this question, we analyzed long-term B cell responses in the participants enrolled in our previously described observational study of 43 healthy participants (13 with a history of SARS-CoV-2 infection) who received two doses of BNT162b2 (Extended Data Tables 1)[6,7]. Long-term blood samples (n=42) and fine needle aspirates (FNAs) of the draining axillary lymph nodes (n=15) were collected 29 weeks post-vaccination (Fig. 1a). Bone marrow aspirates were collected 29 (n=11) and 40 weeks (n=2) post-vaccination, with the latter time point used only for B cell receptor (BCR) repertoire profiling (Fig. 1a). None of the participants who contributed FNA or bone marrow specimens had SARS-CoV-2 infection history.

GC B cells were detected in FNAs from all 15 participants (Fig. 1b, c, **left panels**, Extended Data Fig. 1a, Extended Data Table 2). All 14 participants with FNAs collected prior to week 29 generated S-binding GC B cell responses of varying magnitudes (Fig 1b, c, **right panels**, and Extended Data Table 2). Strikingly, S-binding GC B cells were detected in FNAs from 10 of 15 participants at week 29 (Fig. 1b, c, **right panels**, Extended Data Table 2), demonstrating that two thirds of the sampled participants maintained an antigen-specific GC B cell response for at least 6 months post-vaccination. S-binding lymph node plasma cells (LNPCs) were also detected in FNAs from all 15 participants and exhibited similar dynamics to S-binding GC B cells, albeit at lower frequencies within the total B cell population (Extended Data Fig. 1a, b, Extended Data Table 2). None of the FNAs demonstrated significant contamination with peripheral blood

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA. [2]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO, USA. [3]Division of Allergy and Immunology, Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO, USA. [4]Clinical Trials Unit, Washington University School of Medicine, St. Louis, MO, USA. [5]Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO, USA. [6]Division of Infectious Diseases, Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO, USA. [7]Department of Emergency Medicine, Washington University School of Medicine, St. Louis, MO, USA. [8]Center for Vaccines and Immunity to Microbial Pathogens, Washington University School of Medicine, St. Louis, MO, USA. [9]Division of Oncology, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA. [10]The Andrew M. and Jane M. Bursky Center for Human Immunology & Immunotherapy Programs, Washington University School of Medicine, St. Louis, MO, USA. [11]These authors contributed equally: Wooseob Kim, Julian Q. Zhou. ✉e-mail: ellebedy@wustl.edu

based upon the nearly complete absence of myeloid cells (Extended Data Table 2).

Frequencies of BMPCs secreting IgG or IgA antibodies against either the 2019-2020 inactivated influenza virus vaccine, the tetanus-diphtheria vaccine or S protein were assessed in bone marrow aspirates collected 29 weeks after vaccination by enzyme-linked immunosorbent spot assay (ELISpot) (Fig. 1d, e, Extended Data Fig. 1c). Influenza- and tetanus-diphtheria vaccine-specific IgG-secreting BMPCs were detectable (median frequencies of 1.4% and 0.15%, respectively) in all 11 participants (Fig. 1e). S-binding IgG-secreting BMPCs were detected in 9 of 11 participants (median frequency of 0.06%). IgA-secreting BMPCs specific to influenza vaccine were detected in 10 of 11 participants, but IgA-secreting BMPCs directed against the tetanus-diphtheria vaccine and the S protein were largely below the limit of detection (Extended Data Fig. 1c). All participants had detectable plasma anti-S IgG antibodies and circulating S-binding MBCs at the 29-week time point (Fig. 1f–h). Anti-S IgG titers at 29 weeks were higher than titers observed in a cohort of unvaccinated SARS-CoV-2 convalescent subjects measured 29 weeks post-infection[9–11] (Extended Data Fig. 1d). Vaccinated participants with SARS-CoV-2 infection history had significantly higher titers of anti-S IgG antibodies at five and 29 weeks compared to their naive counterparts[9,11,12] (Fig. 1f). Similar trends were observed for plasma anti-S IgM and IgA (Extended Data Fig. 1e). S-binding MBCs were detected in all participants, with a median frequency of 0.23% of total circulating B cells (Fig. 1g, h, Extended Data Fig. 1f).

To track S-specific B cell evolution and clonal distribution within blood, lymph node and bone marrow, we performed single-cell RNA sequencing (scRNA-seq) and concurrent BCR sequencing of immune cells from 8 participants who contributed specimens from the three compartments. We first sorted PBs from samples collected at their peak frequencies, one week after the second immunization[6] (Fig. 2a, **top panel**, Extended Data Fig. 2a). We then interrogated the dynamics of the immune response in draining axillary lymph nodes. Single-cell transcriptional analysis of lymph nodes revealed distinct immune cell populations, as previously described[13–16] (Fig. 2a, **bottom left panel**, Extended Data Fig. 2b, c, Extended Data Table 3). To further distinguish distinct B cell subsets in the lymph node, we performed unbiased secondary clustering of the B cell populations from the total cellular analysis (Fig. 2a, **bottom right panel**, Extended Data Fig. 2d, e, Extended Data Table 3). Around 40% and ~7.9% of the B cells in the lymph node had GC B cell and LNPC transcriptomic profiles, respectively.

We next generated recombinant monoclonal antibodies (mAbs) from expanded clones detected in FNA samples seven and 15 weeks after vaccination, representing early and late time points. For two of the eight participants from whom the late point was unavailable due to insufficient specimens, we analyzed two separate early time points – weeks five and seven for participant 02a, and weeks four and seven for participant 04. A total of 2099 recombinant mAbs were generated, of which 1503 (71.6%) bound SARS-CoV-2 S by enzyme-linked immunosorbent assay (ELISA) (Fig. 2b, Extended Data Table 4). In subsequent analyses, we included 37 previously identified S-binding mAbs generated from GC B cells at week 4 from participants 07, 20, and 22[6]. Clonal relationships were computationally inferred using heavy chains from scRNA-seq BCR libraries (Extended Data Table 5); bulk-seq BCR libraries for GC B cells, LNPCs (Extended Data Fig. 2g) and BMPCs (Extended Data Table 5); as well as previously published bulk-seq BCR libraries of sorted PBs and GC B cells[6], and magnetically enriched IgD^{low} activated B cells or MBCs from PBMC[17]. B cell clones with experimentally-validated S-binding B cells were designated S-binding clones (Extended Data Fig. 2f) and accounted for 43.1% and 64.4%, respectively of the single-cell profiled GC B cells and LNPCs (Extended Data Fig. 2h, Extended Data Table 3). B cells that were clonally related to S-binding B cells were also found in the PB compartment in blood (6.7%) and the MBC compartment in lymph nodes (0.3%) (Extended Data Fig. 2h, Extended Data Table 3).

B cell maturation in the germinal centre

We analyzed the proportion of S-binding GC B cells clonally related to week 4 circulating PBs. The frequencies of PB-related, S-binding GC B cells varied broadly among participants, ranging from 12.7% to 82.5% (Fig. 3a). Consistent with our flow cytometry results (Fig. 1c), GC B cells from long-lasting S-binding clones were observed for at least 29 weeks – more than 6 months – after vaccination (Extended Data Fig. 3a). In addition, we detected the presence of clonally related MBCs in blood at 29 weeks post-vaccination (Extended Data Fig. 3b). S-binding GC B cells accumulated significantly higher levels of somatic hypermutation (SHM) compared to clonally related PBs, and this difference increased over time (Fig. 3b). We observed a 3.5-fold increase in SHM frequency among all S-binding GC B cells between weeks 4 and 29 (Fig. 3c, Extended Data Fig. 3c). S-binding MBCs detected at 29 weeks post-vaccination, however, had slightly lower SHM frequencies than their clonally related GC B cell counterparts (Extended Data Fig. 3d). The relative proportion of S-binding GC B cells expressing BCR of IgA isotype increased in the lymph node over time (Extended Data Fig. 3e). Clonal analysis revealed a high degree of overlap between S-binding GC and LNPC compartments (Fig. 3e). Furthermore, SHM frequencies of both S-binding LNPCs and GC B cells increased over time at a remarkably similar rate with small differences (Fig. 3f) in contrast to those between S-binding PB and GC B cells (Fig. 3b).

Affinity maturation of antibody response

To determine whether the increase in SHM frequencies of S-specific GC B cells and LNPCs over time is reflected in increased circulating anti-S antibody binding affinity, we measured the avidity of plasma anti-S IgG. In participants without SARS-CoV-2 infection history, anti-S IgG avidity increased at 29 weeks compared to the 5 weeks' time point. Interestingly, participants with a history of SARS-CoV-2 infection had comparable plasma anti-S IgG avidity at five and 29 weeks post-vaccination (Fig. 4a). Consistently, SHM frequencies of S-binding LNPCs increased over time (Fig. 4b). S-binding BMPCs from 29- and 40-weeks post-vaccination exhibited a degree of SHM that was comparable to LNPCs from 15- and 29-week post-vaccination (Fig. 4b) and higher than any other S-binding B cell population except for MBCs (Extended Data Fig. 4a). To understand the evolutionary trajectory of vaccine-induced B cell lineages, we analyzed S-specific clones using a phylogenetic model tailored for BCR repertoires[18]. Consistent with their SHM frequencies (Fig. 4b), PBs tended to locate closer to the germline on the phylogenetic trees, whereas LNPCs and BMPCs tended to be evolutionarily more distant (Fig. 4c, Extended Data Fig. 4b). In contrast to PBs, which clustered to a separate branch of their own, BMPCs and LNPCs co-located on shared branches, suggesting a closer evolutionary relationship between BMPCs and LNPCs (Fig. 4c). Together, these results support a model where S-specific BMPCs are the products of affinity-matured, GC-derived LNPCs.

We next expressed mAbs derived from clonally related PBs and BMPCs and their corresponding monomeric antigen-binding fragments (Fabs) (Extended Data Table 6). We then examined binding affinity and *in vitro* neutralization capacity using biolayer interferometry (BLI) and high-throughput GFP-reduction neutralization test[19], respectively. BMPC-derived Fabs exhibited significantly higher binding affinity against S protein compared to PB-derived Fabs (Extended Data Fig. 4c, d). Of the 21 S-specific clones we detected among BMPCs, seven potently neutralized the SARS-CoV-2 D614G strain (Extended Data Fig. 4e). Importantly, these BMPC-derived mAbs showed higher neutralizing potency than their clonally related, PB-derived counterparts (Fig. 4d), consistent with the significantly increased binding affinity of the BMPC-derived Fabs to S protein (Fig. 4e). Overall, the increased frequency of SHM observed over time and the correlated functional improvements in neutralization suggest that the GC reactions induced by SARS-CoV-2 mRNA vaccination facilitate the development of affinity-matured BMPCs.

Discussion

This study evaluated whether the persistent GC response induced by SARS-CoV-2 mRNA-based vaccines in humans[6] results in the generation of affinity-matured MBCs and BMPCs[1,3,13,20,21]. The two-dose series of BNT162b2 induced a robust S-binding GC B cell response that lasted for at least 29 weeks post-vaccination. The fruits of such persistent GC reactions were evident in the form of circulating S-binding MBCs in all participants and S-specific BMPCs 29 weeks post-vaccination in all but two of the sampled participants. It is likely that S-specific BMPCs in those two participants are present but below the assay detection limit. Longitudinal tracking of over 1500 vaccine-induced B cell clones revealed the gradual accumulation of SHM and isotype switching to IgA within the GC B cell compartment. We also show that GC B cells differentiate into affinity-matured LNPCs within the lymph node, with some of these cells potentially migrating to the bone marrow where they establish long-term residence. The enhanced maturity of the secreted antibodies was reflected in the significantly increased avidity of circulating anti-S IgG antibodies over time. It is also evident from increased affinity of BMPC-derived mAbs detected six months after vaccination in comparison to that of their corresponding PB-derived mAbs. Our data corroborate multiple reports demonstrating the maturation of circulating MBC responses after SARS-CoV-2 mRNA vaccination in humans[9,10,12,22–24].

This is the first study to show that a persistent vaccine induced GC response in humans culminates in the induction of affinity-matured, antigen-specific BMPCs. Notably, none of the 11 bone marrow specimens came from participants with SARS-CoV-2 infection history. An intriguing finding in our study is that the S-specific BMPCs detected more than six months after vaccination exhibited high SHM frequencies relative to other B cell compartments. These data corroborate similar observations made in the mouse model[25,26]. The murine data led to a proposal of a division of labor between memory B cells and long-lived BMPCs[27,28]. Under that framework, BMPCs secrete highly specific, high-affinity antibodies that provide the first layer of protection against the invading pathogen upon re-exposure while MBCs would only be engaged in the event that the pathogen is not fully neutralized by BMPC-derived antibodies. Consistent with this notion, multiple reports have recently documented the evolution of circulating MBCs induced by SARS-CoV-2 mRNA vaccination in humans[9,10,12,23]. These reports have shown that not only the frequency of circulating S-binding MBCs increased over time, but their ability to recognize S proteins from emerging SARS-CoV-2 variants seems to have expanded as well[22,23]. These data indicate an important role for affinity maturation of responding B cell clones beyond increasing binding affinity to the immunizing antigen.

Our study raises a number of important questions that will need to be addressed in future studies concerning the effects of an additional homologous or heterologous immunization on the dynamics and products of ongoing GCs, particularly with respect to breadth of induced B cell responses. It also remains to be addressed whether the IgA^+ GC B cell compartment induced by this systemic immunization can give rise to long-term IgA^+ MBCs and BMPCs. Overall, our data demonstrate the remarkable capacity of mRNA-based vaccines to induce robust and persistent GC reactions that culminate in affinity-matured MBC and BMPC populations.

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/s41586-022-04527-1.

1. Victora, G. D. & Nussenzweig, M. C. Germinal Centers. *Annual Review of Immunology* **30**, 429–457, https://doi.org/10.1146/annurev-immunol-020711-075032 (2012).
2. Cyster, J. G. & Allen, C. D. C. B Cell Responses: Cell Interaction Dynamics and Decisions. *Cell* **177**, 524–540, https://doi.org/10.1016/j.cell.2019.03.016 (2019).
3. Radbruch, A. et al. Competence and competition: the challenge of becoming a long-lived plasma cell. *Nature Reviews Immunology* **6**, 741–750, https://doi.org/10.1038/nri1886 (2006).
4. Slifka, M. K., Antia, R., Whitmire, J. K. & Ahmed, R. Humoral Immunity Due to Long-Lived Plasma Cells. *Immunity* **8**, 363–372, https://doi.org/10.1016/s1074-7613(00)80541-5 (1998).
5. Tarlinton, D., Radbruch, A., Hiepe, F. & Dörner, T. Plasma cell differentiation and survival. *Current Opinion in Immunology* **20**, 162–169, https://doi.org/10.1016/j.coi.2008.03.016 (2008).
6. Turner, J. S. et al. SARS-CoV-2 mRNA vaccines induce persistent human germinal centre responses. *Nature* **596**, 109–113, https://doi.org/10.1038/s41586-021-03738-2 (2021).
7. Mudd, P. A. *et al*. SARS-CoV-2 mRNA vaccination elicits a robust and persistent T follicular helper cell response in humans. Cell, https://doi.org/10.1016/j.cell.2021.12.026 (2021).
8. Laidlaw, B. J. & Ellebedy, A. H. The germinal centre B cell response to SARS-CoV-2. *Nature Reviews Immunology* **22**, 7–18, https://doi.org/10.1038/s41577-021-00657-1 (2021).
9. Goel, R. R. et al. mRNA vaccines induce durable immune memory to SARS-CoV-2 and variants of concern. *Science* **374**, abm0829, https://doi.org/10.1126/science.abm0829 (2021).
10. Cho, A. et al. Anti-SARS-CoV-2 receptor binding domain antibody evolution after mRNA vaccination. *Nature* **600**, 517–522, https://doi.org/10.1038/s41586-021-04060-7 (2021).
11. Chen, Y. *et al*. Differential antibody dynamics to SARS-CoV-2 infection and vaccination (Cold Spring Harbor Laboratory, 2021).
12. Pape, K. A. et al. High-affinity memory B cells induced by SARS-CoV-2 infection produce more plasmablasts and atypical memory B cells than those primed by mRNA vaccines. *Cell Reports* **37**, 109823, https://doi.org/10.1016/j.celrep.2021.109823 (2021).
13. Turner, J. S. et al. Human germinal centres engage memory and naive B cells after influenza vaccination. *Nature* **586**, 127–132, https://doi.org/10.1038/s41586-020-2711-0 (2020).
14. King, H. W. et al. Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Science Immunology* **6**, eabe6291, https://doi.org/10.1126/sciimmunol.abe6291 (2021).
15. Haebe, S. et al. Single-cell analysis can define distinct evolution of tumor sites in follicular lymphoma. *Blood* **137**, 2869–2880, https://doi.org/10.1182/blood.2020009855 (2021).
16. Mourcin, F. et al. Follicular lymphoma triggers phenotypic and functional remodeling of the human lymphoid stromal cell landscape. *Immunity* **54**, 1901, https://doi.org/10.1016/j.immuni.2021.07.018 (2021).
17. Schmitz, A. J. et al. A vaccine-induced public antibody protects against SARS-CoV-2 and emerging variants. *Immunity* **54**, 2159–2166.e2156, https://doi.org/10.1016/j.immuni.2021.08.013 (2021).
18. Hoehn, K. B., Lunter, G. & Pybus, O. G. A Phylogenetic Codon Substitution Model for Antibody Lineages. *Genetics* **206**, 417–427, https://doi.org/10.1534/genetics.116.196303 (2017).
19. Liu, Z. et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host & Microbe* **29**, 477–488.e474, https://doi.org/10.1016/j.chom.2021.01.014 (2021).
20. Elsner, R. A. & Shlomchik, M. J. Germinal Center and Extrafollicular B Cell Responses in Vaccination, Immunity, and Autoimmunity. *Immunity* **53**, 1136–1150, https://doi.org/10.1016/j.immuni.2020.11.006 (2020).
21. Turner, J. S. et al. SARS-CoV-2 infection induces long-lived bone marrow plasma cells in humans. *Nature* **595**, 421–425, https://doi.org/10.1038/s41586-021-03647-4 (2021).
22. Tong, P. et al. Memory B cell repertoire for recognition of evolving SARS-CoV-2 spike. *Cell* **184**, 4969–4980.e4915, https://doi.org/10.1016/j.cell.2021.07.025 (2021).
23. Sokal, A. et al. mRNA vaccination of naive and COVID-19-recovered individuals elicits potent memory B cells that recognize SARS-CoV-2 variants. *Immunity* **54**, 2893–2907.e5, https://doi.org/10.1016/j.immuni.2021.09.011 (2021).
24. Lucas, C. et al. Impact of circulating SARS-CoV-2 variants on mRNA vaccine-induced immunity. *Nature* **600**, 523–529, https://doi.org/10.1038/s41586-021-04085-y (2021).
25. Smith, K. G. C. The extent of affinity maturation differs between the memory and antibody-forming cell compartments in the primary immune response. *The EMBO Journal* **16**, 2996–3006, https://doi.org/10.1093/emboj/16.11.2996 (1997).
26. Florian, Griselda, Chikina, M. & Mark, A Temporal Switch in the Germinal Center Determines Differential Output of Memory B and Plasma Cells. *Immunity* **44**, 116–130, https://doi.org/10.1016/j.immuni.2015.12.004 (2016).
27. Purtha, W. E., Tedder, T. F., Johnson, S., Bhattacharya, D. & Diamond, M. S. Memory B cells, but not long-lived plasma cells, possess antigen specificities for viral escape mutants. *Journal of Experimental Medicine* **208**, 2599–2606, https://doi.org/10.1084/jem.20110740 (2011).
28. Pape, K. A., Taylor, J. J., Maul, R. W., Gearhart, P. J. & Jenkins, M. K. Different B Cell Populations Mediate Early and Late Memory During an Endogenous Immune Response. *Science* **331**, 1203–1207, https://doi.org/10.1126/science.1201730 (2011).

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

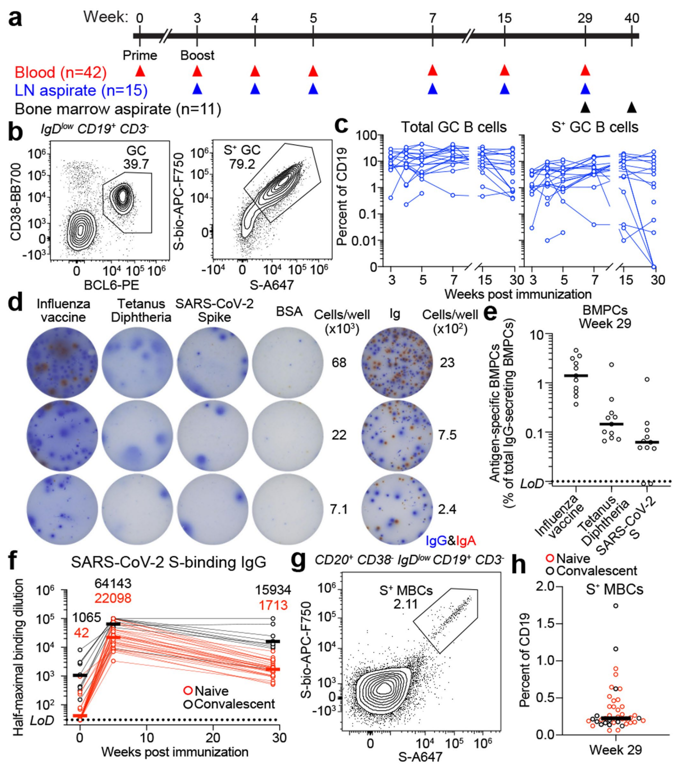


Fig. 1 | Persistence of humoral immune responses to SARS-CoV-2 mRNA vaccination. **a**, Forty-three participants (13 with SARS-CoV-2 infection history) were enrolled, followed by vaccination. Blood (n=42) was collected before and at indicated time points after vaccination. For 15 participants without infection history, aspirates of draining axillary lymph nodes were collected at indicated time points after vaccination. For 11 participants without infection history, aspirates of bone marrow were collected at 29 and 40 weeks post-vaccination. **b**, Representative flow cytometry plots of GC B cells ($CD19^{+}$ $CD3^{-}$ IgD^{low} $BCL6^{+}$ $CD38^{int}$) and S-binding GC B cells in lymph nodes 29 weeks post-vaccination.0 **c**, Kinetics of total (left) and S-binding GC B cells (right) as gated in **b**. **d**, Representative ELISpot wells coated with the indicated antigens, bovine serum albumin or anti-immunoglobulin and developed in blue (IgG) and red (IgA) after plating the indicated numbers of BMPCs. **e**, Frequencies of IgG-secreting BMPCs specific for the indicated antigens 29 weeks post-vaccination. Symbols at each time point represent one sample in **c** (n=15) and **e** (n=11). **f**, Plasma anti-S IgG titers measured by ELISA in participants without (red, n=29) and with (black, n=9) infection history. Horizontal lines and numbers indicate geometric means. Results from one experiment performed in duplicate. Dotted lines indicate detection limit in **e** and **f**. **g**, Representative flow cytometry plot of S-binding MBCs ($CD20^{+}$ $CD38^{-}$ IgD^{low} $CD19^{+}$ $CD3^{-}$) in blood 29 weeks post-vaccination. **h**, Frequencies of S-specific MBCs in participants without (red, n=29) and with (black, n=13) infection history as gated in **g**. Horizontal lines indicate medians in **e** and **h**.

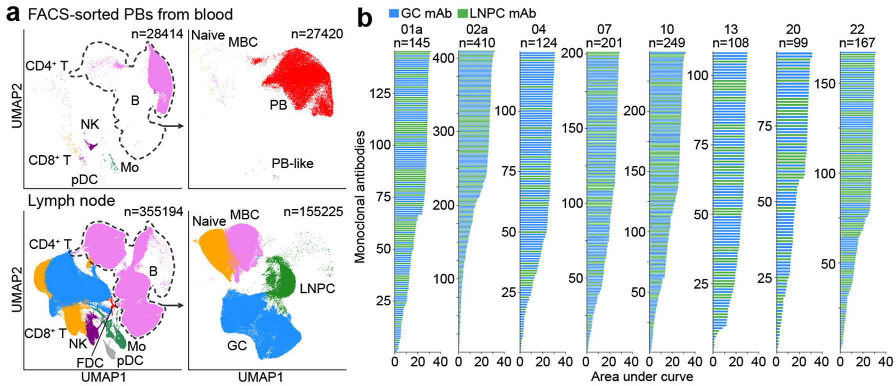


Fig. 2 | Identification of SARS-CoV-2 S-binding B cell clones in draining axillary lymph nodes. a, Uniform manifold approximation and projection (UMAP) showing scRNA-seq transcriptional clusters of total cells (left) and of B cells (right) from PBs sorted from PBMC (upper) and from FNA of lymph nodes (lower). Each dot represents a cell, colored by phenotype as defined by transcriptomic profile. Total numbers of cells are at the top right corner. FDC, follicular dendritic cell; GC, GC B cell; Mo, monocyte; NK, natural killer cell; LNPC, lymph node plasma cell; PB, plasmablast; pDC, plasmacytoid dendritic cell; MBC, memory B cell. **b**, Positive binding of recombinant monoclonal antibodies (mAbs) derived from GC B cells (blue) or LNPCs (green) to SARS-CoV-2 S measured by ELISA. Results are from one experiment performed in duplicate.

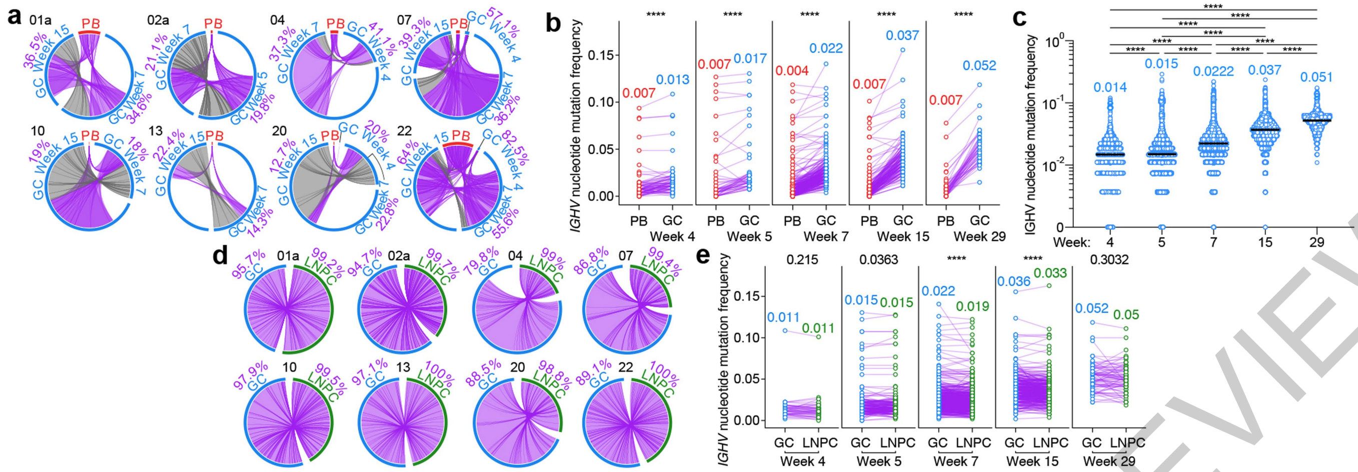


Fig. 3 | Maturation of SARS-CoV-2 S-binding B cells in the lymph node. **a**, Circos diagrams showing clonal overlap between S-binding PBs and GC B cells at indicated time points. Purple and gray chords correspond to, respectively, clones spanning both compartments, and clones spanning only the GC compartment. Percentages are of GC B cell clones related to PBs at each time point. **b**, Immunoglobulin heavy chain variable (*IGHV*) region nucleotide mutation frequency of clonally related PBs and GC B cells at Week 4 (n=81), 5 (n=52), 7 (n=289), 15 (n=162) and 29 (n=47). **c**, *IGHV* nucleotide mutation frequency of S-binding GC B cells at Week 4 (n=1701), 5 (n=21543), 7 (n=62927), 15 (n=49837) and 29 (n=3314). Horizontal lines and numbers represent medians. *P* values were determined by Kruskal-Wallis test followed by Dunn's multiple comparison test. **d**, Circos diagrams showing clonal overlap (purple) between S-binding GC B cells and LNPCs over combined time points. Percentages are of GC B cell clones overlapping with LNPCs or *vice versa*. Arc length corresponds to the number of BCR sequences and chord width corresponds to clone size in **a** and **d**. **e**, *IGHV* nucleotide mutation frequency of clonally related GC B cells and LNPCs at Week 4 (n=48), 5 (n=224), 7 (n=877), 15 (n=449) and 29 (n=76). Each dot represents the median SHM frequency of a clone within the indicated compartment, and medians are presented on the top of each data set in **b** and **e**. *P* values were determined by paired two-sided Mann-Whitney test and corrected for multiple testing using Benjamini and Hochberg's method in **b** and **e**. ****$P < 0.0001$.

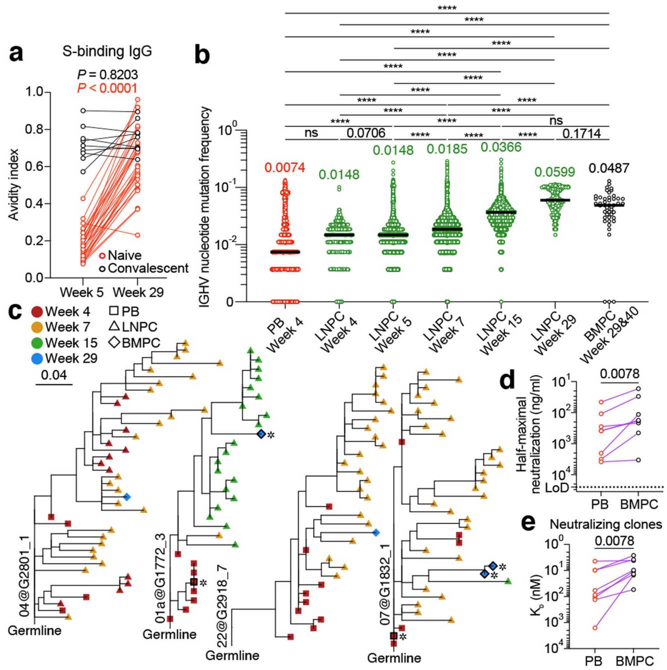


Fig. 4 | Evolution of B cell clones induced by SARS-CoV-2 vaccination. **a**, Avidity indices of plasma anti-S IgG between the indicated time points in participants without (red, n=29) and with (black, n=9) infection history. Results are from one experiment performed in duplicate. **b**, *IGHV* nucleotide mutation frequency of S-binding PBs (n=2735), LNPCs at Week 4 (n=552), 5 (n=11253), 7 (n=45436), 15 (n=24538) and 29 (n=571), and BMPCs (n=47). Horizontal lines and numbers represent median values. *P* values were determined by Kruskal-Wallis test followed by Dunn's multiple comparison test. **c**, Representative phylogenetic trees showing inferred evolutionary relationships between PBs (squares), LNPCs (triangles) and BMPCs (diamonds). Horizontal branch length represents the expected number of substitutions per codon in V-region genes, corresponding to the scale bar. Clone IDs are displayed near the root of the trees. Asterisks denote neutralizing mAbs. **d**, Neutralizing activity of clonally related PB- and BMPC-derived mAbs (n=8) against SARS-CoV-2 D614G strain. Dotted line indicates detection limit. Results are from one experiment with duplicates in **a** and **d**. **e**, Equilibrium dissociation constant (K_D) of neutralizing clone-derived Fabs (n=8) interacting with immobilized S protein measured by BLI. Symbols indicate K_D values of clonally related, PB (red)- and BMPC (black)-derived Fabs, respectively. *P* values were determined by two-tailed Wilcoxon matched-pairs signed rank test in **a**, **d** and **e**. ns > 0.9999, ****P < 0.0001.

Methods

Sample collection, preparation, and storage

All studies were approved by the Institutional Review Board of Washington University in St Louis. Written consent was obtained from all participants. Forty-three healthy volunteers were enrolled, of whom 13 had a history of confirmed SARS-CoV-2 infection (Extended Data Table 1). Fifteen out of 43 healthy participants provided FNAs of draining axillary lymph nodes. In 6 out of the 15 participants, a second draining lymph node was identified and sampled following secondary immunization. One participant (15) received the boost vaccination in the contralateral arm; draining lymph nodes were identified and sampled on both sides. Eleven out of 43 healthy participants provided bone marrow aspirates. Forty-eight participants who had recovered from mild SARS-CoV-2 infection but had not been vaccinated within 7 months of illness were previously described[21].

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

Ultrasound-guided FNA of draining axillary lymph nodes was performed by a radiologist or a qualified physician's assistant under the supervision of a radiologist. Scans were performed with a commercially available ultrasound unit (Loqic E10, General Electric, Milwaukee, WI) using an L2-9 linear array transducer with transmit frequencies of 7, 8, and 9 MHz or a L6-15 linear array transducer with transmit frequencies of 10, 12, and 15 MHz. Lymph node dimensions and cortical thickness were measured, and the presence and degree of cortical vascularity and location of the lymph node relative to the axillary vein were determined before each FNA. For each FNA sample, six passes were made under continuous real-time ultrasound guidance using 25-gauge needles, each of which was flushed with 3 ml of RPMI 1640 supplemented with 10% FBS and 100 U/ml penicillin-streptomycin, followed by three 1 ml rinses. Red blood cells were lysed with ammonium chloride buffer (Lonza), washed with washing buffer (phosphate-buffered saline supplemented with 2% FBS and 2 mM EDTA), and immediately used or cryopreserved in 10% dimethyl sulfoxide in FBS. Participants reported no adverse effects from phlebotomies or serial FNAs.

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

Antigens

Recombinant soluble spike (S) protein derived from SARS-CoV-2 was expressed as previously described[29]. In brief, a mammalian cell codon-optimized nucleotide sequences coding for the soluble version of S (GenBank: MN908947.3, amino acids 1-1,213) including a C-terminal thrombin cleavage site, T4 fold trimerization domain and hexahistidine tag was cloned into the mammalian expression vector pCAGGS. The S protein sequence was modified to remove the polybasic cleavage site (RRAR to A) and two stabilizing mutations were introduced (K986P and V987P, wild-type numbering). Recombinant proteins were produced in Expi293F cells (Thermo Fisher Scientific) by transfection with purified plasmid using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). Supernatants from transfected cells were collected 3 days after transfection, and recombinant proteins were purified using Ni-NTA agarose (Thermo Fisher Scientific), then buffer-exchanged into PBS and concentrated using Amicon Ultra centrifugal filters (MilliporeSigma). For flow cytometry staining, recombinant S was labeled with Alexa Fluor 7647-NHS ester or biotinylated using the EZ-Link Micro NHS-PEG4-Biotinylation Kit (Thermo Fisher Scientific); excess Alexa Fluor 647 and biotin were removed using 7-kDa Zeba desalting columns (Thermo Fisher Scientific).

For expression of biotinylated SARS-CoV-2 S Avitag, the CDS of pCAGGS vector containing recombinant soluble SARS-CoV-2 S protein was modified to encode 3' Avitag insert after the 6xHIS tag (5'-HIS tag-GGCTCCGGGCTGAACGACATCTTCGAAGCCCAGAAGATTG AGTGGCATGAG-Stop-3'; HHHHHHGSGLNDIFEAQKIEWHE-) using inverse PCR mutagenesis in a method described previously[30]. Protein expression and purification of SARS-CoV-2 S-Avitag was performed using the same methods as above. Immediately, after purification, site-specific biotinylation was performed similar to Avidity recommendations. Specifically, SARS-CoV-2 S-Avitag substrate was at 40 uM concentration with 15 ug BirA enzyme/ml in a 0.05 M Bicine buffer at pH 8.3 containing 10 mM ATP, 10 mM MgOAc and 50 uM Biotin, and the reaction was performed for 30 °C for 1 h. The protein was then concentrated/buffer exchanged with PBS using a 100 kDa Amicon Ultra centrifugal filter (MilliporeSigma).

Flow cytometry and cell sorting

Staining for flow cytometry analysis and sorting was performed using freshly isolated or cryo-preserved PBMCs or FNAs. For FNA staining, cells were incubated for 30 min on ice with biotinylated and Alexa Fluor 647-conjugated recombinant soluble S and PD-1-BB515 (EH12.1, BD Horizon, 1:100) in 2% FBS and 2 mM EDTA in PBS (P2), washed twice, then stained for 30 min on ice with IgG-BV480 (goat polyclonal, Jackson ImmunoResearch, 1:100), IgA-FITC (M24A, Millipore, 1:500), CD45-A532 (HI30, Thermo, 1:50), CD38-BB700 (HIT2, BD Horizon, 1:500), CD20-Pacific Blue (2H7, 1:400), CD27-BV510 (O323, 1:50), CD8-BV570 (RPA-T8, 1:200), IgM-BV605 (MHM-88, 1:100), HLA-DR-BV650 (L243, 1:100), CD19-BV750 (HIB19, 1:100), CXCR5-PE-Dazzle 594 (J252D4, 1:50), IgD-PE-Cy5 (IA6-2, 1:200), CD14-PerCP (HCD14, 1:50), CD71-PE-Cy7 (CY1G4, 1:400), CD4-Spark685 (SK3, 1:200), streptavidin-APC-Fire750, CD3-APC-Fire810 (SK7, 1:50) and Zombie NIR (all BioLegend) diluted in Brilliant Staining buffer (BD Horizon). Cells were washed twice with P2, fixed for 1 h at 25 °C using the True Nuclear fixation kit (BioLegend), washed twice with True Nuclear Permeabilization/Wash buffer, stained with FOXP3-BV421 (206D, BioLegend, 1:15), Ki-67-BV711 (Ki-67, BioLegend, 1:200), T-bet-BV785 (4B10, BioLegend, 1:400), BCL6-PE (K112-91, BD Pharmingen, 1:25), and BLIMP1-A700 (646702, R&D, 1:50) for 1 h at 25 °C, washed twice with True Nuclear Permeabilization/Wash buffer and resuspended in P2 for acquisition. For memory B cell staining, PBMC were incubated for 30 min on ice with biotinylated and Alexa Fluor 647-conjugated recombinant soluble S in P2, washed twice, then stained for 30 min on ice with IgG-BV480 (goat polyclonal, Jackson ImmunoResearch, 1:100), IgD-Super Bright 702 (IA6-2, Thermo, 1:50), IgA-FITC (M24A, Millipore, 1:500), CD45-A532 (HI30, Thermo, 1:50), CD38-BB700 (HIT2, BD Horizon, 1:500), CD24-BV421 (ML5, 1:100), CD20-Pacific Blue (2H7, 1:400), CD27-BV510 (O323, 1:50), CD8-BV570 (RPA-T8, 1:200), IgM-BV605 (MHM-88, 1:100), CD19-BV750 (HIB19, 1:100), FcRL5-PE (509f6, 1:100), CXCR5-PE-Dazzle 594 (J252D4, 1:50), CD14-PerCP (HCD14, 1:50), CD71-PE-Cy7 (CY1G4, 1:400), CD4-Spark685 (SK3, 1:200), streptavidin-APC-Fire750, CD3-APC-Fire810 (SK7, 1:50) and Zombie NIR (all BioLegend) diluted in Brilliant Staining buffer (BD Horizon). Cells were washed twice with P2 and resuspended in P2 for acquisition. All samples were acquired on an Aurora using SpectroFlo v.2.2 (Cytek). Flow cytometry data were analyzed using FlowJo v.10 (BD Biosciences).

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

ELISA

Assays were performed in MaxiSorp 96-well plates (Thermo Fisher) coated with 100 ul of recombinant SARS-CoV-2 S, Donkey anti-human IgG (H+L) antibody (Jackson ImmunoResearch, 709-005-149) or BSA diluted to 1ug/ml in PBS, and plates were incubated at 4 °C overnight. Plates then were blocked with 10% FBS and 0.05% Tween 20 in PBS. Plasma or purified monoclonal antibodies were serially diluted in blocking buffer and added to the plates. Monoclonal antibodies and plasma samples were tested at 10 ug/ml and 1:30 starting dilution, respectively, followed by 7 additional 3-fold serial dilutions. Plates were incubated for 90 min at room temperature and then washed 3 times with 0.05% Tween 20 in PBS. Secondary antibodies were diluted in blocking buffer before adding to wells and incubating for 60 min at room temperature. HRP-conjugated goat anti-human IgG (H+L) antibody (Jackson ImmunoResearch, 109-035-088, 1:2500) was used to detect monoclonal antibodies. HRP-conjugated goat anti-Human IgG Fcγ fragment (Jackson ImmunoResearch, 109-035-190, 1:1500), HRP-conjugated goat anti-human serum IgA α chain (Jackson ImmunoResearch, 109-035-011, 1:2500), and HRP-conjugated goat anti-human IgM (Caltag, H15007, 1:4000) were used to detect plasma antibodies. Plates were washed 3 times with PBST and 3 times with PBS before the addition of *o*-phenylenediamine dihydrochloride peroxidase substrate (MilliporeSigma). Reactions were stopped by the addition of 1M hydrochloric acid. Optical density measurements were taken at 490 nm. The threshold of positivity for recombinant mAbs was set as two times the optical density of background binding to BSA at the highest concentration of each mAb. The area under the curve for each monoclonal antibody and half-maximal binding dilution for each plasma sample were calculated using GraphPad Prism v.9. Plasma antibody avidity was measured as previously described[31]. Areas under the curve were calculated by setting the mean + three times the s.d. of background binding to BSA as a baseline. Briefly, plasma dilutions that would give an optical density reading of 2.5 were calculated from the serial dilution ELISA. S-coated plates were incubated with this plasma dilution as above and then washed one time for 5 minutes with either PBS or 8M urea in PBS, followed by 3 washes with PBST and developed as above. The avidity index was calculated for each sample as the optical density ratio of the urea-washed to PBS-washed wells.

ELISpot

ELISpot plates were coated overnight at 4 °C with Flucelvax Quadrivalent 2019/2020 seasonal influenza virus vaccine (Seqirus, 1:100), tetanus/diphtheria vaccine (Grifols, 1:20), SARS-CoV-2 S (10 ug/ml), anti-human Ig (Cellular Technology Limited) and BSA. A direct ex vivo ELISpot assay was performed to determine the number of total, vaccine-binding or recombinant S-binding IgG- and IgA-secreting cells present in PBMCs or enriched BMPCs using Human IgA/IgG double-color ELISpot kits (Cellular Technology Limited) according to the manufacturer's protocol. ELISpot plates were analyzed using an ELISpot analyzer (Cellular Technology Limited).

Single-cell RNA-seq library preparation and sequencing

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

Bulk B cell receptor sequencing

Sorted GC B cells and LNPCs from FNA, enriched BMPCs from bone marrow or enriched MBCs from PBMCs from blood were used for library preparation for bulk BCR sequecning. Circulating MBCs were magnetically isolated by first staining with IgD-PE and MojoSort anti-PE Nanobeads (BioLegend), and then processing with the EasySep Human B Cell Isolation Kit (StemCell Technologies) to negatively enrich IgDlo B cells. RNA was prepared from each sample using the RNeasy Plus Micro kit (Qiagen). Libraries were prepared using the NEBNext Immune Sequencing Kit for Human (New England Biolabs) according to the manufacturer's instructions without modifications. High-throughput 2×300-bp paired-end sequencing was performed on the Illumina MiSeq platform with a 30% PhiX spike-in according to manufacturer's recommendations, except for performing 325 cycles for read 1 and 275 cycles for read 2.

Preprocessing of bulk sequencing BCR reads

Preprocessing of demultiplexed pair-end reads were performed using pRESTO v.0.6.2[32] as previously described[6], with the exception that sequencing errors were corrected using the UMIs as they were without additional clustering (Extended Data Table 5). Previously preprocessed unique consensus sequences from reported samples[6] were included as they were. Previously preprocessed unique consensus sequences from reported samples[17] corresponding to participants 01, 02a, 04, 07, 10, 13, 20, and 22 were subset to those with at least two contributing reads and included.

Preprocessing of 10× Genomics single-cell BCR reads

Demultiplexed pair-end FASTQ reads were preprocessed using the 'cellranger vdj' command from 10× Genomics' Cell Ranger v.6.0.1 for alignment against the GRCh38 human reference v.5.0.0 ('refdata-cellranger-vdj-GRCh38-alts-ensembl-5.0.0'). The resultant 'filtered_contig.fasta' files were used as preprocessed single-cell BCR reads (Extended Data Table 5).

V(D)J gene annotation and genotyping

Initial germline V(D)J gene annotation was performed on the preprocessed BCRs using IgBLAST v.1.17.1[33] with the deduplicated version of IMGT/GENE-DB release 202113-2[34]. IgBLAST output was parsed using MakeDb.py from Change-O v.1.0.2[35]. For the single-cell BCRs, isotype annotation was pulled from the 'c_call' column in the 'filtered_contig_annotations.csv' files outputted by Cell Ranger.

For both bulk and single-cell BCRs, sequence-level quality control was performed, requiring each sequence to have non-empty V and J gene annotations; exhibit chain consistency in all annotations; bear fewer than 10 non-informative (non-A/T/G/C, such as N or -) positions; and carry a non-empty CDR3 with no N and a nucleotide length that is a multiple of 3. For single-cell BCRs, cell-level quality control was also performed, requiring each cell to have either exactly one heavy chain and at least one light chain, or at least one heavy chain and exactly one light chain. Within a cell, for the chain type with more than one sequence, the most abundant sequence in terms of UMI count (when tied, the sequence that appeared earlier in the file) was kept. Ultimately, exactly one heavy chain and one light chain per cell were kept. Additionally, quality control against cross-sample contamination was performed by examining the extent, if any, of pairwise overlapping between samples in terms of BCRs with both identical UMIs and identical non-UMI nucleotide sequences.

Individualized genotypes were inferred based on sequences that passed all quality control using TIgGER v.1.0.0[36] and used to finalize V(D)J annotations. Sequences annotated as non-productively rearranged by IgBLAST were removed from further analysis.

Clonal lineage inference

B cell clonal lineages were inferred on a by-individual basis based on productively rearranged sequences using hierarchical clustering with single linkage[37]. When combining both bulk and single-cell BCRs, heavy chain-based clonal inference was performed[38]. First, heavy chain sequences were partitioned based on common V and J gene annotations and CDR3 lengths using the groupGenes function from Alakazam v1.1.0[35]. Within each partition, heavy chain sequences with CDR3s that were within 0.15 normalized Hamming distance from each other were clustered as clones using the hclust function from fastcluster v1.2.3[39]. When using only single-cell BCRs, clonal inference was performed based on paired heavy and light chains. First, paired heavy and light chains were partitioned based on common V and J gene annotations and CDR3 lengths. Within each partition, pairs whose heavy chain CDR3s were within 0.15 normalized Hamming distance from each other were clustered as clones.

Following clonal inference, full-length clonal consensus germline sequences were reconstructed using CreateGermlines.py from Change-O v.1.0.2[35] for each clone with the D-segment (for heavy chains) and the N/P regions masked with Ns, resolving any ambiguous gene assignments by majority rule. Within each clone, duplicate IMGT-aligned V(D)J sequences from bulk sequencing were collapsed using the collapseDuplicates function from Alakazam v1.1.0[35] except for duplicates derived from different time points, tissues, B cell compartments, or isotypes.

BCR analysis

BCR analysis was performed in R v4.1.0 with visualization performed using base R, ggplot2 v3.3.5[40], and GraphPad Prism v9.

For the B cell compartment label, gene expression-based cluster annotation was used for single-cell BCRs; FACS-based sorting was used in general for bulk BCRs, except that PB sorts from lymph nodes were labelled LNPCs, Week 5 IgDlo sorts from blood were labelled activated, and Week 7 IgDlo sorts from blood were labelled memory. For the time point label, one blood PB sample that pooled collections in both Week 4 and Week 5 was treated as Week 4; and one blood memory sort sample that pooled collections in both Week 29 and Week 30 was treated as Week 29. For analysis involving the memory compartment, the memory sequences were restricted to bulk-sequenced Week 29 memory sorts from blood.

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

Clonal overlap between B cell compartments was visualized using circlize v.0.4.13[41].

Somatic hypermutation (SHM) frequency was calculated for each heavy chain sequence by counting the number of nucleotide mismatches from the germline sequence in the variable segment leading up to the CDR3, while excluding the first 18 positions that could be error-prone due to the primers used for generating the mAb sequences. Calculation was performed using the calcObservedMutations function from SHazaM v.1.0.2[35].

Phylogenetic trees for S-specific clones containing BMPCs were constructed on a by-participant basis using IgPhyML v1.1.3[18] with the HLP19 model[42]. Only heavy chain sequences from Week 4 PB compartment, the GC B cell, LNPC, and MBC compartments up to and including Week 15, and the Week 29 or 40 BMPC compartment were considered. For clones with >100 sequences, subsampling was applied with probabilities proportional to the proportions of sequences from different compartments, in addition to keeping all sequences corresponding to synthesized mAbs and all BMPC sequences. Only subsampled sequences from the PB, LNPC, and BMPC compartments were used for eventual tree-building. Trees were visualized using ggtree v3.0.4[43].

Human housekeeping genes

A list of human housekeeping genes was compiled from the 20 most stably expressed genes across 52 tissues and cell types in the Housekeeping and Reference Transcript (HRT) Atlas v1.0[44]; 11 highly uniform and strongly expressed genes reported[45]; and some of the most commonly used housekeeping genes[46]. The final list includes 34 genes: *ACTB, TLE5 (AES), AP2M1, BSG, C1orf43, CD59, CHMP2A, CSNK2B, EDF1, EEF2, EMC7, GABARAP, GAPDH, GPI, GUSB, HNRNPA2B1, HPRT1, HSP90AB1, MLF2, MRFAP1, PCBP1, PFDN5, PSAP, PSMB2, PSMB4, RAB11B, RAB1B, RAB7A, REEP5, RHOA, SNRPD3, UBC, VCP*, and *VPS29*.

Processing of 10× Genomics single-cell 5′ gene expression data

Demultiplexed pair-end FASTQ reads were first preprocessed on a by-sample basis using the 'cellranger count' command from 10× Genomics' Cell Ranger v.6.0.1 for alignment against the GRCh38 human reference v.2020-A ('refdata-gex-GRCh38-2020-A'). To avoid a batch effect introduced by sequencing depth, the 'cellranger aggr' command was used to subsample from each sample so that all samples had the same effective sequencing depth, which was measured in terms of the number of reads confidently mapped to the transcriptome or assigned to the feature IDs per cell. Gene annotation on human reference chromosomes and scaffolds in Gene Transfer Format ('gencode.v32.primary_assembly.annotation.gtf') was downloaded (2021-06-02) from GENCODE v32[47], from which a biotype ('gene_type') was extracted for each feature. Quality control was performed as follows on the aggregate gene expression matrix consisting of 432,713 cells and 36,601 features using SCANPY v1.7.2[48] and Python v3.8.8. (1) To remove presumably lysed cells, cells with mitochondrial content greater than 12.5% of all transcripts were removed. (2) To remove likely doublets, cells with more than 8,000 features or 80,000 total UMIs were removed. (3) To remove cells with no detectable expression of common endogenous genes, cells with no transcript for any of the 34 housekeeping genes were removed. (4) The feature matrix was subset, based on their biotypes, to protein-coding, immunoglobulin, and T cell receptor genes that were expressed in at least 0.1% of the cells in any sample. The resultant feature matrix contained 15,842 genes. (5) Cells with detectable expression of fewer than 200 genes were removed. After quality control, there were a total of 383,708 cells from 56 single-cell samples (Extended Data Table 5).

Single-cell gene expression analysis

Single-cell gene expression analysis was performed in SCANPY v1.7.2[48]. UMI counts measuring gene expression were log-normalized. The top 2,500 highly variable genes (HVGs) were identified using the 'scanpy.pp.highly_variable_genes' function with the 'seurat_v3' method, from which immunoglobulin and T cell receptor genes were removed. The data were scaled and centred, and principal component analysis (PCA) was performed based on HVG expression. PCA-guided neighborhood graphs embedded in Uniform Manifold Approximation and Projection (UMAP) were generated using the top 20 principal components via the 'scanpy.pp.neighbors' and 'scanpy.tl.umap' functions.

Overall clusters (Extended Data Table 3, **top**) were identified using Leiden graph-clustering via the 'scanpy.tl.leiden' function with resolution 0.23 (Extended Data Fig. 2b). UMAPs were faceted by batch, by participant, and by participant followed by sample; and inspected for convergence across batches, participants, and samples within participants, to assess whether there was a need for integration (Extended Data Fig. 2b). Cluster identities were assigned by examining the expression of a set of marker genes for different cell types (Extended Data Fig. 2c):

MS4A1, *CD19* and *CD79A* for B cells; *CD3D, CD3E, CD3G, IL7R* and *CD4* or *CD8A* for CD4$^+$ or CD8$^+$ T cells, respectively; *GZMB, GNLY, NKG7* and *NCAM1* for natural killer (NK) cells; *CD14, LYZ, CST3* and *MS4A7* for monocytes; *IL3RA* and *CLEC4C* for plasmacytoid dendritic cells (pDCs); and *FDCSP, CXCL14*[15] and *FCAMR*[16] for follicular dendritic cells (FDCs). One group of 27 cells labelled 'B and T' was excluded. To remove potential contamination by platelets, 73 cells with a log-normalized expression value of >2.5 for PPBP were removed. All 644 cells from the FDC cluster were confirmed to have originated from FNA samples instead of blood.

Cells from the overall B cell cluster (Extended Data Table 3, **bottom**) were further clustered to identify B cell subsets using Leiden graph-clustering via the 'scanpy.tl.leiden' function with resolution 0.18 (Extended Data Fig. 2d). Cluster identities were assigned by examining the expression of a set of marker genes for different B cell subsets (Extended Data Fig. 2e) along with the availability of BCRs. The following marker genes were examined: *BCL6, RGS13, MEF2B, STMN1, ELL3* and *SERPINA9* for GC B cells; *XBP1, IRF4, SEC11C, FKBP11, JCHAIN* and *PRDM1* for PBs and LNPCs; *TCL1A, IL4R, CCR7, IGHM*, and *IGHD* for naive B cells; and *TNFRSF13B, CD27* and *CD24* for MBCs. Although one group clustered with B cells during overall clustering, it was labelled 'B and T' as its cells tended to have both BCRs and relatively high expression levels of *CD2* and *CD3E*; and was subsequently excluded from the final B cell clustering. 18 cells that were found in the GC B cell cluster but came from blood samples were labelled 'PB-like'. 223 cells that were found in the PB cluster but came from FNA samples were re-assigned as LNPCs. 40 cells that were found in the LNPC cluster but came from blood samples were re-assigned as PBs. Heavy chain SHM frequency and isotype usage of the B cell subsets were assessed for consistency with expected values to further confirm their assigned identities.

Selection of single-cell BCRs from GC B cell or LNPC clusters for expression

Single-cell gene expression analysis was performed using lymph node samples up to and including Week 15 on a by-participant basis. Clonal inference was performed based on paired heavy and light chains from the same samples. From every clone with a clone size of >3 cells that contained cells from the GC B cell and/or LNPC clusters, one GC B cell or LNPC was selected. For selection, where a clone spanned both the GC B cell and LNPC compartments, and/or multiple time points, a compartment and a timepoint were first randomly selected. Within that clone, the cell with the highest heavy chain UMI count was then selected, breaking ties based on *IGHV* SHM frequency. In all selected cells, native pairing was preserved.

Selection of BCRs from S-specific BMPC clones for expression

From each heavy chain-based S-specific clone containing both PBs and BMPCs, where possible, one PB heavy chain was selected, and, together with all BMPC heavy chains, were paired with the same light chain for expression. For the PB heavy chain, if single-cell paired PBs were available, the single-cell paired PB whose *IGHV* mutation frequency was closest to the median mutation frequency of other single-cell paired PBs in the same clone (breaking ties by UMI count), and whose light chain V gene, J gene, and CDR3 length (VJL) combination was consistent with the clonal majority, was used as the source. The natively paired light chain of the PB from which the heavy chain was selected was used. In clones in which two PBs had inconsistent light chain VJL combinations, both PBs were used. Clones in which there was light chain uncertainty due to more than two PBs or due to LNPCs were generally excluded.

Curation of selected BCRs for expression

The selected BCRs were curated prior to synthesis. First, artificial gaps introduced under the IMGT unique numbering system[49] were removed from the IMGT-aligned observed V(D)J sequences. IMGT gaps were identified as positions containing in-frame triplet dots ('...') in the IMGT-aligned germline sequences. Second, any non-informative (non-A/T/G/C, such as N or -) positions in the observed sequences, with the exception of potential in-frame indels, were patched by the nucleotides at their corresponding germline positions. Third, if applicable, the 3' end of the observed sequences were trimmed so that the total nucleotide length would be a multiple of 3. Finally, potential in-frame indels were manually reviewed. For a given potential in-frame indel from a selected cell, its presence or lack thereof in the unselected cells from the same clone was considered. Barring strong indications that an in-frame indel was due to sequencing error rather than the incapability of the IMGT unique numbering system of capturing it, the in-frame indels were generally included in the final curated sequences.

Transfection for recombinant mAbs and Fab production

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

GFP-reduction neutralization test

Serial dilutions of each mAb diluted in DMEM were incubated with 10^2 plaque-forming unit (PFU) of VSV-SARS-CoV-2 D614G for 1 h at 37 °C. Antibody-virus complexes were added to Vero cell monolayers in 96-well plates and incubated at 37 °C for 7.5 h. Cells were fixed at room temperature in 2% formaldehyde (Millipore Sigma) containing 10 ug/mL of Hoechst 33342 nuclear stain (Invitrogen) for 45 min at room temperature. Fixative was replaced with PBS prior to imaging. Images were acquired using an IN Cell 2000 Analyzer automated microscope (GE Healthcare) in both the DAPI and FITC channels to visualize nuclei and infected cells. Images were analyzed using the Multi Target Analysis Module of the IN Cell Analyzer 1000 Workstation Software (GE Healthcare). GFP-positive cells were identified using the top-hat segmentation method and subsequently counted within the IN Cell workstation software. The initial dilution of mAb started at 25 ug/mL and was three-fold serially diluted in 96-well plate over eight dilutions.

Affinity analysis via biolayer interferometry (BLI)

We used the Octet Red instrument (ForteBio) with shaking at 1,000 r.p.m. The kinetic analysis using Octet SA biosensors (Sartorius) was performed as follows: (1) Baseline: 120 sec immersion in buffer (10mM HEPES and 1% BSA). (2) Loading: 130 sec immersion in solution with biotinylated SARS-CoV-2 S Avitag 10 ug/ml. (3) Baseline: 120 sec immersion in buffer. (4) Association: 300 sec immersion in solution with serially diluted recombinant Fab. (5) Dissociation: 600 sec immersion in buffer. The BLI signal was recorded and analyzed using BIAevaluation Software (Biacore). The 1:1 binding model with a drifting baseline was employed for the equilibrium dissociation constant (K_D).

Reporting summary

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

Data availability

Raw sequencing data and transcriptomics count matrix are deposited at Sequence Read Archive and Gene Expression Omnibus respectively under BioProject PRJNA777934. Processed transcriptomics and BCR data are deposited at https://doi.org/10.5281/zenodo.5895181 on Zenodo. Previously reported bulk-sequenced BCR data used in this study were deposited under PRJNA731610 and PRJNA741267 on SRA, and at https://doi.org/10.5281/zenodo.5042252 and https://doi.org/10.5281/zenodo.5040099 on Zenodo.

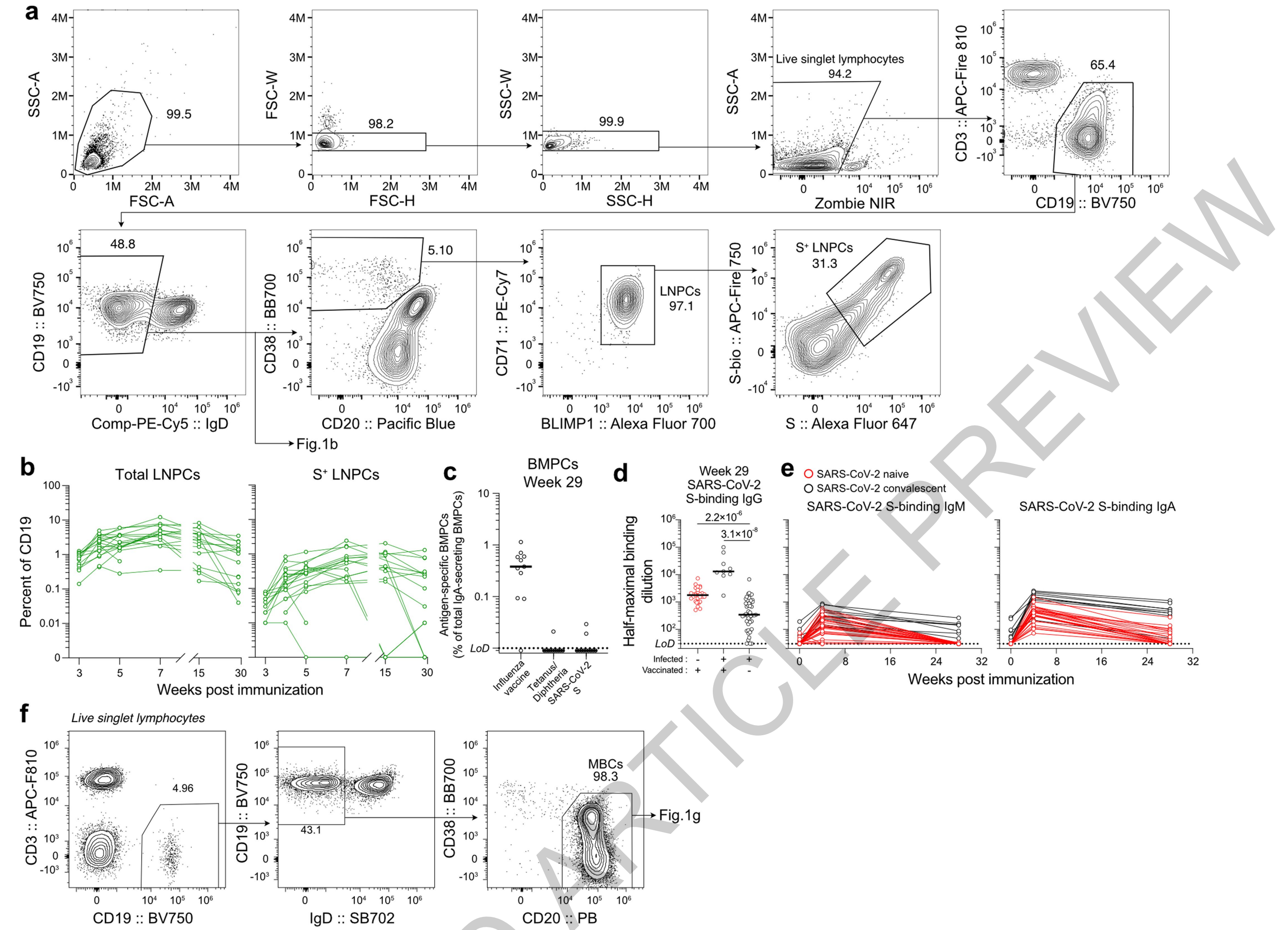
29. Stadlbauer, D. et al. SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. *Current Protocols in Microbiology* **57**, e100, https://doi.org/10.1002/cpmc.100 (2020).
30. Fairhead, M. & Howarth, M. Site-specific biotinylation of purified proteins using BirA. *Methods Mol Biol* **1266**, 171–184, https://doi.org/10.1007/978-1-4939-2272-7_12 (2015).
31. Davis, C. W. et al. Longitudinal Analysis of the Human B Cell Response to Ebola Virus Infection. *Cell* **177**, 1566–1582.e1517, https://doi.org/10.1016/j.cell.2019.04.036 (2019).
32. Vander Heiden, J. A. et al. pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics* **30**, 1930–1932, https://doi.org/10.1093/bioinformatics/btu138 (2014).
33. Ye, J., Ma, N., Madden, T. L. & Ostell, J. M. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Research* **41**, W34–W40, https://doi.org/10.1093/nar/gkt382 (2013).
34. Giudicelli, V., Chaume, D. & Lefranc, M. P. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res* **33**, D256–261, https://doi.org/10.1093/nar/gki010 (2005).
35. Gupta, N. T. et al. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data: Table 1. *Bioinformatics* **31**, 3356–3358, https://doi.org/10.1093/bioinformatics/btv359 (2015).
36. Gadala-Maria, D., Yaari, G., Uduman, M. & Kleinstein, S. H. Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. *Proceedings of the National Academy of Sciences* **112**, E862–E870, https://doi.org/10.1073/pnas.1417683112 (2015).
37. Gupta, N. T. et al. Hierarchical Clustering Can Identify B Cell Clones with High Confidence in Ig Repertoire Sequencing Data. *The Journal of Immunology* **198**, 2489–2499, https://doi.org/10.4049/jimmunol.1601850 (2017).
38. Zhou, J. Q. & Kleinstein, S. H. Cutting Edge: Ig H Chains Are Sufficient to Determine Most B Cell Clonal Relationships. *The Journal of Immunology* **203**, 1687–1692, https://doi.org/10.4049/jimmunol.1900666 (2019).
39. Müllner, D. fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. *Journal of Statistical Software* **53**, 1–18, https://doi.org/10.18637/jss.v053.i09 (2013).
40. Wickham, H. ggplot2: Elegant graphics for data analysis. *Springer-Verlag New York* (2016).
41. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812, https://doi.org/10.1093/bioinformatics/btu393 (2014).
42. Hoehn, K. B. et al. Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination. *Proceedings of the National Academy of Sciences* **116**, 22664–22672, https://doi.org/10.1073/pnas.1906020116 (2019).
43. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**, 28–36, https://doi.org/10.1111/2041-210x.12628 (2017).
44. Hounkpe, B. W., Chenou, F., Franciele, & Erich, HRT Atlas v1.0 database: redefining human and mouse housekeeping genes and candidate reference transcripts by mining massive RNA-seq datasets. *Nucleic Acids Research* **49**, D947–D955, https://doi.org/10.1093/nar/gkaa609 (2021).
45. Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends in Genetics* **29**, 569–574, https://doi.org/10.1016/j.tig.2013.05.010 (2013).
46. Valente, V. et al. Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR. *BMC Molecular Biology* **10**, 17, https://doi.org/10.1186/1471-2199-10-17 (2009).
47. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research* **47**, D766–D773, https://doi.org/10.1093/nar/gky955 (2019).
48. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**, 15, https://doi.org/10.1186/s13059-017-1382-0 (2018).
49. Lefranc, M.-P. et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Developmental & Comparative Immunology* **27**, 55–77, https://doi.org/10.1016/s0145-305x(02)00039-3 (2003).

Acknowledgements We thank the generous study participants for providing these precious specimens. We thank Dr. Brian J. Laidlaw for critically reading of the manuscript and Dr. Christopher A. Nelson for help with the binding affinity analyses. The Ellebedy laboratory was supported by National Institute of Allergy and Infectious Diseases (NIAID) grants U01AI141990 and 1U01AI150747, NIAID Centers of Excellence for Influenza Research and Surveillance contract HHSN272201400006C and HHSN272201400008C, and NIAID Collaborative Influenza Vaccine Innovation Centers contract 75N93019C00051. W.K. was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1A6A3A03041509). J.S.T. was supported by NIAID 5T32CA009547. This study utilized samples obtained from the Washington University School of Medicine's COVID-19 biorepository supported by the NIH/National Center for Advancing Translational Sciences, grant number UL1 TR002345. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official view of NIAID or NIH. The WU353 and WU368 studies were reviewed and approved by the Washington University Institutional Review Board (approval no. 202003186 and 202012081, respectively).

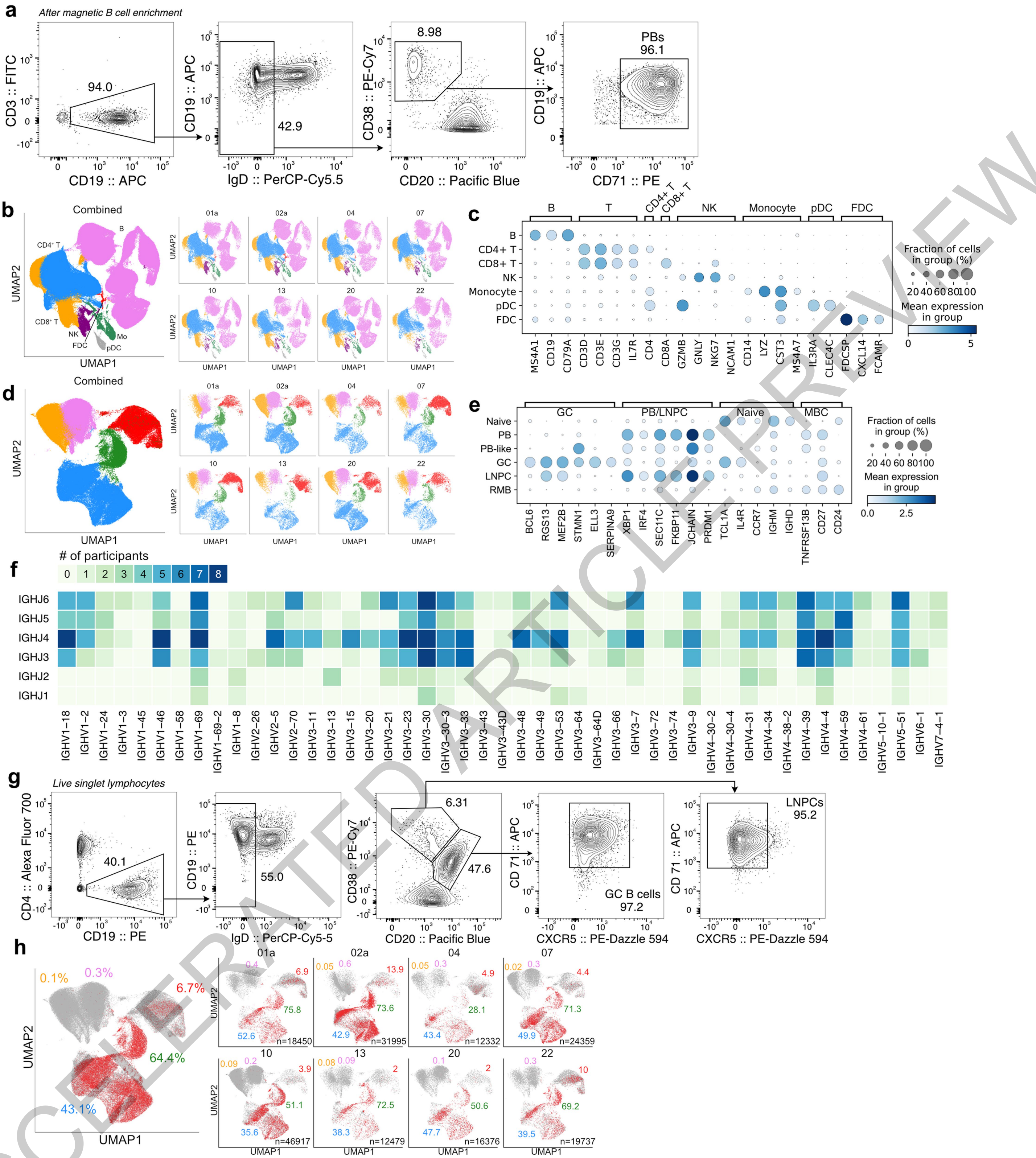
Author contributions A.H.E. conceived and designed the study. A.H., M.K.K., L.P., P.A.M., I.P., J.A.O., and R.M.P. wrote and maintained the IRB protocol, recruited, and phlebotomized participants, coordinated sample collection, and analyzed clinical data. W.K., E.K, and J.S.T. processed specimens. W.K., E.K., W.B.A., and J.S.T. performed ELISA and ELISpot. W.K., S.C.H, A.J.Schmitz, T.L., M.T., and W.B.A. generated and characterized monoclonal antibodies. W.K.and A.J.Sturtz prepared libraries for scRNA sequencing. A.J.Schmitz performed RNA extractions and library preparation for bulk BCR sequencing. J.Q.Z. analyzed scRNA sequencing and bulk BCR sequencing data. A.J.Schmitz expressed SARS-CoV-2 S proteins. J.S.T. sorted cells and collected and analysed the flow cytometry data. T.S. and W.D.M performed FNA. W.D.M. and S.A.T. supervised lymph node evaluation prior to FNA and specimen collection and evaluated lymph node ultrasound data. Z.L. and S.P.J.W performed and analyzed *in vitro* neutralization assay. W.K., J.Q.Z, J.S.T., and A.H.E. analyzed the data. A.H.E. supervised experiments and obtained funding. W.K., J.Q.Z, P.A.M, J.S.T., and A.H.E. composed the manuscript. All authors reviewed the manuscript.

Competing interests The Ellebedy laboratory received funding under sponsored research agreements from Emergent BioSolutions and from AbbVie. J.S.T. is a consultant for Gerson Lehrman Group.

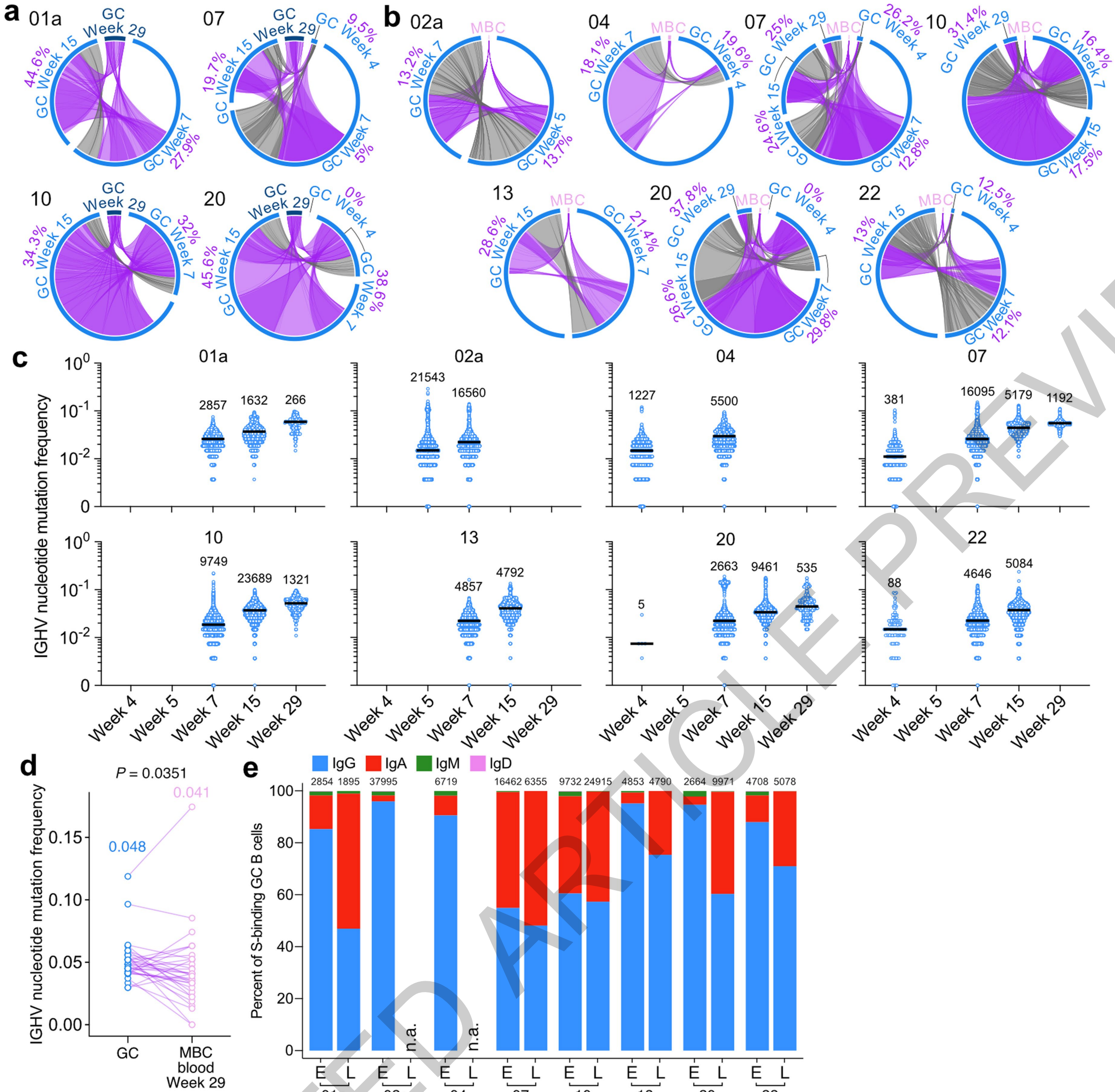
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-04527-1.
Correspondence and requests for materials should be addressed to Ali H. Ellebedy.
Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permissions information is available at http://www.nature.com/reprints.



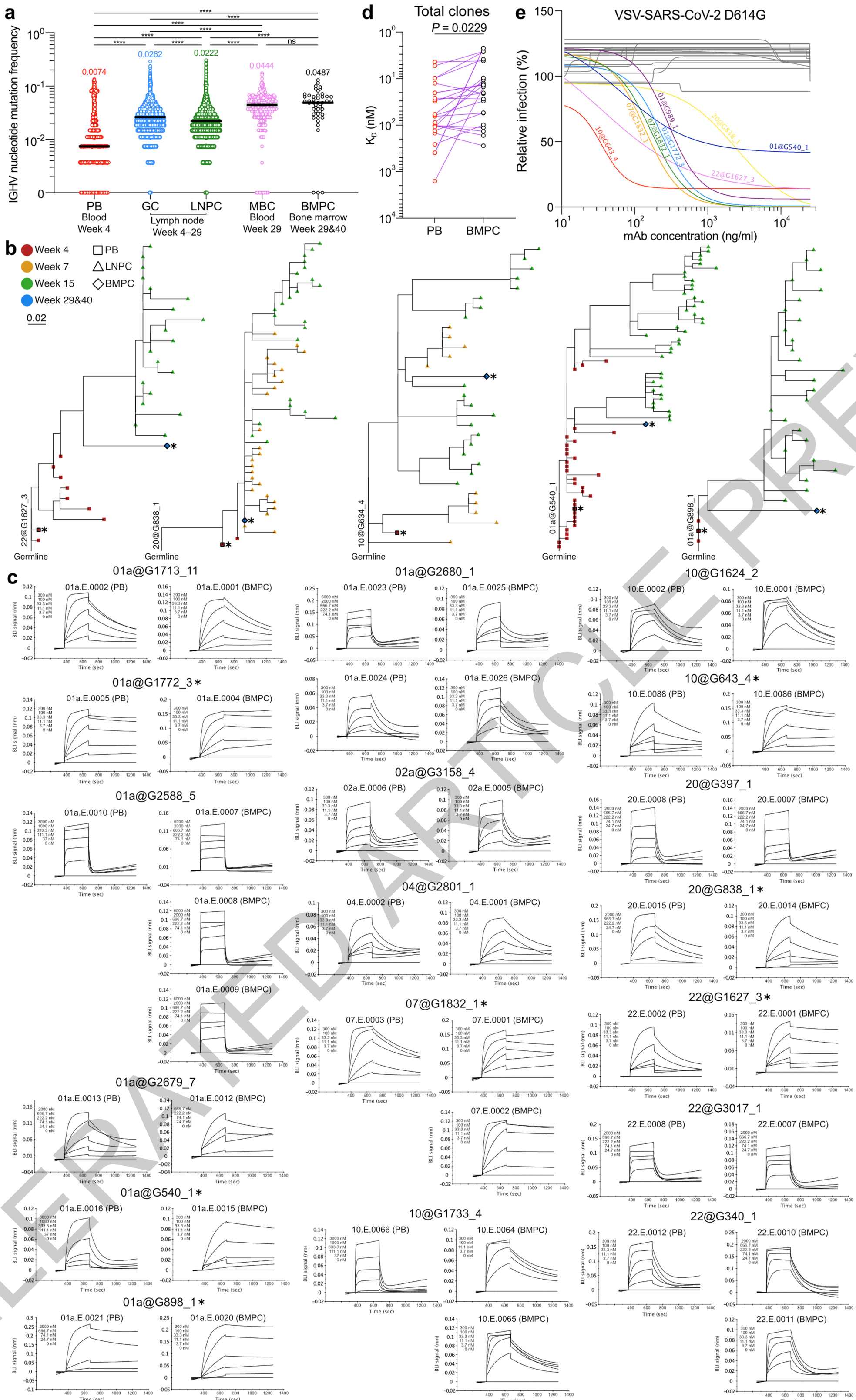
Extended Data Fig. 1 | Persistence of humoral immune responses to SARS-CoV-2 mRNA vaccination. a, Flow cytometry gating strategies for GC B cells (Fig. 1b) and LNPCs (defined as CD19^{+} CD3^{-} IgD^{low} CD20^{low} CD38^{+} BLIMP1^{+} CD71^{+} live singlet lymphocytes) in the lymph node. **b**, Kinetics of total (left) and S-specific LNPCs (right) as gated in **a**. **c**, Frequencies of BMPCs secreting IgA antibodies specific for the indicated antigens 29 weeks after immunization. Symbols represent one sample in **b** (n=15) and **c** (n=11). **d**, **e**, Plasma antibody titers against SARS-CoV-2 S measured by ELISA in participants without (red, n=29) and with (black, n=9) a history of SARS-CoV-2 infection in SARS-CoV-2 vaccinated (left, center) and unvaccinated (right, n=48) participants 29 weeks after the first vaccine dose or symptom onset (**d**) and in vaccinated participants (red, n=29; black, n=9) over time (**e**). *P* values were determined by Kruskal-Wallis test followed by Dunn's multiple comparison test between unvaccinated and both vaccinated groups (**d**), and by two-sided Mann-Whitney test (**e**). Horizontal lines indicate median values and dotted lines indicate detection limit in **c** and **e**. **f**, Flow cytometry gating strategies for MBCs (CD19^{+} CD3^{-} IgD^{low} CD20^{+} CD38^{-} live singlet lymphocytes) and S-binding MBCs (Fig. 1g) in blood.



Extended Data Fig. 2 | Identification of SARS-CoV-2 S-binding B cell clones in the lymph node. a, Flow cytometry gating strategies for sorting PBs (defined as CD19^{+} CD3^{-} IgD^{low} CD20^{low} CD38^{+} CD71^{+} live singlet lymphocytes) from blood. **b**, **d** UMAPs showing scRNA-seq transcriptional clusters of total cells (**b**) and of B cells (**d**) from PBs sorted from blood and FNA of draining axillary lymph nodes combined. **c, e**, Dot plots for the marker genes used for identifying annotated clusters. **f**, Heatmap of paired *IGHV* and *IGHJ* gene usage in S-binding clones. Color indicates the number of participants in which clones using a combination of *IGHV* and *IGHJ* genes were found. **g**, Flow cytometry gating strategies for sorting GC B cells (CD19^{+} CD4^{-} IgD^{low} CD20^{+} CD38^{int} CXCR5^{high} CD71^{+} live singlet lymphocytes) and LNPCs (CD19^{+} CD4^{-} IgD^{low} CD20^{low} CD38^{+} CXCR5^{low} CD71^{+} live singlet lymphocytes) from FNAs. **h**, SARS-CoV-2 S-binding clones visualized in red on UMAP of B cell clusters. Percentages are of S-binding clones within GC B cells (blue), LNPCs (green), PBs (red), MBCs (pink) or naive B cells (yellow). Total numbers of cells are at the bottom right corner.



Extended Data Fig. 3 | Maturation of SARS-CoV-2 S-binding B cells in the lymph node. a, Circos diagrams showing clonal overlap between S-binding GC B cells at indicated time points. Purple and grey chords correspond to, respectively, clones spanning both 29 weeks post-vaccination and other time points, and clones spanning one or more of 4, 7 and 15 weeks post-vaccination. Percentages are of GC B cell clones related to GC B cells detected at 29 weeks post-vaccination. **b**, Circos diagrams showing clonal overlap between S-binding MBCs in blood 29 weeks post-vaccination and GC B cells at indicated time points. Purple and grey chords correspond to, respectively, clones spanning both the MBC and GC B cell compartments, and clones spanning only the GC B cell compartment. Percentages are of GC B cell clones overlapping with MBCs in blood 29 weeks post-vaccination. Arc length corresponds to the number of BCR sequences and chord with corresponds to clone size in **a** and **b**. **c**, Comparison of *IGHV* nucleotide mutation frequency of SARS-CoV-2 S-binding GC B cells in each participant at the indicated time points. Horizontal lines represent median values. Cell numbers are presented on the top of each data set. **d**, Comparison of *IGHV* region nucleotide mutation frequencies between clonally related, SARS-CoV-2 S-binding GC B cells and MBCs (n=33) detected at 29 weeks post-vaccination. Each dot represents the median SHM frequency of a clone within the indicated compartment. Median values are presented on the top of each data set. *P* value was determined by a paired two-sided non-parametric Mann-Whitney test. **e**, Percentages of GC B cells expressing BCRs of isotype IgG (blue), IgA (red), IgM (green) or IgD (pink) at the early (E) or the late (L) time point. The early and late time points represent, respectively, 4, 5 or 7 weeks, and 15 or 29 weeks after immunization. Cell numbers are at the top.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Evolution of B cell clones induced by SARS-CoV-2 vaccination. **a**, **a**, Comparison of *IGHV* nucleotide mutation frequency of PBs (n=2735), GC B cells (n=139322), LNPCs (n=82350), MBCs (n=341) and BMPCs (n=47). Horizontal lines represent median values. *P* values were determined by Kruskal-Wallis test followed by Dunn's multiple comparison test. **b**, Phylogenetic trees of neutralizing clones showing inferred evolutionary relationships between PBs (squares), LNPCs (triangles) and BMPCs (diamonds). Horizontal branch length represents the expected number of substitutions per codon in V-region genes, corresponding to the scale bar. Clone IDs are displayed near the root of the trees. Asterisks denote neutralizing mAbs. **c**, Kinetic curves of BLI signal for clonally related, PB- and BMPC-derived Fabs interacting with immobilized SARS-CoV-2 S. Clone IDs, Fab IDs and cell types are presented on the top of each data set. Asterisks denote neutralizing clones. **d**, Equilibrium dissociation constant (K_D) of Fabs (n=24) interacting with immobilized SARS-CoV-2 S measured by biolayer interferometry (BLI). Red and black dots indicate K_D values of clonally related, PB- and BMPC-derived Fabs, respectively. *P* value was determined by Wilcoxon matched-pair signed rank test. **e**, Neutralization curves of VSV-SARS-CoV-2 D614G with BMPC-derived mAbs. Colored and grey lines represent neutralizing and non-neutralizing clones, respectively. Neutralizing clone IDs are indicated on each curve. ns > 0.9999, ****$P < 0.0001$.

Extended Data Table 1 | Demographics of participants and vaccine side effects

Variable	SARS-CoV-2 mRNA vaccination				Variable	Convalescent (N=13) N (%)
	Total (N=43) N (%)	Blood (N=42) N (%)	Lymph node (N=15) N (%)	Bone marrow (N=11) N (%)		
Age (median [range])	38 (28-73)	37.5 (28-73)	37 (28-52)	36 (28-48)	**Age (median [range])**	50 (21-69)
Sex					**Sex**	
Female	21 (48.8)	20 (47.6)	7 (53.8)	6 (54.5)	Female	23 (47.9)
Male	22 (51.2)	22 (52.4)	6 (46.2)	5 (45.5)	Male	25 (52.1)
Race					**Race/Ethnicity**	
White	34 (79.1)	34 (81)	12 (80)	11 (100)	White	47 (97.9)
Black	6 (14)	6 (14.3)	2 (13.3)	0 (0)	Black	0 (0)
Asian	1(2.3)	0 (0)	1 (6.7)	0 (0)	Asian	1 (2.1)
Other	2 (4.7)	2 (4.8)	0 (0)	0 (0)	Hispanic	0 (0)
Ethnicity						
Not of Hispanic, Latinx, or Spanish origin	41 (95.3)	40 (95.2)	14 (93.3)	10 (90.9)		
Hispanic, Latinx, Spanish origin	2 (4.7)	2 (4.8)	1 (6.7)	1 (9.1)		
BMI (media [range])	26.8 (21.4-67.4)	26.9 (21.4-67.4)	24.1 (21.4-40.1)	23.9 (21.4-40.1)		
Comorbidities					**Comorbidities**	
Lung disease	1 (2.3)	1 (2.4)	0 (0)	0 (0)	Asthma	10 (20.8)
Diabetes mellitus	0 (0)	0 (0)	0 (0)	0 (0)	Other lung disease	0 (0)
Hypertension	7 (16.3)	6 (14.3)	2 (13.3)	1 (9.1)	Heart disease	0 (0)
Cardiovascular	0 (0)	0 (0)	0 (0)	0 (0)	Hypertension	7 (14.6)
Liver disease	0 (0)	0 (0)	0 (0)	0 (0)	Diabetes mellitus	3 (6.3)
Chronic kidney disease	0 (0)	0 (0)	0 (0)	0 (0)	Cancer	6 (12.5)
Cancer on chemotherapy	0 (0)	0 (0)	0 (0)	0 (0)	Autoimmune disease	4 (8.3)
Hematological malignancy	0 (0)	0 (0)	0 (0)	0 (0)	Hyperlipidaemia	2 (4.2)
Pregnancy	0 (0)	0 (0)	0 (0)	0 (0)	GERD	5 (10.4)
Neurological	0 (0)	0 (0)	0 (0)	0 (0)	Other	16 (33.3)
HIV	0 (0)	0 (0)	0 (0)	0 (0)	Solid Organ Transplant	1 (2.1)
Hyperlipidemia	1 (2.3)	1 (2.4)	0 (0)	0 (0)		
Gastroesophageal reflux disease	-	-	-	-		
Autoimmune disease	-	-	-	-		
Solid organ transplant recipient	0 (0)	0 (0)	0 (0)	0 (0)		
Bone marrow transplant recipient	0 (0)	0 (0)	0 (0)	0 (0)		
Confirmed SARS-CoV-2 infection	13 (30.2)	13 (31)	0 (0)	0 (0)	**Hospitalized for treatment of COVID-19**	3 (6.3)
Time from SARS-CoV-2 infection to baseline visit in days (median [range])	234 (70-370)	234 (70-370)	-	-		

Variable	Total N=43 N (%)	Total N=43 N (%)
	First dose	**Second dose**
None	2 (4.7)	0 (0)
Chills	9 (20.9)	18 (41.9)
Fever	6 (14)	12 (27.9)
Headache	10 (23.3)	16 (37.2)
Injection site pain/redness/swelling	37 (86)	39 (90.7)
Muscle or joint pain	10 (23.3)	25 (58.1)
Fatigue	13 (30.2)	27 (62.8)
	Duration of side effects in hours (median	
Chills	18 (6-72)	24 (0.2-48)
Fever	8 (3-72)	30 (1-48)
Headache	36 (6-120)	24 (4-72)
Injection site pain	48 (0.2-168)	48 (2-144)
Muscle or joint pain	24 (3-72)	24 (1-48)
Fatigue	48 (12-120)	24 (3-144)

Extended Data Table 2 | Frequencies of GC B cells, LNPCs and CD14⁺ myeloid cells in draining axillary lymph nodes

Participant	LN #	Total GC B cells (% of CD19)						SARS-CoV-2 S-binding GC B cells (% of CD19)					
		Week 3	Week 4	Week 5	Week 7	Week 15	Week 29	Week 3	Week 4	Week 5	Week 7	Week 15	Week 29
01a	1	15.25	13.69	11.14	31.67	29.13	14.37	3.23	5.41	4.01	13.75	18.06	8.00
02a	1	8.28	34.16	44.92	21.90	7.17	0.85	1.36	9.66	11.55	5.98	0.01	0.06
	2		14.08	13.54	23.13				4.66	3.85	9.33		
04	1	4.71	14.02	11.21	43.94	3.75	0.38	0.57	2.84	5.59	6.37	0.70	0.00
	2					1.09	0.41					0.03	0.00
07	1	21.14	19.68	11.48	39.22	28.93	28.10	4.59	4.92	3.87	14.55	19.92	14.77
08	1	9.91	1.31	3.99	12.15		31.34	1.21	0.41	0.91	4.21		23.33
10	1	7.72	5.92	2.82	7.34	19.33	19.41	1.10	1.47	0.90	3.10	14.19	15.37
	2		5.70	3.40	4.28	7.14	9.19		1.36	1.04	1.05	2.79	5.90
13	1	14.99	8.69	7.23	16.02	20.21	0.77	3.53	3.00	2.14	3.70	10.03	0.00
15	1	13.06	26.89	44.56				2.96	8.01	9.58			
	2		0.23	0.61					0.01	0.02			
16	1	5.20	7.61	17.04		5.82	4.40	0.74	1.50	3.34		2.57	1.48
	2		1.22	7.44		9.61	3.64		0.24	1.84		5.81	1.06
20	1	0.41	7.49	2.10	0.52	0.47	0.71	0.03	3.07	0.31	0.04	0.04	0.02
	2		7.79	3.36	13.67	16.93	20.23		1.28	0.72	4.43	11.08	10.56
21	1		20.18		14.50		7.03		5.92		5.20		3.33
22	1	24.69	20.63	19.66	25.86	22.01	0.72	4.44	5.35	4.92	7.33	9.69	0.00
26	1						1.69						0.29
28	1	6.40	6.25			12.82	9.46	1.12	1.16			7.26	4.73
43	1	29.01	29.38	26.09	29.43	35.19		4.43	9.42	6.28	15.26	15.38	

Participant	LN #	Total LNPCs (% of CD19)						SARS-CoV-2 S-binding LNPCs (% of CD19)					
		Week 3	Week 4	Week 5	Week 7	Week 15	Week 29	Week 3	Week 4	Week 5	Week 7	Week 15	Week 29
01a	1	0.77	4.96	6.00	6.82	8.23	2.33	0.07	0.84	1.15	1.61	2.04	0.78
02a	1	0.35	2.47	3.52	3.20	0.43	0.22	0.01	0.20	0.40	0.36	0.00	0.00
	2		2.30	3.74	3.91				0.25	0.28	0.54		
04	1	0.88	1.29	1.57	2.67	0.29	0.08	0.05	0.09	0.43	0.20	0.00	0.01
	2					0.17	0.14					0.00	0.01
07	1	0.97	2.74	1.73	3.97	2.02	1.87	0.07	0.26	0.26	0.67	0.32	0.22
08	1	1.07	2.18	1.53	4.84		3.49	0.07	0.09	0.15	0.74		1.30
10	1	0.71	4.08	2.31	3.92	4.02	2.42	0.04	0.34	0.33	0.86	1.06	0.76
	2		2.83	3.51	1.75	1.08	1.40		0.18	0.40	0.16	0.19	0.31
13	1	0.39	0.94	1.57	5.25	3.02	0.04	0.02	0.21	0.30	0.64	0.44	0.00
15	1	1.26	1.27	3.60				0.08	0.12	0.29			
	2		1.79	0.64					0.10	0.01			
16	1	0.47	0.88	2.24		1.68	0.60	0.02	0.05	0.12		0.16	0.03
	2		0.43	1.31		2.91	1.06		0.02	0.11		0.43	0.10
20	1	0.14	0.46	0.28	0.35	0.34	0.09	0.01	0.01	0.00	0.00	0.00	0.00
	2		1.73	1.62	4.01	2.48	1.32		0.05	0.08	0.22	0.42	0.29
21	1		3.42		12.18		1.53		0.38		1.70		0.25
22	1	1.01	2.92	5.91	7.84	5.84	0.07	0.03	0.16	0.30	0.77	0.82	0.00
26	1						0.23						0.00
28	1	0.54	1.54			6.26	1.29	0.03	0.13			1.15	0.25
43	1	1.14	3.85	3.66	7.75	3.63		0.05	0.63	0.55	2.44	0.93	

Participant	LN #	CD14 (% of live singlet)					
		Week 3	Week 4	Week 5	Week 7	Week 15	Week 29
01a	1	0.21	0.06	0.15	0.15	0.16	0.21
02a	1	0.16	0.17	0.07	0.25	0.17	0.30
	2		0.14	0.14	0.71		
04	1	0.19	0.37	0.72	0.14	1.00	0.16
	2					0.33	0.25
07	1	0.19	0.15	0.42	0.66	0.04	4.38
08	1	0.27	0.17	0.34	0.81		1.01
10	1	0.22	0.07	0.08	0.21	0.03	0.23
	2		0.09	0.05	0.13	0.03	0.26
13	1	0.57	0.34	0.43	0.33	1.63	1.04
15	1	0.07	0.19	0.08			
	2		0.14	0.24			
16	1	0.26	0.14	0.06		1.21	0.50
	2		0.11	0.14		0.51	0.32
20	1	0.21	0.15	0.29	0.41	0.84	0.11
	2		0.11	0.16	0.11	0.29	0.16
21	1		0.13		0.11		0.51
22	1	0.18	0.16	0.20	0.36	0.28	0.24
26	1						0.03
28	1	0.14	0.22			0.66	0.66
43	1	0.52	0.32	0.98	0.09	0.01	

Extended Data Table 3 | Cell counts and frequencies of transcriptional clusters and of SARS-CoV-2 S binding cells in scRNA-seq of PBs from blood and FNA from lymph nodes

Sample	Overall cluster	Cell count (% of whole cells)	B cell cluster	Cell count (% of whole B cell clusters)
PBs	B	27420 (96.5%)	PB	27231 (99.3%)
	CD4$^+$ T	49 (0.2%)	MBC	150 (0.5%)
	CD8$^+$ T	81 (0.3%)	Navie	21 (0.1%)
	NK	718 (2.5%)	PB-like	18 (0.1%)
	Monocyte	141 (0.5%)		
	pDC	5 (0.02%)		
Lymph node	B	166022 (46.7%)	GC B cells	62156 (40%)
	CD4$^+$ T	136929 (38.6%)	LNPC	12299 (7.9%)
	CD8$^+$ T	36532 (10.3%)	MBC	42105 (27.1%)
	NK	6268 (1.8%)	Naïve	38665 (24.9%)
	Monocyte	6379 (1.8%)		
	pDC	2420 (0.7%)		
	FDC	644 (0.2%)		
Combined	B	193442 (50.4%)	GC B cell	62156 (34%)
	CD4$^+$ T	136978 (35.7%)	LNPC	12299 (6.7%)
	CD8$^+$ T	36613 (9.5%)	PB	27231 (14.9%)
	NK	6986 (1.8%)	MBC	42255 (23.1%)
	Monocyte	6520 (1.7%)	Naïve	38686 (21.2%)
	pDC	2425 (0.6%)	PB-like	18 (0.01%)
	FDC	644 (0.2%)		

Participants	B cell cluster	Cell count (% of whole cells)	SARS-CoV-2 S-binding cell count (% in each B cell cluster)
01a	GC	4831 (26.2%)	2540 (52.6%)
	LNPC	1246 (6.8%)	944 (75.8%)
	Naïve	2622 (14.2%)	0 (0%)
	PB	6179 (33.5%)	424 (6.9%)
	PB-like	2 (0%)	0 (0%)
	MBC	3570 (19.3%)	14 (0.4%)
02a	GC	17894 (55.9%)	7679 (42.9%)
	LNPC	3092 (9.7%)	2277 (73.6%)
	Naïve	4127 (12.9%)	2 (0%)
	PB	1838 (5.7%)	255 (13.9%)
	PB-like	1 (0%)	0 (0%)
	MBC	5043 (15.8%)	32 (0.6%)
04	GC	3693 (29.9%)	1601 (43.4%)
	LNPC	395 (3.2%)	111 (28.1%)
	Naïve	2075 (16.8%)	1 (0%)
	PB	4364 (35.4%)	213 (4.9%)
	PB-like	0 (0%)	-
	MBC	1805 (14.6%)	6 (0.3%)
07	GC	9790 (40.2%)	4889 (49.9%)
	LNPC	1199 (4.9%)	855 (71.3%)
	Naïve	4707 (19.3%)	1 (0.02%)
	PB	3667 (15.1%)	161 (4.4%)
	PB-like	8 (0%)	0 (0%)
	MBC	4988 (20.5%)	16 (0.3%)
10	GC	12459 (26.6%)	4432 (35.6%)
	LNPC	3063 (6.5%)	1564 (51.1%)
	Naïve	16470 (35.1%)	14 (0.1%)
	PB	1507 (3.2%)	59 (3.9%)
	PB-like	2 (0%)	0 (0%)
	MBC	13416 (28.6%)	26 (0.2%)
13	GC	3639 (29.2%)	1393 (38.3%)
	LNPC	934 (7.5%)	677 (72.5%)
	Naïve	2602 (20.9%)	2 (0.1%)
	PB	1868 (15%)	38 (2%)
	PB-like	2 (0%)	0 (0%)
	MBC	3434 (27.5%)	3 (0.1%)
20	GC	4564 (27.9%)	2178 (47.7%)
	LNPC	806 (4.9%)	408 (50.6%)
	Naïve	4019 (24.5%)	0 (0%)
	PB	1177 (7.2%)	23 (2%)
	PB-like	0 (0%)	-
	MBC	5810 (35.5%)	7 (0.1%)
22	GC	5286 (26.8%)	2089 (39.5%)
	LNPC	1564 (7.9%)	1082 (69.2%)
	Naïve	2064 (10.5%)	0 (0%)
	PB	6631 (33.6%)	661 (10%)
	PB-like	3 (0%)	0 (0%)
	MBC	4189 (21.2%)	12 (0.3%)
Combined	GC	62156 (34%)	26801 (43.1%)
	LNPC	12299 (6.7%)	7918 (64.4%)
	Naïve	38686 (21.2%)	20 (0.1%)
	PB	27231 (14.9%)	1834 (6.7%)
	PB-like	18 (0%)	0 (0%)
	MBC	42255 (23.1%)	116 (0.3%)

Extended Data Table 4 | Description of SARS-CoV-2 S-binding mAbs derived from GC B cells and LNPCs

		01a (N=145) N (%)	02a (N=410) N (%)	04 (N=124) N (%)	07 (N=201) N (%)	10 (N=249) N (%)	13 (N=108) N (%)	20 (N=99) N (%)	22 (N=167) N (%)
Heavy chain isotype	IGHG	109 (75.2%)	387 (94.4%)	102 (82.3%)	146 (72.6%)	200 (80.3%)	102 (94.4%)	90 (90.9%)	144 (86.2%)
	IGHA	35 (24.1%)	11 (2.7%)	15 (12.1%)	54 (26.9%)	43 (17.3%)	5 (4.6%)	6 (6.1%)	22 (13.2%)
	IGHM	1 (0.7%)	11 (2.7%)	7 (5.6%)	1 (0.5%)	2 (0.8%)	1 (0.9%)	3 (3%)	1 (0.6%)
	IGHD	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Light chain isotype	IGKC	102 (70.3%)	223 (54.4%)	61 (49.2%)	133 (66.2%)	183 (73.5%)	66 (61.1%)	69 (69.7%)	73 (43.7%)
	IGLC	43 (29.7%)	187 (45.6%)	63 (50.8%)	68 (33.8%)	66 (26.5%)	42 (38.9%)	30 (30.3%)	94 (56.3%)
Compartment	GC B cell	85 (58.6%)	248 (60.5%)	95 (76.6%)	134 (66.7%)	151 (60.6%)	68 (63%)	59 (59.6%)	82 (49.1%)
	LNPC	60 (41.4%)	162 (39.5%)	29 (23.4%)	67 (33.3%)	98 (39.4%)	40 (37%)	40 (40.4%)	85 (50.9%)
Time point	Week 4			45 (36.3%)					
	Week 5		180 (43.9%)						
	Week 7	84 (57.9%)	230 (56.1%)	79 (63.7%)	142 (70.6%)	115 (46.2%)	67 (62%)	34 (34.3%)	108 (64.7%)
	Week 15	61 (42.1%)			59 (29.4%)	134 (53.8%)	41 (38%)	65 (65.7%)	59 (35.3%)

Extended Data Table 5 | Processing of BCR and 5′ gene expression data from scRNA-seq and BCR reads from bulk-seq

Participant	Timepoint	Compartment	Replicate	BCR		5′ gene expression			
				Pre-QC number of cells	Post-QC number of cells	Pre-QC number of cells	Post-QC number of cells	Median number of UMIs per cell	Median number of genes per cell
01a	Week 4	PB	1	6211	5562	8106	6247	14035	2471
	Week 7	FNA	1	2691	2484	7656	6967	5093	2002
			2	2946	2714	8837	8089	4977	1901
	Week 15	FNA	1	2515	2320	7040	6583	4923	1958
			2	2529	2347	7504	7078	4884.5	1952
	Week 29	FNA	1	1447	1362	6727	6451	5254	1875
			2	1470	1387	6624	6310	5358.5	1906
02a	Week 4	PB	1	3981	3016	2847	1954	8579.5	1964.5
	Week 5	FNA_1	1	3864	3402	6394	5906	5900.5	2220.5
			2	4069	3582	6017	5603	6737	2397
		FNA_2	1	5555	4919	8243	7519	4648	1862
			2	5455	4761	6411	5790	5491	2045.5
	Week 7	FNA_1	1	7573	5977	11562	10605	5093	1972
			2	7686	6047	10021	9247	5652	2094
		FNA_2	1	5644	4376	9612	8334	4816	1851
			2	5807	4582	8098	6872	5340.5	2023.5
04	Week 4	PB	1	5347	4321	8571	4488	12850.5	2545.5
		FNA	1	2608	2432	5069	4611	5597	2106
			2	2597	2436	6203	5702	5796	2166
	Week 7	FNA	1	3097	2752	7312	6817	5315	1952
			2	2674	2390	6681	6240	5406	1975
07	Week 4	PB	1	4123	3610	6284	3744	14062.5	2606
	Week 7	FNA	1	3527	3081	6977	6330	5545	2105
			2	3757	3434	7489	6950	5433	2071
	Week 15	FNA	1	4483	3992	8787	8380	5269	2038
			2	3860	3465	7580	6951	5469	2142
	Week 29	FNA	1	3541	3117	10301	8774	4944	1802
			2	3382	3003	9628	8174	4896.5	1800.5
10	Week 4	PB	1	3228	2535	2336	1569	13015	2374
	Week 7	FNA_1	1	5148	4596	7843	7322	5509	2052
			2	5420	4697	8767	8169	5258	1987
		FNA_2	1	4862	4473	8375	7610	5839	2189
			2	4724	4322	8193	7445	5837	2223
	Week 15	FNA_1	1	6634	5330	9797	8949	5392	2045
			2	6780	5550	9022	8058	5476	2098
		FNA_2	1	4842	4516	8051	7196	6123	2261
			2	5163	4737	8516	7557	6123	2261
	Week 29	FNA_1	1	5293	4633	8255	7850	4840	1767
		FNA_2	1	5962	5330	8089	7432	5226.5	1925.5
			2	5314	4822	7224	6724	5226.5	1948.5
13	Week 4	PB	1	2241	1897	3072	2260	12389	2435.5
	Week 7	FNA	1	4185	3643	6751	6259	5605	2100
			2	4263	3677	7375	7008	5371.5	2043.5
	Week 15	FNA	1	3098	2716	8730	6931	5612	2186
20	Week 4	PB	1	2962	2175	2011	1345	13238	2347
	Week 7	FNA	1	2591	2317	7791	7440	5460.5	1852
			2	2495	2278	7380	7017	5516	1906
	Week 15	FNA	1	4526	4021	9490	8693	5736	2022
			2	4228	3718	9386	8613	5777	2043
	Week 29	FNA	1	2552	2348	7451	6859	5979	2016
			2	2470	2307	7611	7018	5787.5	1991
22	Week 4	PB	1	6517	5313	9504	6841	11719	2254
	Week 7	FNA	1	4754	3899	11945	11272	5085.5	1921
			2	4826	4005	10934	10413	4612	1835
	Week 15	FNA	1	2501	2206	8378	6926	4725	1795
			2	2279	1932	7855	6216	5301	2008

Participant	Time point	Compartment	Cell Count	Sequence count			
				Input Reads	Preprocessed Reads	Post-QC Productive Heavy Chains	Unique Heavy Chain VDJs
01a	Week 7	GC B cell	16920	1234529	4950	4018	2510
		LNPC	3307	1350772	11799	9851	3592
	Week 15	GC B cell	10440	1031639	1274	897	612
		LNPC	2139	1659983	6337	5348	2122
	Week 29	BMPC_1	8000000	955463	59652	53194	37471
		BMPC_2	8000000	1269266	83277	74407	27918
02a	Week 5	GC B cell_1	56741	1088685	41917	35093	20712
		LNPC_1	4047	952080	40227	30713	7146
		GC B cell_2	18898	1234333	18388	15015	7951
		LNPC_2	3936	1110676	40000	30904	7090
	Week 7	GC B cell_1	58665	1161523	45691	37631	20461
		LNPC_1	7679	1077630	55443	42626	9864
		GC B cell_2	6507	1103370	1450	1003	742
		LNPC_2	1051	963386	16419	12170	3338
	Week 29	MBC	95493	1390109	84345	64093	42243
		BMPC	200000	1132888	63177	55882	35438
04	Week 4	GC B cell	27014	1191023	3115	2207	1779
		LNPC	2312	1012560	9807	8276	2955
	Week 7	GC B cell	37948	1354770	16773	13798	8757
		LNPC	2495	917541	11273	9403	3361
	Week 29	MBC	67760	1632908	68593	57422	38597
		BMPC	200000	1526703	11664	10357	9241
07	Week 7	GC B cell	67030	1330762	37376	32737	18798
		LNPC	7131	1360131	29602	21560	6666
	Week 15	GC B cell	36501	1160090	6994	5377	3762
		LNPC	2108	1798209	19416	14404	3801
	Week 29	MBC	133775	1657286	100577	82382	57530
		BMPC	100000	1212296	19112	17647	14378
10	Week 7	GC B cell_1	22459	1481878	25145	21119	10891
		LNPC_1	8919	1453599	71499	53870	12474
		GC B cell_2	15111	1289713	18400	15237	8337
		LNPC_2	4212	1293442	48343	36409	8473
	Week 15	GC B cell_1	85250	1185039	65048	54102	26862
		LNPC_1	12701	1368903	68759	52349	14086
		GC B cell_2	18438	1414562	13853	11288	7370
		LNPC_2	3056	1103274	34068	25447	6379
	Week 29	MBC	75871	1434312	60385	52787	35516
	Week 40	BMPC	100000	7661613	55918	48290	26952
13	Week 7	GC B cell	26603	1448419	13749	11614	6413
		LNPC	6798	1528468	35222	25799	6997
	Week 15	GC B cell	38318	1424118	13795	11862	7757
		LNPC	3801	1263065	26314	19482	5107
	Week 29	MBC	123272	1289409	98490	83720	57241
		BMPC	100000	3206444	21006	19123	14894
20	Week 7	GC B cell	16773	1249336	10208	8084	4593
		LNPC	3967	1248043	38803	28898	7185
	Week 15	GC B cell	63961	999961	24525	20745	12230
		LNPC	6877	1324665	3196	1998	1323
	Week 29	MBC	103001	1280425	89864	73338	47584
		BMPC	200000	1307413	602	491	380
	Week 40	BMPC	100000	3962648	157106	133610	51104
22	Week 7	GC B cell	17676	1344080	15060	12556	5925
		LNPC	4823	1068783	26503	20779	4907
	Week 15	GC B cell	48264	891196	22310	17775	8612
		LNPC	10443	1290422	28342	19770	6665
	Week 29	MBC	123721	1252951	73542	62160	43396
		BMPC	200000	2185196	115201	105375	64106

Extended Data Table 6 | Description of recombinant mAbs and Fabs derived from clonally related PBs and BMPCs

Participant	Clone ID	Compartment	mAb/Fab ID	Heavy chain ID	Light chain ID	Area under curve	K_D (nM)	Half-maximal neutralization (ng)
01a	01a@G1713_11	PB	01a.E.0002	01a.PH.0002	01a.PL.0001	3.19	12.50	
	01a@G1713_11	BMPC	01a.E.0001	01a.PH.0001	01a.PL.0001	2.98	21.00	
	01a@G1772_3	PB	01a.E.0005	01a.PH.0005	01a.PL.0002	28.74	4.39	283.50
	01a@G1772_3	BMPC	01a.E.0004	01a.PH.0004	01a.PL.0002	29.64	4.14	215.30
	01a@G2588_5	PB	01a.E.0010	01a.PH.0009	01a.PL.0003	2.49	102.00	
	01a@G2588_5	BMPC	01a.E.0007	01a.PH.0006	01a.PL.0003	1.90	167.00	
	01a@G2588_5	BMPC	01a.E.0008	01a.PH.0007	01a.PL.0003	2.21	135.00	
	01a@G2588_5	BMPC	01a.E.0009	01a.PH.0008	01a.PL.0003	1.54	113.00	
	01a@G2679_7	PB	01a.E.0013	01a.PH.0012	01a.PL.0004	22.25	259.00	
	01a@G2679_7	BMPC	01a.E.0012	01a.PH.0011	01a.PL.0004	15.25	97.90	
	01a@G2680_1	PB	01a.E.0023	01a.PH.0019	01a.PL.0007	25.46	64.50	
	01a@G2680_1	BMPC	01a.E.0025	01a.PH.0021	01a.PL.0007	25.85	54.70	
	01a@G2680_1	PB	01a.E.0024	01a.PH.0020	01a.PL.0008	26.59	30.60	
	01a@G2680_1	BMPC	01a.E.0026	01a.PH.0021	01a.PL.0008	27.00	12.40	
	01a@G540_1	PB	01a.E.0016	01a.PH.0014	01a.PL.0005	4.74	1590.00	2033.00
	01a@G540_1	BMPC	01a.E.0015	01a.PH.0013	01a.PL.0005	25.54	15.00	114.30
	01a@G898_1	PB	01a.E.0021	01a.PH.0018	01a.PL.0006	24.60	108.00	3162.00
	01a@G898_1	BMPC	01a.E.0020	01a.PH.0017	01a.PL.0006	27.79	10.60	448.70
02a	02a@G3158_4	PB	02a.E.0006	02a.PH.0005	02a.PL.0001	6.12	28.80	
	02a@G3158_4	BMPC	02a.E.0005	02a.PH.0004	02a.PL.0001	4.42	20.20	
04	04@G1379_17	PB	04.E.0004	04.PH.0004	04.PL.0002	20.76		
	04@G1379_17	BMPC	04.E.0005	04.PH.0005	04.PL.0002	22.40		
	04@G2801_1	PB	04.E.0002	04.PH.0002	04.PL.0001	3.94	61.10	
	04@G2801_1	BMPC	04.E.0001	04.PH.0001	04.PL.0001	6.43	30.40	
07	07@G1832_1	PB	07.E.0003	07.PH.0003	07.PL.0001	25.60	9.75	385.40
	07@G1832_1	BMPC	07.E.0001	07.PH.0001	07.PL.0001	26.66	3.82	184.10
	07@G1832_1	BMPC	07.E.0002	07.PH.0002	07.PL.0001	27.30	2.64	177.10
	07@G2586_1	PB	07.E.0006	07.PH.0006	07.PL.0002	27.61		
	07@G2586_1	BMPC	07.E.0005	07.PH.0005	07.PL.0002	0.00		
10	10@G1624_2	PB	10.E.0002	10.PH.0002	10.PL.0001	4.47	5.25	
	10@G1624_2	BMPC	10.E.0001	10.PH.0001	10.PL.0001	5.41	2.60	
	10@G1733_4	PB	10.E.0066	10.PH.0014	10.PL.0007	1.39	391.00	
	10@G1733_4	BMPC	10.E.0064	10.PH.0012	10.PL.0007	10.30	2.20	
	10@G1733_4	BMPC	10.E.0065	10.PH.0013	10.PL.0007	11.90	2.22	
	10@G643_4	PB	10.E.0088	10.PH.0020	10.PL.0010	22.21	58.00	108.90
	10@G643_4	BMPC	10.E.0086	10.PH.0019	10.PL.0010	25.61	9.63	30.09
20	20@G397_1	PB	20.E.0008	20.PH.0007	20.PL.0001	23.26	149.00	
	20@G397_1	BMPC	20.E.0007	20.PH.0006	20.PL.0001	24.40	277.00	
	20@G838_1	PB	20.E.0015	20.PH.0014	20.PL.0002	25.26	131.00	3851.00
	20@G838_1	BMPC	20.E.0014	20.PH.0013	20.PL.0002	24.82	53.80	3356.00
22	22@G1627_3	PB	22.E.0002	22.PH.0002	22.PL.0001	18.31	87.70	45.63
	22@G1627_3	BMPC	22.E.0001	22.PH.0001	22.PL.0001	22.39	13.70	16.97
	22@G2918_7	PB	22.E.0005	22.PH.0004	22.PL.0002	0.37		
	22@G2918_7	BMPC	22.E.0004	22.PH.0003	22.PL.0002	1.51		
	22@G3017_1	PB	22.E.0008	22.PH.0007	22.PL.0003	7.17	26.30	
	22@G3017_1	BMPC	22.E.0007	22.PH.0006	22.PL.0003	7.54	27.30	
	22@G340_1	PB	22.E.0012	22.PH.0011	22.PL.0004	28.20	20.20	
	22@G340_1	BMPC	22.E.0010	22.PH.0009	22.PL.0004	27.81	13.30	
	22@G340_1	BMPC	22.E.0011	22.PH.0010	22.PL.0004	27.54	13.60	

nature portfolio

Corresponding author(s): Ali H. Ellebedy

Last updated by author(s): Jan 31, 2022

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	Flow cytometry data were acquired using SpectroFlo software v.2.2.
Data analysis	Flow cytometry data were analyzed using FlowJo v.10 and Prism v.9. ELISA and ELISpot were analyzed using Prism v.9. Sequencing data were analyzed using pRESTO v.0.6.2, Cell Ranger v.6.0.1, GRCh38 human reference, IgBLAST v.1.17.1, IMGT/GENE-DB release 202113-2, Change-O v.1.0.2, TIgGER v.1.0.0, Alakazam v.1.1.0, fastcluster v.1.2.3, R v.4.1.0, ggplot2 v.3.3.5, Prism v.9, circlize v.0.4.13, SHazaM v.1.0.2, IgPhyML v.1.1.3, ggtree v.3.0.4, SCANPY v.1.7.2, and Python v.3.8.8. GRNT data were analyzed using IN Cell Analyzer 1000 Workstation Software v3.7. BLI data were analyzed using BIAevaluation v4.1.

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 and transcriptomics count matrix are deposited at Sequence Read Archive and Gene Expression Omnibus respectively under BioProject PRJNA777934. Processed transcriptomics and BCR data are deposited at https://doi.org/10.5281/zenodo.5895181 on Zenodo. Previously reported bulk-sequenced

BCR data used in this study were deposited under PRJNA731610 and PRJNA741267 on SRA, and at https://doi.org/10.5281/zenodo.5042252 and https://doi.org/10.5281/zenodo.5040099 on Zenodo.

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	Total 43 healthy participants were enrolled based on recruitment, of whom 42 provided peripheral blood, 15 provided axillary lymph node samples and 11 provided bone marrow samples. Thirteen out of 43 healthy participants have a history of prior SARS-CoV-2 infection. Non-vaccinated 48 convalescent patients were enrolled based on recruitment. Sample size was not determined by statistical methods, but gave sufficient statistics of the effect sizes of interest.
Data exclusions	No data were excluded
Replication	Human samples were collected from 43 participants. ELISA and GFP-reduction neutralization test were performed once with two technical replicates. Affinity analysis via biolayer interferometry was performed at least two technical replicates according to the fitting-curve. ELISpot and flow cytometry experiments were performed once for each sample at each time point due to insufficient specimens. All atttempts at replication were successful.
Randomization	Different experimental groups were not used.
Blinding	This is not relevant, as this is an observational study.

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. Donkey anti-human IgG (H+L) (Jackson ImmunoResearch, 709-005-149)
2. HRP-conjugated goat anti-human IgG (H+L) (Jackson ImmunoResearch, 109-035-088)
3. HRP-conjugated goat anti-Human IgG Fcγ fragment (Jackson ImmunoResearch, 109-035-190)
4. HRP-conjugated goat anti-human serum IgA α chain (Jackson ImmunoResearch, 109-035-011)
5. HRP-conjugated goat anti-human IgM (Caltag, H15007)
6. BCL6-PE (K112-91, BD Pharmingen, 561522)
7. BLIMP1-A700 (646702, R&D, IC36081N)
8. CD38-BB700 (HIT2, BD Horizon, 566445)
9. CD45-A532 (HI30, Thermo, 58-0459-42)
10. IgA-FITC (M24A, Millipore, CBL114F)
11. IgG-BV480 (goat polyclonal, Jackson ImmunoResearch, 109-685-098)
12. PD-1-BB515 (EH12.1, BD Horizon, 564494)
13. CD3-FITC (HIT3a, BioLegend, 300306)
14. CD3-APC-Fire810 (SK7, BioLegend, 344858)
15. CD4-Spark685 (SK3, BioLegend, 344658)
16. CD4-Alexa-Fluor-700 (SK3, BioLegend, 344622)
17. CD8-BV570 (RPA-T8, BioLegend, 301038)

18. CD14-PerCP (HCD14, BioLegend, 325632)
19. CD19-BV750 (HIB19, BioLegend, 302262)
20. CD19-PE (HIB19, BioLegend, 302254)
21. CD19-APC (HIB19, BioLegend, 302212)
22. CD20-Pacific Blue (2H7, BioLegend, 302320)
23. CD27-BV510 (O323, BioLegend, 302836)
24. CD38-PE-Cy7 (HIT2, BioLegend, 303516)
25. CD71-PE (CY1G4, BioLegend, 334106)
26. CD71-PE-Cy7 (CY1G4, BioLegend, 334112)
27. CD71-APC (CY1G4, BioLegend, 334108)
28. CXCR5-PE-Dazzle 594 (J252D4, BioLegend, 356928)
29. FOXP3-BV421 (206D, BioLegend, 320124)
30. HLA-DR-BV650 (L243, BioLegend, 307650)
31. IgD-PE-Cy5 (IA6-2, BioLegend, 348250)
32. IgD-PerCP-Cy5.5 (IA6-2, BioLegend, 348208)
33. IgM-BV605 (MHM-88, BioLegend, 314524)
34. Ki-67-BV711 (Ki-67, BioLegend, 350516)
35. T-bet-BV785 (4B10, BioLegend, 644835)

Validation

All commercial antibodies were validated by their manufacturers as detailed in their product information and titrated in the lab for the indicated assay by serial dilution. We validated PB-, GC B cell-, LNPC, and BMPC-derived mAbs generated in our lab in preliminary ELISAs to SARS-CoV-2 spike, bovine serum albumin, and anti-Ig. The threshold of positivity for mAbs was set as two times the optical density of background binding to BSA at 10 ug/ml of each mAb.
1. https://www.jacksonimmuno.com/catalog/products/709-005-149
2. https://www.jacksonimmuno.com/catalog/products/109-035-088
3. https://www.jacksonimmuno.com/catalog/products/109-035-190
4. https://www.jacksonimmuno.com/catalog/products/109-035-011
5. https://www.thermofisher.com/antibody/product/Goat-anti-Human-IgM-Secondary-Antibody-Polyclonal/SA5-10293?imageId=710449
6. https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/pe-mouse-anti-bcl-6.561522
7. https://www.rndsystems.com/products/human-blimp1-prdm1-alexa-fluor-700-conjugated-antibody-646702_ic36081n
8. https://www.bdbiosciences.com/en-pt/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/bb700-mouse-anti-human-cd38.566446
9. https://www.thermofisher.com/antibody/product/CD45-Antibody-clone-HI30-Monoclonal/58-0459-42
10. https://www.emdmillipore.com/US/en/product/Mouse-Anti-Human-IgA-Antibody-clone-M24A-heavy-chain-FITC-conjugated,MM_NF-CBL114F?ReferrerURL=https%3A%2F%2Fwww.google.co.kr%2F
11. https://www.jacksonimmuno.com/catalog/products/109-685-098
12. https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/bb515-mouse-anti-human-cd279-pd-1.564494
13. https://www.biolegend.com/nl-be/products/fitc-anti-human-cd3-antibody-751?Clone=HIT3a
14. https://www.biolegend.com/en-us/search-results/apc-fire-810-anti-human-cd3-antibody-19515?GroupID=BLG7576&Clone=SK7
15. https://www.biolegend.com/en-us/products/spark-nir-685-anti-human-cd4-antibody-18516?Clone=SK3
16. https://www.biolegend.com/en-us/search-results/alexa-fluor-700-anti-human-cd4-antibody-9354?Clone=SK3
17. https://www.biolegend.com/fr-ch/products/brilliant-violet-570-anti-human-cd8a-antibody-7371?Clone=RPA-T8
18. https://www.biolegend.com/en-us/products/percp-anti-human-cd14-antibody-9564?Clone=HCD14
19. https://www.biolegend.com/en-us/search-results/brilliant-violet-750-anti-human-cd19-antibody-15900?GroupID=BLG10095&Clone=HIB19
20. https://www.biolegend.com/en-us/products/pe-anti-human-cd19-antibody-719?Clone=HIB19
21. https://www.biolegend.com/en-us/products/apc-anti-human-cd19-antibody-715?Clone=HIB19
22. https://www.biolegend.com/en-us/products/pacific-blue-anti-human-cd20-antibody-3330?Clone=2H7
23. https://www.biolegend.com/en-us/search-results/brilliant-violet-510-anti-human-cd27-antibody-8005?Clone=O323
24. https://www.biolegend.com/en-us/products/pe-cyanine7-anti-human-cd38-antibody-5418?Clone=HIT2
25. https://www.biolegend.com/en-us/products/pe-anti-human-cd71-antibody-4908?Clone=CY1G4
26. https://www.biolegend.com/en-us/products/pe-cyanine7-anti-human-cd71-antibody-9328?Clone=CY1G4
27. https://www.biolegend.com/en-us/products/apc-anti-human-cd71-antibody-7517?Clone=CY1G4
28. https://www.biolegend.com/en-us/products/pe-dazzle-594-anti-human-cd185-cxcr5-antibody-9860?Clone=J252D4
29. https://www.biolegend.com/en-us/products/brilliant-violet-421-anti-human-foxp3-antibody-12045?Clone=206D
30. https://www.biolegend.com/en-us/search-results/brilliant-violet-650-anti-human-hla-dr-antibody-8875?Clone=L243
31. https://www.biolegend.com/en-us/products/pe-cyanine5-anti-human-igd-antibody-19969?Clone=IA6-2
32. https://www.biolegend.com/en-us/products/percp-cyanine5-5-anti-human-igd-antibody-6811?Clone=IA6-2
33. https://www.biolegend.com/en-us/products/brilliant-violet-605-anti-human-igm-antibody-8746?Clone=MHM-88
34. https://www.biolegend.com/en-us/search-results/brilliant-violet-711-anti-human-ki-67-antibody-7946?Clone=Ki-67
35. https://www.biolegend.com/fr-ch/products/brilliant-violet-785-anti-t-bet-antibody-15077?Clone=4B10
We validated mAbs generated in our lab in preliminary ELISAs to SARS-CoV-2 spike, bovine serum albumin, and anti-Ig.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Expi293F, Vero
Authentication	The cell line was not authenticated.

Mycoplasma contamination	Cell lines were not tested for mycoplasma contamination. Growth rates were consistent with manufacturer's published data.
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	Study participants demographics are detailed in Extended Data Table 1.
Recruitment	Study participants were recruited from the St. Louis metropolitan area by the Washington University Clinical Trials Unit. Recruitment was open to all eligible adults who were supposed to receive the two-dose series of Pfizer-BioNTech SARS-CoV-2 mRNA vaccine (BNT162b2). All participants privided written informed consent before participation in the study. Participants were asked to provide details of SARS-CoV-2 infection history, and side effects and comorbidities after vaccination. Other than the criteria listed herein, no other parameters were used to select participants. Potential self-selection and recruiting biases are unlikely to affect the parameters we measured.
Ethics oversight	The study was approved by the Washington University IRB.

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

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	Peripheral blood and bone marrow mononuclear cells were isolated from EDTA anticoagulated blood and bone marrow aspirates, respectively using density gradient centrifugation, and remaining RBCs were lysed with ammonium chloride lysis buffer. Bone marrow plasma cells were magnetically enriched from bone marrow mononuclear cells and immediately used for ELISpot or cryopreserved in 10% dimethylsufoxide in FBS for flow cytometric analysis. PBMCs were immediately used or cryopreserved in 10% DMSO in FBS. Fine needle aspirates of axillary LNs were flushed from needles with 3 mL of RPMI supplemented with 10% FBS and 100 U/mL penicillin/streptomycin, followed by three 1 mL rinses. Red blood cells were lysed with ammonium chloride buffer, washed twice with PBS supplemented with 2% FBS, 2mM EDTA and immediately used or cryopreserved in 10% DMSO in FBS.
Instrument	Cytek Aurora
Software	Flow cytometry data was acquired using Cytek SpectroFlo and analyzed using FlowJo (Treestar) v10.
Cell population abundance	Bulk sorts directly into lysis buffer were not amenable to post-sort purity analysis.
Gating strategy	Details of gating strategies are shown in Extended Data Figures. Briefly, forward and side scatter parameters (FSC-A/H/W, SSC-A,H,W), and Zombie dyes (BioLegend) were used to select for live singlet lymphocytes. For analyzing SARS-CoV-2 S binding cells, GC B cells were gated on CD3-/CD19+/IgD-/CD38int/BCL6+ live singlet lymphocytes; LNPCs were gated on CD3-/CD19+/IgD-/CD20low/CD38+/BLIMP+/CD71+ live singlet lymphocytes; and MBCs from PBMCs were gated on CD3-/CD19+/IgD-/CD20+/CD38- live singlet lymphocytes. Blood and tonsillectomy samples collected before the COVID-19 pandemic were stained as a negative control for S-binding cells. For sorting, plasmablasts from PBMCs were gated on CD3-/CD19+/IgD-/CD20-/CD38+/CD71+ live singlet lymphocytes; GC B cells were gated on CD4-/CD19+/IgD-/CD20high/CD38int live singlet lymphocytes; and LNPCs were gated on CD4-/CD19+/IgD-/CD20low/CD38+ live singlet lyphocytes.

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