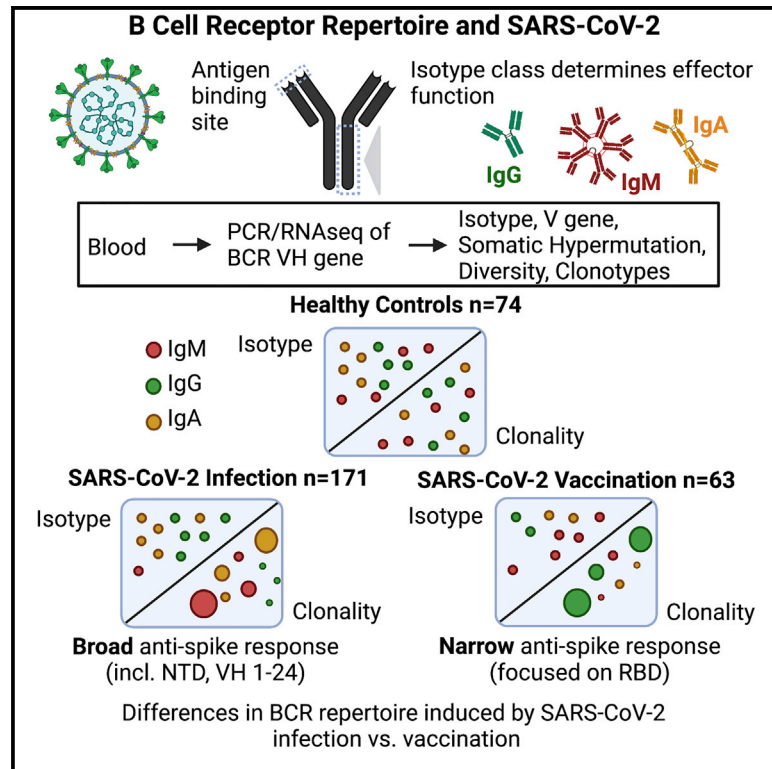


# B cell receptor repertoire kinetics after SARS-CoV-2 infection and vaccination

## Graphical abstract



## Authors

Prasanti Kotagiri, Federica Mescia, William M. Rae, ..., Rachael J.M. Bashford-Rogers, Paul A. Lyons, Kenneth G.C. Smith

## Correspondence

pk488@cam.ac.uk (P.K.),  
pal34@cam.ac.uk (P.A.L.),  
kgcs2@cam.ac.uk (K.G.C.S.)

## In brief

Kotagiri et al. find that SARS-CoV-2 infection versus vaccination induces distinct changes in the B cell receptor repertoire, including prominent clonal expansion in IgA and IgM after infection, but IgG after vaccination. A broad anti-spike response to infection contrasts with a narrower RBD-focused one after vaccination, potentially informing vaccination strategies.

## Highlights

- BCR repertoire changes after SARS-CoV-2 infection versus vaccination differ
- Clonal expansion is prominent in IgA and M post infection, IgG post vaccination
- COVID-19 is associated with a broad anti-spike response, including NTD and VH1-24
- SARS-CoV-2 vaccination may drive narrower spike specificity targeting RBD



## Resource

# B cell receptor repertoire kinetics after SARS-CoV-2 infection and vaccination

Prasanti Kotagiri,<sup>1,2,\*</sup> Federica Mescia,<sup>1,2</sup> William M. Rae,<sup>1,2</sup> Laura Bergamaschi,<sup>1,2</sup> Zewen K. Tuong,<sup>2,3</sup> Lorinda Turner,<sup>1,2</sup> Kelvin Hunter,<sup>1,2</sup> Pehuén P. Gerber,<sup>1,2</sup> Myra Hosmillo,<sup>4</sup> Cambridge Institute of Therapeutic Immunology and Infectious Disease-National Institute of Health Research (CITIID-NIHR) COVID BioResource Collaboration, Christoph Hess,<sup>1,2,5,6</sup> Menna R. Clatworthy,<sup>1,2,3</sup> Ian G. Goodfellow,<sup>4</sup> Nicholas J. Matheson,<sup>1,2,7</sup> Eoin F. McKinney,<sup>1,2</sup> Mark R. Wills,<sup>1,2</sup> Ravindra K. Gupta,<sup>1,2</sup> John R. Bradley,<sup>2,8</sup> Rachael J.M. Bashford-Rogers,<sup>9</sup> Paul A. Lyons,<sup>1,2,\*</sup> and Kenneth G.C. Smith<sup>1,2,10,\*</sup>

<sup>1</sup>Cambridge Institute of Therapeutic Immunology and Infectious Disease, Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge CB2 0AW, UK

<sup>2</sup>Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge CB2 0QQ, UK

<sup>3</sup>Cellular Genetics, Wellcome Sanger Institute, Hinxton, Cambridge CB10 1RQ, UK

<sup>4</sup>Division of Virology, Department of Pathology, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK

<sup>5</sup>Department of Biomedicine, University and University Hospital Basel, Basel 4031, Switzerland

<sup>6</sup>Botnar Research Centre for Child Health (BRCC) University Basel and ETH Zurich, Basel 4059, Switzerland

<sup>7</sup>NHS Blood and Transplant, Cambridge CB2 1PT, UK

<sup>8</sup>NIHR BioResource, Cambridge University Hospitals NHS Foundation, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK

<sup>9</sup>Nuffield Department of Medicine, Wellcome Centre for Human Genetics, Oxford OX3 7BN, UK

<sup>10</sup>Lead contact

\*Correspondence: [pk488@cam.ac.uk](mailto:pk488@cam.ac.uk) (P.K.), [pal34@cam.ac.uk](mailto:pal34@cam.ac.uk) (P.A.L.), [kgcs2@cam.ac.uk](mailto:kgcs2@cam.ac.uk) (K.G.C.S.)

<https://doi.org/10.1016/j.celrep.2022.110393>

## SUMMARY

B cells are important in immunity to both severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) infection and vaccination, but B cell receptor (BCR) repertoire development in these contexts has not been compared. We analyze serial samples from 171 SARS-CoV-2-infected individuals and 63 vaccine recipients and find the global BCR repertoire differs between them. Following infection, immunoglobulin (Ig)G1/3 and IgA1 BCRs increase, somatic hypermutation (SHM) decreases, and, in severe disease, IgM and IgA clones are expanded. In contrast, after vaccination, the proportion of IgD/M BCRs increase, SHM is unchanged, and expansion of IgG clones is prominent. VH1-24, which targets the N-terminal domain (NTD) and contributes to neutralization, is expanded post infection except in the most severe disease. Infection generates a broad distribution of SARS-CoV-2-specific clones predicted to target the spike protein, while a more focused response after vaccination mainly targets the spike's receptor-binding domain. Thus, the nature of SARS-CoV-2 exposure differentially affects BCR repertoire development, potentially informing vaccine strategies.

## INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) resulting in coronavirus disease 2019 (COVID-19) has caused over 5 million deaths as of November 2021 (<https://covid19.who.int/>). It primarily infects respiratory epithelial cells, and results in a range of clinical manifestations from asymptomatic disease to multi-organ failure. B cells play a vital role in anti-viral defense (Quast and Tarlinton, 2021). B cell depletion can result in persisting viremia (Buckland et al., 2020; Kemp et al., 2021), SARS-CoV-2 neutralizing monoclonal antibodies and convalescent plasma may have a therapeutic role (Joyner et al., 2021; Libster et al., 2021), and neutralizing antibodies may prevent re-infection and transmission (Kim et al., 2021b).

These observations make a strong case for a central role for B cells in the defense against SARS-CoV-2. There is strong

evidence that neutralizing SARS-CoV-2-specific antibodies can protect against disease onset and progression (Cox and Brokstad, 2020; Garcia-Beltran et al., 2021; Stephens and McElrath, 2020) and potentially also through non-SARS-CoV-2-specific "natural" antibodies, or antibodies generated in response to other coronaviruses that may also cross-react with SARS-CoV-2 (Hernandez and Holodick, 2017; Song et al., 2021; Yang et al., 2021). It is also likely that B cells play a role through other functions, including antigen presentation to T cells, cytokine production, and other regulatory mechanisms.

Severe COVID-19 is typified by major perturbations in circulating immune cells (Arunachalam et al., 2020; Bergamaschi et al., 2021; Hadjadj et al., 2020; Laing et al., 2020; Mann et al., 2020; Mathew et al., 2020; Rydzynski Moderbacher et al., 2020; Schultheiß et al., 2020; Stephenson et al., 2021; Wen et al., 2020). Together with other groups, we have shown that



COVID-19 has a profound impact on B cell subsets. Increased numbers of recently generated circulating plasmablasts are seen early in disease irrespective of severity, and indeed is one of the few cellular abnormalities observed in asymptomatic SARS-CoV-2 infection (Bergamaschi et al., 2021). Absolute numbers of almost all other B cell subpopulations are reduced, including naive B cells, both switched and unswitched memory B cells, transitional B cells, and, more recently, marginal zone-like B cells (Bergamaschi et al., 2021). All of these B cell subsets are maximally reduced soon after symptom onset, with most gradually resolving thereafter (with the exception of transitional B cells, which continue to decline over the first 2 months after infection) (Bergamaschi et al., 2021). Early histology reports also demonstrated reduced germinal centers in secondary lymphoid organs in COVID-19, and, consistent with this, circulating TFH-like cells are markedly reduced (Kaneko et al., 2020). Most initial reports have underestimated the impact of COVID-19 on the B cell immune response, having examined proportions rather than absolute numbers of B cell subsets (Bergamaschi et al., 2021). Changes between these subsets, as well as within them, will be reflected in the B cell receptor (BCR) repertoire.

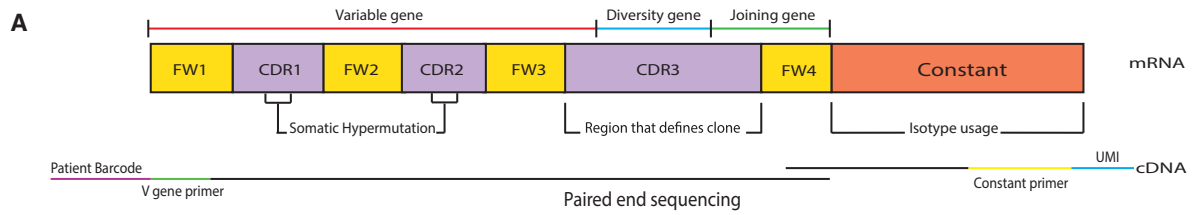
The BCR repertoire refers to the range of individual BCRs that collectively provide the diversity of antigen receptors required by B cells to recognize new antigens, to minimize interaction with autoantigens, and, when certain specificities are expanded, to provide increased protection in the context of B cell memory. BCR diversity is driven by the rearrangement of the immunoglobulin receptor genes during B cell development in the bone marrow. During B cell development, single variable (V), diversity (D), and joining (J) genes are selected from multiple distinct copies and imprecisely joined to create a BCR (Schatz and Swanson, 2011). To prevent self-reactivity, B cells go through both central and peripheral tolerance checkpoints (Theofilopoulos et al., 2017; Wardemann et al., 2003). Further diversification of the BCR repertoire occurs post antigen exposure through somatic hypermutation and subsequent selection of high-affinity clones (Smith et al., 1997; Tonegawa, 1983; Victora and Nussenzweig, 2012). B cells may undergo a process termed immunoglobulin class switching where, through stepwise DNA deletion and recombination of the constant region, downstream isotypes are generated (Stavnezer et al., 2008). During this process, the antigen-binding region remains the same, and so therefore does antigen affinity, but isotype switching confers a range of different effector functions (Xu et al., 2012). High-throughput bulk RNA sequencing of BCR heavy-chain genes allows us to assess isotype use, somatic hypermutation, V gene usage, and clonality.

The study of the repertoire has been illuminating in immune-mediated disease, infection, and vaccination. In previous work, we described increased clonality, IgA proportion, and shared IGHV gene usage in systemic lupus erythematosus and Crohn's disease (Bashford-Rogers et al., 2019). Early reports have similarly revealed substantial changes in the BCR repertoire in severe COVID-19. An increased representation of IgG1 and reduced IgM isotypes is seen, as is the over-representation of some specific heavy-chain genes (such as the VH3 family). A global reduction in somatic hypermutation has also been observed compared with health (Galson et al., 2020; Kreer

et al., 2020; Kuri-Cervantes et al., 2020; Nielsen et al., 2020; Schulte-Schrepping et al., 2020; Seydoux et al., 2020). Analysis of SARS-CoV-2-specific B cells has demonstrated some changes consistent with those seen in the global BCR repertoire, in particular low SHM early in disease with a subsequent increase in the memory population (Gaebler et al., 2021; Robbiani et al., 2020; Sokal et al., 2021; Wang et al., 2021a). Reduced SHM levels in BCR repertoires have been seen early after Ebola (Davis et al., 2019), and Dengue infection (Godoy-Lozano et al., 2016), with the pattern of early low SHM followed by a late increase in SHM-high clones being seen in other infections (for example, vesicular stomatitis virus; Kalinke et al., 1996). This has been attributed to a prominent early extrafollicular response, which is characterized by the rapid production of short-lived unmutated plasmablasts differentiating outside germinal centers. The subsequent increase in total SHM is then likely to be driven by a combination of the death of these unmutated plasmablasts, accompanied by increased production of mutated B cells as the germinal center response expands.

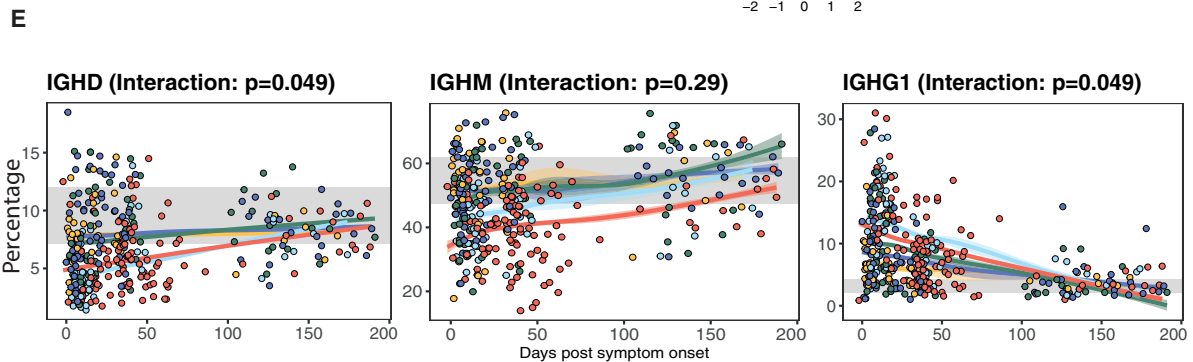
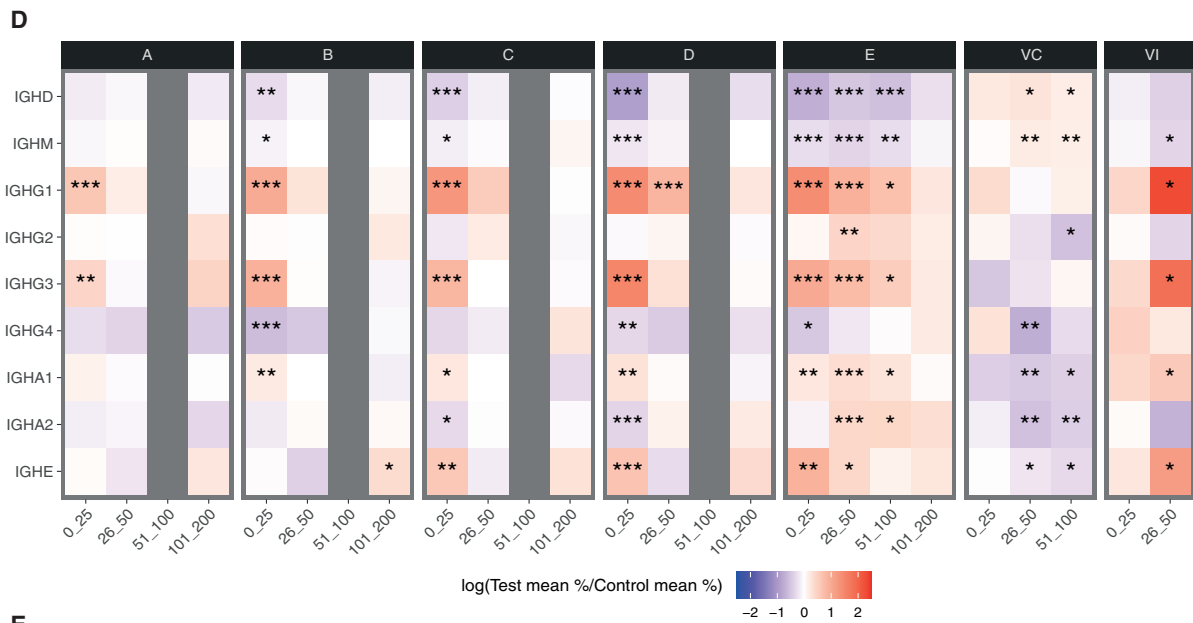
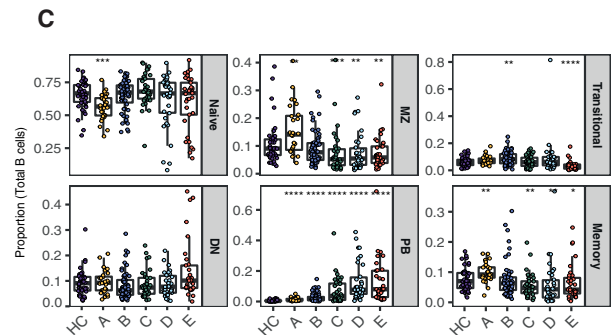
Increasingly, more work is being conducted on the BCR repertoire post SARS-CoV-2 vaccination. Studies show that SARS-CoV-2 mRNA vaccination induces antibodies against NTD, receptor-binding domain (RBD), and S2 (Amanat et al., 2021), with anti-RBD clones showing high use of IGHV3-30 and IGHV3-53, similar to that seen in natural infection (Wang et al., 2021b). However, neutralizing ability post vaccination appears targeted to the RBD domain with removal of RBD-specific antibodies abolishing neutralization of Wuhan-Hu-1 virus (Greaney et al., 2021). Vaccine-elicited antibodies appear more broadly distributed across the RBD compared with natural infection, potentially preventing loss of efficacy when point mutations occur in the virus (Stamatatos et al., 2021). We analyzed the global BCR repertoire following vaccination with mRNA vaccine BNT162b22, revealing an increase in SARS-CoV-2-specific sequences and higher level of SHM in the <80-year-olds compared with >80-year-olds (Collier et al., 2021).

Increasing our understanding of the B cell immune response in the context of COVID-19 is important given its role in defense against SARS-CoV-2 infection, and potentially in the prevention of secondary infection, re-infection, and autoimmunity. While we have some understanding of this early after SARS-CoV-2 infection: little is known about how the BCR repertoire changes over time, varies with disease severity, or compares with that generated by vaccination. Studying the global BCR repertoire allows us to study not only antigen-specific B cells but also "bystander" viral-associated clones that are often mobilized post infection and vaccination (Horns et al., 2020). We have analyzed the BCR repertoire in a large cohort of patients with varying disease severity, sampling at a number of timepoints to 6 months post symptom onset, and comparing these with BCR repertoire changes following vaccination with the BNT162B2 SARS-CoV-2 vaccine (Collier et al., 2021), and the trivalent influenza vaccine (TIV) (specific for influenza A [H3N2], A [H1N1], and B). Differences in the BCR repertoire might be expected between responses to natural infection and vaccination, which could affect the efficacy of both the acute response to viral infection and the quality and longevity of the memory response. For example, natural infection is focused on the mucosa, and has different antigen



**B**

	Days				
		0-25	26-50	51-100	101-200
Healthy Donors	HC	74 (74)			
Screening asymptomatic	A	26 (17)	8 (8)		7 (7)
Screening symptomatic	B	48 (34)	14 (13)		20 (20)
Hospital: room air	C	36 (33)	14 (14)	3 (3)	17 (17)
Hospital: supplemental O <sub>2</sub>	D	37 (30)	10 (10)	2 (2)	9 (9)
Hospital: assisted vent	E	29 (24)	46 (38)	19 (16)	13 (13)
Vaccinated: SARS-CoV-2	VC	11 (11)	21 (21)	31 (31)	
Vaccinated: INFLUENZA	VI	14 (14)	9 (9)		



(legend on next page)

kinetics and persistence compared with vaccination. This could result in, for example, reduced SHM (due to prominent early plasmablast responses) and an increase in IgA versus IgG use in infection. In contrast, differences in antigen delivery in vaccination could lead to a more prominent IgG response, and an earlier and more sustained increase in SHM, potentially providing both qualitative and quantitative difference in the memory response that could affect future protection. In providing such information, comparison of the BCR repertoire between infection and vaccination may explain susceptibility to severe disease and inform vaccination strategies.

## RESULTS

### Patient cohorts and BCR repertoire analysis

SARS-CoV-2 PCR-positive subjects ( $n = 171$ ) were recruited between 31 March and 20 July 2020 and divided into five categories according to peak clinical severity; those studied here were members of the cohorts described by Bergamaschi et al. (2021), and BCR repertoires were generated (Figures 1A and 1B, S1A and S1B):

- Asymptomatic healthcare workers (HCWs) recruited from routine screening.
- HCWs either still working with mild symptoms, or symptomatic and self-isolating.
- Patients who presented to hospital but never required oxygen supplementation.
- Admitted patients whose maximal respiratory support was supplemental oxygen.
- Patients who required assisted ventilation (56) or died without ventilation (3).

Patients were bled weekly while inpatients, and less frequently thereafter. Patient time courses are measured since symptom onset for groups B to E, and from the first positive swab for group A (not having symptoms to trigger presentation, patients in group A were likely sampled, on average, later after infection than B–E). Recipients ( $n = 63$ ) of the BNT162B2 (Pfizer/BioNTech) SARS-CoV-2 vaccine were bled after the initial dose, prior to boosting. Recipients ( $n = 14$ ) of the TIV were bled before vaccination, and then at 7 and 30 days after. Healthy controls were recruited across a range of ages and included the TIV recipients prior to vaccination. To account for the effect of aging on the BCR repertoire, all healthy control comparisons were age matched (Figure S1C). This was achieved by examining the age

distribution of participants after grouping according to both disease severity/vaccination and time. A subset of healthy controls was then selected to mirror this distribution and used in the comparison.

To understand how compositional changes in B cell subsets might influence the global BCR repertoire, we compared B cell proportions in patients within 25 days from symptom onset to healthy controls (Figure 1C). The proportion of plasmablasts was increased in all severity groups and, in addition, marginal zone, transitional, and memory B cells were decreased in group E.

### Altered isotype use after SARS-CoV-2 infection

We assessed the proportion of unique B cell clones of different isotypes, counting each unique VDJ region only once to ensure results were not skewed by the differential mRNA content of B cell subsets (in particular plasmablasts, which have increased immunoglobulin mRNA content) (Table S1). BCR repertoire analysis cannot fully capture the entirety of a patient's repertoire, as only a small proportion of B cells are sampled in a given blood draw, and the analysis is also sensitive to sequencing read depth. We therefore confirmed reproducibility of the BCR repertoire metrics reported using biological replicates derived from peripheral blood mononuclear cells (PBMCs) and whole blood (Figures S1D–S1F and Table S2).

IgG1 and IgG3 proportions were increased across all severity groups (IgG1, 1.87–4.27-fold; IgG3, 1.56–4.57-fold), and were the only isotypes increased in the asymptomatic group A. IgA1 was increased in a similar pattern (1.14–1.34-fold), although changes were less pronounced, and IgE was increased in groups C to E (1.85–2.85-fold) (Figure 1D). IgD and IgM were reduced, particularly in those with severe disease (IgD, 0.40–0.85-fold; IgM, 0.76–0.94-fold). All isotype proportions returned to normal over time, with recovery being delayed in more severe groups. These changes in isotype proportion were illustrated assessing time as a continuous variable using a linear mixed-effects model analysis (Figure 1E).

Correlation of BCR isotype use derived from BCR repertoire sequencing with B cell subset numbers and with serum immunoglobulin titers was performed (Figure S1G). The strongest positive correlations were seen between IgG1, IgG3, and IgA1 and plasmablast numbers, and the strongest negative correlation between IgM levels and plasmablast numbers, suggesting the increased proportion of these switched isotypes was, in large part, driven by an increase in clonally distinct plasmablasts (Figure 1C). Consistent with this was the increased IgG1 and IgG3 proportions seen in group A, in which an early rise in

### Figure 1. Altered isotype use after SARS-CoV-2 infection

- (A) Cartoon of BCR sequencing method. UMI denoted in blue, patient barcode in red, forward variable gene primer in green, and reverse constant primer in yellow. Region used to assess somatic hypermutation, clone identity, and isotype marked.
- (B) Study participant and sample numbers split by severity categories and time after screening (cat. A), symptom onset (cat. B–E), or vaccination (cat. VC and VI). The number of samples are shown, with the number of individuals from whom they were collected shown in brackets.
- (C) Boxplots showing B cell subset proportions according to disease severity at 0–25 days from symptom onset. Naive ( $CD19^+IgD^+CD27^-$ ), marginal zone B cells (MZ) ( $CD19^+IgD^+CD27^+$ ), double-negative B cells (DN) ( $CD19^+IgD^-CD27^-$ ), transitional ( $CD19^+IgD^+CD27^+CD24^+CD38^+$ ), memory ( $CD19^+IgD^-CD27^+CD24^+CD38^+$ ), and plasmablasts (PB) ( $CD19^-CD20^+CD27^+CD24^+CD38^+$ ). Comparison with HC, unadjusted Wilcoxon test p value: \* $p < 0.05$ , \*\* $p < 0.005$ , \*\*\* $p < 0.0005$ .
- (D) Heatmap showing log<sub>2</sub> fold change in mean proportion between SARS-CoV-2 and vaccine cases and HC, within severity categories and across time bins. Wilcoxon test false discovery rate (FDR) adjusted p value: \* $p < 0.05$ , \*\* $p < 0.005$ , \*\*\* $p < 0.0005$ .
- (E) Linear mixed-effects model showing the longitudinal expression of IGHD, IGHM, and IGHD proportions over time, grouped by severity. Gray band indicates the interquartile range of the corresponding isotype in HCs. Nominal p values for the time  $\times$  severity group interaction term are reported.



plasmablasts is the only prominent change in B cell subpopulations seen (Bergamaschi et al., 2021). IgD and IgM correlate strongly with naive B cell number, suggesting that their decline is in part a reflection of reduced naive and transitional B cell numbers in moderate to severe COVID-19. Correlation between IgA1 isotype use and serum IgA was seen, but no such correlation was seen between the IgG isotypes and serum IgG (Figure S1G). The lack of correlation between BCR repertoire and serum IgG titers is likely to be due to differences in both cellular and immunoglobulin kinetics. “Steady-state” serum IgG is made predominantly by long-lived plasma cells in the bone marrow or in chronically inflamed tissue, with acute increases after infection being driven largely by extrafollicular plasmablasts. Neither would necessarily correlate with the circulating memory B cells and plasmablasts contributing to the BCR repertoire assessments. This will be compounded by immunoglobulin kinetics: IgG takes time to build up in serum in immune responses, lagging behind the cellular response, and has a serum half-life of 21 days, so will persist after cellular resolution begins. Together these mean that correlation between serum Ig and BCR repertoire may not occur, particularly in the setting of acute disease.

We quantified the level of switching between classes in different disease subsets, by assessing the frequency of unique VDJ regions that shared two different isotypes, having corrected for read depth by subsampling. This demonstrated increased switching to IgG1 and IgA1 in all severity groups in the first 25 days after symptom onset. Beyond 25 days, increased switching was prominent only to IgA1, and predominantly in those with more severe disease (Figure S1H). This is unlikely to be due to persisting virus, as clearance occurs within the first 25 days as measured by nasal/throat swab but rather could be associated with ongoing inflammation in group E shown by continued elevation in CRP (Bergamaschi et al., 2021).

Isotype changes in response to BNT162B2 SARS-CoV-2 vaccine were very different to those seen in SARS-CoV-2 infection. Increased IGHD (1.29-fold) and M (1.2-fold) isotype proportions were apparent only after 25 days from vaccination with concurrent decreases in IGHG2/4 (0.77-fold, 0.47-fold), IGHA1/2 (0.63-fold, 0.58-fold), and IGHE (0.81-fold) (Figure 1D). Similarly, isotype changes were only seen in response to the TIV beyond 25 days after vaccination but the prominent increases in IgG1 (10-fold), IgG3 (6.26-fold), and IgA1 (1.83-fold) proportions mirrored that of SARS-CoV-2 infection (Figure 1D).

### Somatic hypermutation in the SARS-CoV-2 BCR repertoire

SHM is the mechanism by which the BCR repertoire is diversified during the germinal center reaction, with the subsequent selection of high-affinity mutants resulting in “affinity maturation,” and potentially also in an increased breadth of the memory B cell repertoire (Smith et al., 1997; Tonegawa, 1983; Victora and Nussenzweig, 2012). Reduced SHM has been seen in SARS-CoV-2 infection by others (Galson et al., 2020; Kreer et al., 2020; Kuri-Cervantes et al., 2020; Nielsen et al., 2020; Schultheiß et al., 2020; Seydoux et al., 2020), and this is confirmed in our cohort (Figures 2A and S2A). Reduced SHM is most pronounced in IgG1 (0.72–0.78-fold change), IgG3 (0.68–0.86-fold change), and IgA1 (0.86–0.90-fold change), and to a lesser extent IgA2 (0.92–0.94-fold change) and IgE (0.80–0.94-fold change). This is most prominent early after symptom onset, occurs across all severity groups, and recovers over time. SHM is reduced in the isotypes most increased in the BCR repertoire, suggesting that most expansion occurs outside the germinal center. When the kinetics of SHM reduction are considered in more detail, it can be seen that SHM reaches its nadir between 11 and 20 days after symptom onset in most groups (Figure 2B). Reduced SHM could reflect the relative increase in the proportion of unmutated B cell clones, highest in early time points compared with late ones (Figure 2C), although less pronounced in group E (Figure 2D).

In contrast, a marked increase in SHM in IgD (~1.8-fold), and to a lesser extent IgM (~1.2-fold), is present in those with moderate to severe COVID-19. This may be reflective of a cellular compositional change in IgM- and IgD-positive cells. When compared with HC, patients with severe COVID-19 have a proportional increase in IgM+ plasmablasts and memory B cells, which are expected to have a higher mutational load compared with naive B cells. This was confirmed with levels of SHM in IGHM/D clones having a negative correlation with CD19 naive B cell proportions and a positive correlation with IgM memory and plasmablast proportions (Figures S2B and S2C). There was also a comparative increase in SHM in expanded compared with unexpanded clones (Figure S2D) consistent with generation in the germinal center. SARS-CoV-2 vaccination did not appreciably alter SHM, while influenza vaccination showed an increase at 25–50 days (Figure 2A).

Finally, the acquisition of anti-SARS-CoV-2 spike IgG antibodies was temporally associated with reduced global IgG1

### Figure 2. Somatic hypermutation in the SARS-CoV-2 BCR repertoire

(A) Heatmap showing the log<sub>2</sub> fold change in mean frequency of replacement mutations covering regions CDR1 and CDR2 between SARS-CoV-2 infection and vaccine cases and HC, within severity categories and across time bins post screening (cat. A), symptom onset (cat. B–E), or vaccination (cat. VC and VI). Wilcoxon test FDR adjusted p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005.

(B) Boxplots showing the mean frequency of replacement mutations covering regions CDR1 and CDR2 in IGHG1 split by severity categories and time bins. Circles represent individual donors.

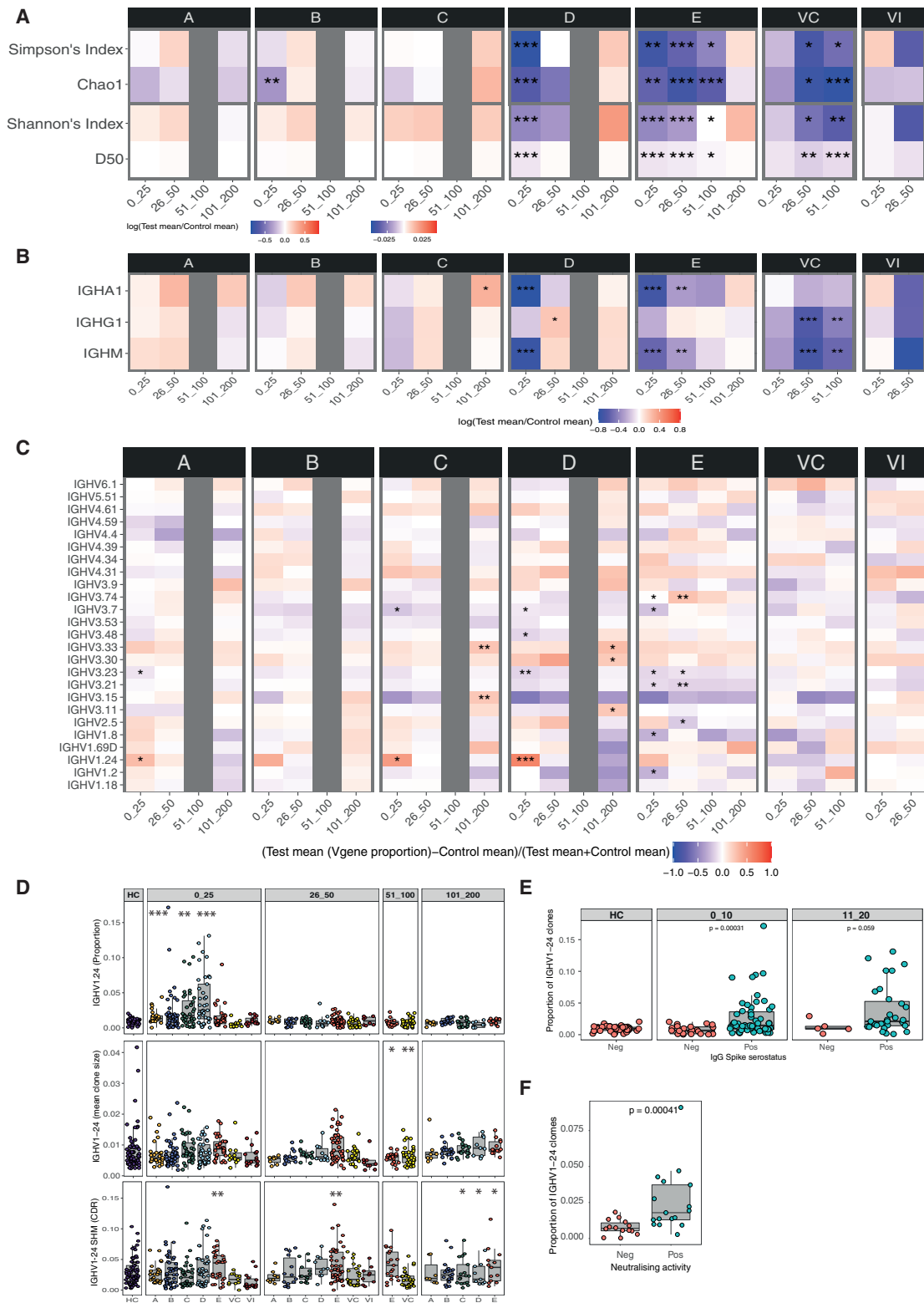
(C) Density plot modeling IGHG1 SHM across HC, SARS-CoV-2 infection and vaccination at 11–20 and 101–200 days from symptom onset.

(D) Boxplots showing the proportion of clones per patient with a mean level of SHM <0.05 nt across HC and infection and vaccination at 11–20 and 101–200 days from symptom onset. Wilcoxon test FDR adjusted p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005. Circles represent individual donors.

(E) Boxplots showing mean IGHG1 SHM per patient split according to days from symptom onset and IgG spike serostatus. Wilcoxon unpaired two-sided test. Circles represent individual donors.

(F) Boxplots showing mean IGHG1 SHM per paired patient pre and post seroconversion. Visual representation of change in somatic hypermutation post seroconversion. Points represent B cell clones. Top row represents seronegative patients. Bottom row represents paired patients post seroconversion. IGHV gene is represented on the x axis, CDR3 length on the y axis, and point color represents level of SHM.

(G) Boxplots showing mean SHM per patient split according to neutralizing activity. Circles represent individual donors. Wilcoxon unpaired two-sided test.



**Figure 3. V gene usage and diversity in COVID-19**

(A) Heatmap showing log<sub>2</sub> fold change in mean diversity indices between SARS-CoV-2 infection and vaccine cases and HC, within severity categories and across time bins post screening (cat. A), symptom onset (cat. B–E), or vaccination (cat. VC and VI). Wilcoxon test FDR adjusted p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005.

(legend continued on next page)



SHM and appeared independent of the time post symptom onset (Figures 2E and 2F). This observation also held true for serum neutralization titers, which were taken from the same bleed samples for BCR repertoire analysis (Figure 2G). Increased switching to IgG1 with seroconversion was also seen (Figure S2E). These observations are consistent with recent evidence suggesting that the early neutralizing anti-spike SARS-CoV-2 antibody response is not mutated (Kreer et al., 2020; Seydoux et al., 2020; Wang et al., 2021a), with this antigen-specific observation reflected in the BCR repertoire as a whole.

### Clonal expansion is induced in severe COVID-19 and after SARS-CoV-2 vaccination

We next assessed clonal diversity and expansion, using a number of standard measures, after subsampling to correct for varying library depth (Figures 3A and 3B). These measurements were the repertoire richness, Simpson, Shannon, and D50 indices. Richness refers to the abundance of unique clones in a repertoire (Chao1). The inverse Simpson index assesses the probability of two randomly sampled reads belonging to the same clone; the more expanded clones within the population, the greater the chance of clonal sharing. The D50 index refers to the number of unique CDR3 sequences that are present in the top 50% of sequences. A small D50 index is suggestive of large dominant clones. Shannon index is a measure of evenness, whereby the proportion of total reads represented by each clone is assessed. This metric is not influenced by the number of unique of clones but rather the distribution of size of clones in the repertoire. Thus a decrease in BCR repertoire diversity, corresponding to an increase in expanded clones, will usually be reflected in a decrease in all four indices.

There were no changes in BCR repertoire diversity in groups A, B, and C. In contrast, there was a profound reduction in diversity in groups D and E. In both groups this was most pronounced in the first 25 days, but in group E the reduction persisted out to 100 days. By 200 days, diversity had been restored in both severity groups (Figure 3A). The increased B cell clonality seen in group E is likely to be the product rather than the cause of severe disease, and it is associated with the persistence of SARS-CoV-2-specific clones (discussed below). It is unlikely to be driven by ongoing overt infection, as in severe disease viral clearance with broadly similar kinetics to milder disease is the rule in most patients (Bergamaschi et al., 2021). It may be that ongoing inflammation, together with persistent sequestered viral antigen, results in ongoing germinal center activity in these patients, thus increasing clonal expansion. Secondary infections, commonly seen in the intensive care unit (ICU) setting in COVID-19, could

also contribute to increased B cell clonality. There were similar reductions in diversity across isotypes, consistent with antigen-driven clonal expansion, after a single dose of SARS-CoV-2 vaccine; this was not observed until 26 days post vaccination and persisted out to 100 days (Figure 3B).

The kinetic recovery of diversity is shown for different isotypes using the Simpson index (Figure S3A). Reduced diversity is most prominent in the severe groups in the IgM and IgA subgroups, and appears less pronounced for IgG, while, in contrast, vaccination induces this reduction in IgM and IgG, and not IgA, perhaps reflecting the fact that vaccination does not engage mucosal immunity (Figure 3B). In keeping with this, there were significantly higher levels of IgG spike-specific antibodies compared with IgA in patients post vaccination, while, in natural infection, levels were comparable (Figure S3B).

### Expansion of IGHV1-24 after SARS-CoV-2 infection and vaccination

An examination of the contribution to the repertoire of various VH genes was then performed (Figures 3C and S3C). Two broad features were apparent. The first was that the majority of statistically significant changes in the VH gene usage were seen only in groups C, D, and E, which were most prominent early. These were thought to be most likely a consequence of the major changes in B cell subsets that occur in those with severe disease (Figure 1C). In contrast, one IGHV gene, VH1-24, was increased in all severity groups in the first time window. This V gene has been shown to be strongly associated with antibodies that recognize the NTD of the SARS-CoV-2 spike protein, conferring neutralization even in the germline state (Voss et al., 2021) (Figures 3D–3F). Consistent with this, VH1-24 proportion was strongly associated with the development of neutralizing antibodies (Figure 3F), an observation not confounded by disease duration (Figure S3D). VH1-24 was increased in proportion in the first time window studied, but at this time there was no evidence of increased SHM or clonal expansion (Figure 3D). There was a significant difference in VH1-24 proportion between groups D and E (Wilcoxon test, p value:  $4.8 \times 10^{-3}$ ) at 0–25 days from symptom onset. No increase in VH1-24 was seen after SARS-CoV-2 vaccination.

### Clonal convergence after SARS-CoV-2 infection and vaccination

We looked for overlap between BCR clones present in our study and the CoV-AbDab database, a resource detailing all published and patented antibodies shown to bind SARS-CoV-2 and other coronaviruses (Raybould et al., 2020). Convergent clones were

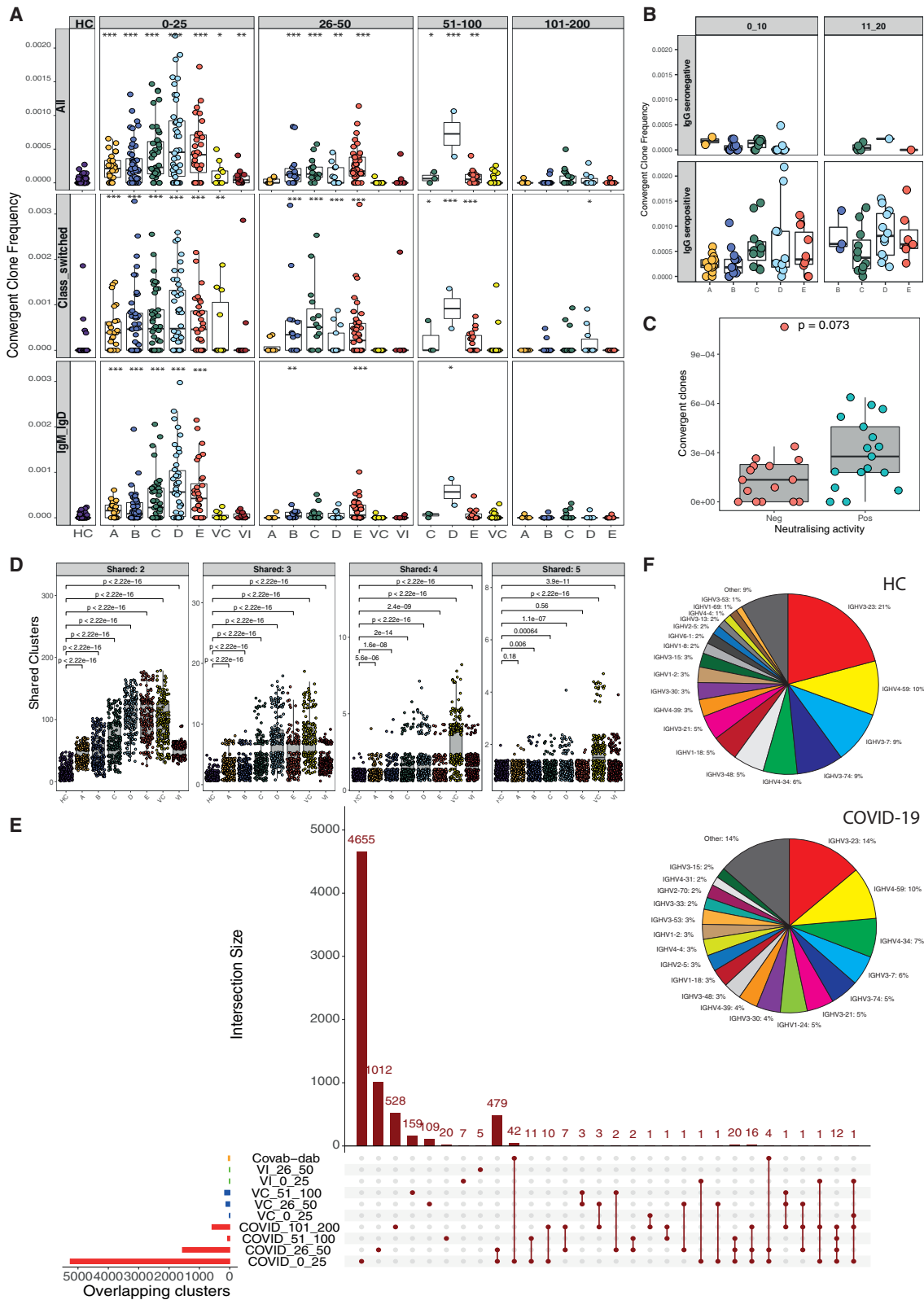
(B) Heatmap showing log<sub>2</sub> fold change in mean Simpson's diversity between SARS-CoV-2 infection and vaccine cases and HC, within severity categories and across isotypes and time bins. Wilcoxon test FDR adjusted p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005.

(C) Heatmap showing the difference between V gene proportion between SARS-CoV-2 infection and vaccine cases and HC, within severity categories and time bins. Difference calculated using the following method: (mean V gene proportion of disease – mean V gene proportion of HC)/(mean V gene proportion of disease + mean V gene proportion of HC). Wilcoxon test FDR adjusted p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005.

(D) Boxplots showing IGHV1-24 proportion and mean clone size and SHM by severity categories and time bins. Uncorrected Wilcoxon test with HC as reference group, p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005. Circles represent individual donors.

(E) Boxplots showing proportion of IGHV1-24-positive clones/per patient split according to days from symptom onset/swab and IgG spike serostatus. Wilcoxon test p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005.

(F) Boxplots showing proportion of IGHV1-24-positive clones/per patient split according to neutralization ability. Wilcoxon test p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005.



(legend on next page)

defined by sharing of IGHV and IGHJ genes, having identical CDR-H3 region length and having CDR-H3 sequences that show 85% amino acid homology, and thus likely to have a similar antigen specificity to the reference antibodies. We found clonotype convergence in both IGHD/IGHM clones in COVID-19 as well as in patients vaccinated against SARS-CoV-2, within the first 25 days from symptom onset/vaccination (median proportion of convergent clones in HC, 0; A, 0.00016; B, 0.000088; C, 0.00023; D, 0.00057; E, 0.00043; VC, 0; VI, 0) and class-switched clones (median proportion HC, 0; A, 0.00039; B, 0.00047; C, 0.00047; D, 0.00048; E, 0.00046; VC, 0; VI, 0). At 26–50 days, clonal convergence decreased in all isotypes but most markedly in IGHD/IGHM (median proportion HC, 0; A, 0; B, 0.000037; C, 0; D, 0; E, 0.000059; VC, 0; VI, 0), consistent with class switching of antigen-specific clones (Figure 4A). With infection, there was a significant convergence of class-switched and non-class-switched clones described to neutralize and target the RBD and NTD of spike. In contrast, after SARS-CoV-2 vaccination, COVID-19-specific clones were mainly class switched and predicted to target the RBD (Figure S4A).

Increased convergence was present in patients who had seroconverted as well as in patients with neutralizing antibodies (Figures 4B and 4C). Clone tracking of COVID-19-specific BCRs sampled at multiple timepoints showed progressive somatic hypermutation with time but variable class switching (Figure S4B).

To assess if there was increased clonal sharing within patients with COVID-19 (0–25 days from swab/symptom onset) and/or patients vaccinated (26–50 days from vaccination) compared with health, we calculated the number of shared clusters pairwise up to five patients. We performed this in nine patients to accommodate the smallest group size. We performed 200 permutations where we randomly selected nine patients within a disease group, and then 4,000 unique clusters per person. This showed greater sharing in patients with COVID-19, suggestive of increased BCR overlap driven by shared antigen through exposure (Figure 4D). Similarly, post SARS-CoV-2 vaccination, there was increased convergence, with a greater number of shared clusters among four or more people compared with natural infection or health (Figure 4D).

In order to identify new clones that might be COVID-19 specific, we looked for convergent clones that were shared among a minimum of three patients at a given time interval and that were not present in healthy controls (Figures 4E and S4C). We found convergence of clonotypes consistent with shared antigen driving selection of clones. Although overlap with the CoV-AbDab database was minimal, likely due to limitations

of the database, clones were present that were shared in over 10 patients (Figure S4D). Assessing V gene usage of these disease-associated clones at 0–25 days from symptom onset revealed an increase in representation of IGHV4-34 and IGHV1-24 compared with health (Figure 4F). IGHV4-34 displays auto-reactivity against self-antigen and is associated with systemic lupus erythematosus (Bashford-Rogers et al., 2019; Jenks et al., 2018; Scharer et al., 2019; Tipton et al., 2015). Parallels have been drawn between both conditions, with similarities in extrafollicular pathway activation and expansion of double-negative B cells and the presence of lowly mutated clones (Consortium, Co.-19 M.B.At. et al., 2021; Galson et al., 2020; Jenks et al., 2018; Woodruff et al., 2020).

## DISCUSSION

To assess the nature of the B cell response to SARS-CoV-2 infection, we measured how BCR repertoires change in patients with COVID-19 stratified by both disease severity and time post infection. An increase in the proportion of BCRs bearing IgG1, IgG3, and IgA1 isotypes was seen in all groups soon after infection, including asymptomatic individuals. In parallel, decreased SHM was also seen in all groups, with the reduction particularly marked in the three expanded isotypes (IgG1, IgG3, and IgA1). These observations are consistent with the initial changes in the BCR repertoire being driven by an early plasmablast expansion, as they are seen in severity groups in which this is the only B cell subset that changes compared with healthy controls (Bergamaschi et al., 2021). The implication, then, is that these early plasmablasts are unmutated, in keeping with the finding that early potent neutralizing antibodies have near-germline sequences (Kim et al., 2021a; Kreer et al., 2020; Seydoux et al., 2020; Sokal et al., 2021; Wang et al., 2021a). In support of this and previous studies, we show that a decrease in SHM correlates with seroconversion (Nielsen et al., 2020) and neutralization. These plasmablasts could be derived from the rapid differentiation of already switched but unmutated precursor cells from, for example, the marginal zone; the relative absence of SHM suggests that differentiated memory B cells are not a major source. They may also be generated rapidly by isotype switching and differentiation of naive B cells to plasmablasts outside the germinal center (Smith et al., 1996). Given both the reduction in germinal centers and circulating as well as tissue CD4 T<sub>FH</sub> cells observed in patients with severe COVID-19, this seems a more likely source of plasmablasts than switching in, and export from, early germinal centers (Bergamaschi et al., 2021; Kaneko

### Figure 4. Clonal convergence after SARS-CoV-2 infection and vaccination

(A) Convergent clone frequency across SARS-CoV-2 infection and vaccine cases and healthy controls with the CoV-AbDab database. This represents the percentage of unique clones in each patient that are also found in the CoV-AbDab database. Samples split by severity categories and time bins post screening (cat. A), symptom onset (cat. B–E), or vaccination. One-sided Wilcoxon test FDR adjusted p value: \*p < 0.05, \*\*p < 0.005, \*\*\*p < 0.0005. Circles represent individual donors.

(B) Boxplots showing convergence per patient split according to days from symptom onset and IgG spike serostatus. Circles represent individual donors.

(C) Boxplots showing convergence per patient split according to neutralization ability. Circles represent individual donors. Two-sided Wilcoxon test.

(D) Bar plot representing the number of clonotypes shared by members within a group. Patients with SARS-CoV-2 infection are within 25 days from symptom onset/swab, and participants post vaccine are within 25–50 days. Two-sided Wilcoxon test.

(E) Graph of convergent IGH clusters. The number of convergent clusters within a disease group are represented by the horizontal bars and across disease groups by lines. An unconnected dot indicates no sharing. The vertical histogram bars represent subtotals.

(F) Pie chart comparing V gene usage of convergent clusters in HC and COVID-19 within 25 days from symptom onset/swab identified from (E).

et al., 2020; Stephenson et al., 2021). Evidence of clonal evolution in the form of an increase in SHM, neutralization potency, and broadening of the repertoire is, however, observed at 6 months post infection (Gaebler et al., 2021; Sokal et al., 2021; Wang et al., 2021a). This, together with the detection of SARS-CoV-2-specific long-lived plasma cells in bone marrow aspirates at 11 months post infection (Turner et al., 2021), demonstrates that an effective germinal center response does occur, despite it not being prominently visible in the early BCR repertoire in the circulation.

In those with more severe disease (groups D and E), increased clonal expansion, as shown by a reduction in repertoire diversity, is seen. The clonal expansion is most prominent in IgM and IgA, and not in IgG. This highlights the important role that IgA B cell memory may play after mucosal infection with SARS-CoV-2, which usually first infects the upper respiratory tract mucosa. In addition, dimerized IgA antibodies, the form predominantly found in mucosal tissues, are more potent neutralizers of SARS-CoV-2 than IgG or monomeric IgA (Wang et al., 2021c), and IgA-virus-immune complexes are potent inducers of netosis via engagement of Fc- $\alpha$ RI on neutrophils, which can enhance anti-viral defense (Stacey et al., 2021).

Across all severity groups, VH1-24 was the dominant VH gene expanded after SARS-CoV-2 infection. Antibodies bearing this VH gene have been noted to make up the majority of neutralizing IgG antibodies arising after SARS-CoV-2 infection, with a specificity for the NTD component of the spike antigen, rather than the RBD, which had first been assumed to be the main target of neutralizing antibodies (Voss et al., 2021). Given these observations, the fact that the proportional increase in VH1-24 in group E is substantially lower than less severe groups raises the possibility that a robust early VH1-24 response might be important in controlling infection, and that a failure to mount one might predispose to severe disease outcomes.

Changes in the repertoire following SARS-CoV-2 vaccination were less pronounced than after COVID-19 infection. Isotype usage showed an increase in IGHM and D and a decrease in IGHA1/2, the opposite pattern to that seen in natural infection. There were no changes in SHM or specific heavy-chain usage. This is in contrast to both natural infections, where SHM rates fall early, and to influenza vaccination, where they rise. The most likely explanation for this is that SARS-CoV-2 vaccination may not induce the early pronounced plasmablast response characteristic of SARS-CoV-2 infection, which is likely driven by a more intense and broadly distributed antigen exposure in the latter. In contrast, influenza vaccination is likely to have high SHM rates due to more prominent memory cell reactivation than seen with SARS-CoV-2, due to previous influenza exposure. The relative role of memory reactivation and adjuvant in explaining this difference could be further explored by examining the BCR repertoire in patients exposed to both SARS-CoV-2 and influenza vaccines that utilize different antigenic strategies, with different documented previous antigenic exposures. In this context, repertoire analysis could assist in development of vaccine strategies, optimizing adjuvants, delivery route, etc. to increase the potency, breadth, and longevity of the memory response. Despite this, convergence analysis with the CoV-AbDab database revealed the generation of

SARS-CoV-2 spike-specific clones post vaccination, and significantly increased clonal sharing compared with healthy controls. When considering clonal overlap in a minimum of five people, there were more shared clones generated after SARS-CoV-2 vaccination than in any post-infection severity group. This suggests a focusing of the immune response onto a narrow range of antigens by SARS-CoV-2 vaccination. In contrast, influenza vaccination resulted in clear global changes in the repertoire, with isotype changes mirroring SARS-CoV-2 infection, and with an increase in SHM, likely reflecting an expansion of cross-reactive antibodies and a reactivation of memory cells generated by previous exposures and vaccinations (Turner et al., 2020). Clonal expansion is, however, seen post SARS-CoV-2 vaccination and occurs with similar kinetics to that seen in response to natural infection. This appears to be driven primarily by expansion of clones bearing IgM and IgG isotypes rather than the IgA that predominates after natural infection, a finding further supported by vaccination inducing higher spike-specific IgG antibody titers compared with IgA. This is likely to reflect the different anatomical compartments involved in early immune responses to infection versus vaccination, with the latter delivered in a way that cannot generate a significant mucosal response (Röltgen and Boyd, 2021). The inability of systemic vaccination to induce mucosal IgA or tissue resident memory T cell responses may limit efficacy against respiratory pathogens (Jeyanathan et al., 2020; Moradi-kalbolandi et al., 2021; Russell et al., 2020). Given mucosal immune responses are compartmentalized, with intranasal vaccines inducing a response in the upper and lower respiratory tracts, a vaccine utilizing this approach would neutralize pathogens at the site of entry (Moradi-kalbolandi et al., 2021; Russell et al., 2020). In support of this, a chimpanzee adenovirus-vectored vaccine administered intranasally to Syrian hamsters resulted in less viral load and lung pathology upon challenge compared with intramuscular administration (Bricker et al., 2021).

Vaccination elicited COVID-19-specific clones that were class switched, likely neutralizing and mainly targeted the RBD of the spike protein. However, the level of convergence with the CoV-AbDab database was far lower than that seen in severe disease, with minimal formation of antibodies targeting NTD compared with natural infection. The reason for the under-representation of NTD-specific clones is unclear given that the BNT162B2 SARS-CoV-2 vaccine utilizes the sequence of the full-length SARS-CoV-2 spike protein, including the NTD. The lack of NTD-convergent clones may represent a limitation of the CoV-AbDab database or that NTD-binding clones are rarer and thus less likely to be sampled. If a true difference in vaccination and infection is present, it might be that future vaccine design strategies might be developed to increase the immunogenicity of the NTD, but it is also worth bearing in mind that there is a divergence of views on the importance of this antigen in the generation of neutralizing antibodies (Greaney et al., 2021; Stamatos et al., 2021).

Global BCR repertoire analysis is a useful adjunct to that of antigen-specific B cell responses, providing additional context that can inform vaccine strategies. We show temporal changes in the BCR repertoire in response to natural SARS-CoV-2

infection that are dynamically and compositionally distinct from those generated by vaccination. SARS-CoV-2 natural infection results in activation of mucosal immunity with clonal expansion in IgM and IgA isotypes and an increase in VH1-24 in the circulation. SARS-CoV-2 vaccination induces very different isotype changes to natural infection, results in clonal expansion of IgM and IgG, and appears to focus the immune response to the RBD.

### Limitations of the study

We recognize that a limitation of the study is that neither the antigenic specificity nor the neutralizing capacity of antibodies encoded by identified BCR sequences was determined experimentally. Instead, function was inferred by similarity to sequences deposited in the CoV-AbDab database. While this is an excellent and growing resource, this approach is limited in breadth, with a particular bias toward the identification of RBD-binding clones, and likely results in the under reporting of SARS-CoV-2-specific clones in our dataset. Future work would include generating monoclonal antibodies from convergent IGH sequences to allow further characterization. In addition, BCR repertoire changes in response to a first dose of the BNT162b2 vaccine were analyzed; future studies would include different vaccines and responses to subsequent “booster” doses. In addition, the analysis of the BCR repertoire of flow-sorted B cell subsets would have enabled a more granular delineation of how SARS-CoV-2 infection affects the BCR repertoire.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - COVID-19
  - SARS-CoV-2 vaccination
  - Influenza vaccination
  - Healthy controls
- **METHOD DETAILS**
  - Clinical data collection
  - Study volunteers were classified in 5 groups
  - Peripheral blood mononuclear cell preparation and flow cytometry immunophenotyping
  - Total immunoglobulin levels
  - SARS-CoV-2 serology
  - SARS-CoV-2 vaccine recipients
  - SARS-CoV-2 neutralisation assay
  - B cell receptor repertoire
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2022.110393>.

### CONSORTIA

The members of The CITIID-NIHR COVID-19 BioResource Collaboration are as follows. Principal investigators: Stephen Baker, John Bradley, Gordon Dougan, Christoph Hess, Ian Goodfellow, Ravi Gupta, Nathalie Kingston, Paul J. Lehner, Paul A. Lyons, Nicholas J. Matheson, Willem H. Owehand, Caroline Saunders, Kenneth G.C. Smith, Charlotte Summers, James E.D. Thaventhiran, Mark Toshner, and Michael P. Weekes; CRF and volunteer research nurses: Ashlea Bucke, Jo Calder, Laura Canna, Jason Domingo, Anne Elmer, Stewart Fuller, Julie Harris, Sarah Hewitt, Jane Kennet, Sherly Jose, Jenny Kourampa, Anne Meadows, Criona O'Brien, Jane Price, Cherry Publico, Rebecca Rastall, Carla Ribeiro, Jane Rowlands, Valentina Ruffolo, and Hugo Tordesillas; sample logistics: Ben Bullman, Benjamin J. Dunmore, Stuart Fawke, Stefan Gräf, Josh Hodgson, Christopher Huang, Kelvin Hunter, Emma Jones, Ekaterina Legchenko, Cecilia Matará, Jennifer Martin, Ciara O'Donnell, Linda Pointon, Nicole Pond, Joy Shih, Rachel Sutcliffe, Tobias Tilly, Carmen Treacy, Zhen Tong, Jennifer Wood and Marta Wylot; sample processing: Laura Bergamaschi, Ariana Betancourt, Georgie Bower, Aloka De Sa, Madeline Epping, Stuart Fawke, Oisín Huhn, Sarah Jackson, Isobel Jarvis, Jimmy Marsden, Francesca Nice, Georgina Okecha, Ommar Omarjee, Marianne Perera, Nathan Richoz, Rahul Sharma, and Lori Turner; clinical data collection: Eckart M.D.D. De Bie, Katherine Bunclark, Masa Josipovic, Michael Mackay, Federica Mescia, Alice Michael, Sabrina Rossi, Mayurun Selvan, Sarah Spencer, and Cissy Yong; Royal Papworth Hospital ICU: Ali Ansaripour, Alice Michael, Lucy Mwaura, Caroline Patterson, and Gary Polwarth; Addenbrooke's Hospital ICU: Petra Polgarova and Giovanni di Stefano; NIHR BioResource: John Allison, Heather Biggs, Helen Butcher, Daniela Caputo, Matt Chandler, Patrick F. Chinnery, Debbie Clapham-Riley, Anne-Maree Dean, Eleanor Dewhurst, Christian Fernandez, Anita Furlong, Anne George, Barbara Graves, Jennifer Gray, Sabine Hein, Tasmin Ivers, Mary Kasanicki, Emma Le Gresley, Rachel Linger, Sarah Meloy, Alexei Moulton, Francesca Muldoon, Nigel Ovington, Sofia Papadia, Roxana Paraschiv, Christopher Penkett, Isabel Phelan, Venkatesh Ranganath, Jennifer Sambrook, Katherine Schon, Hannah Stark, Kathleen E. Stirrups, Paul Townsend, Julie von Ziegenweid, Neil Walker, and Jennifer Webster.

### ACKNOWLEDGMENTS

We thank all the patients and health care workers who consented to take part in this study. We are grateful to CVC Capital Partners, the Evelyn Trust (20/75), Addenbrooke's Charitable Trust, Cambridge University Hospitals (12/20A), the NIHR Cambridge Biomedical Research Centre, and the UKRI/NIHR through the UK Coronavirus Immunology Consortium (UK-CIC) for their financial support. Further support: K.G.C.S., Wellcome Investigator Award (200871/Z/16/Z); C.H., Wellcome COVID-19 Rapid Response DCF and the Fondation Botnar; N.M., MRC (CSF MR/P008801/1), NHSBT (WPA15-02), and Addenbrooke's Charitable Trust, (grant ref. to 900239 N.J.M.); I.G.G. and Wellcome Senior Fellowship and Wellcome grant (Ref: 207498/Z/17/Z). N.J.M. was funded by the MRC (CSF MR/P008801/1), NHSBT (WPA15-02), and Addenbrooke's Charitable Trust (grant ref. to 900239 N.J.M.); R.K.G. is supported by a Wellcome Trust Senior Fellowship in Clinical Science (WT108082AIA). Z.K.T. and M.R.C. are supported by a Medical Research Council Human Cell Atlas Research Grant (MR/S035842/1). M.R.C. is supported by an NIHR Research Professorship (RP-2017-08-ST2-002). P.K. is the recipient of a Jacquot Research Entry Scholarship of the Royal Australasian College of Physicians Foundation. W.M.R. is funded by the Wellcome Trust (216382/Z/19/Z). We would like to thank the NIHR Cambridge Clinical Research Facility outreach team for enrollment of patients; the NIHR Cambridge Biomedical Research Center Cell Phenotyping Hub and the CRUK Cambridge Institute flow cytometry core facility for flow and mass cytometry; and the Cambridge NIHR BRC Stratified Medicine Core Laboratory NGS Hub (supported by an MRC Clinical Infrastructure Award) for BCR sequencing. We thank G. Manferrari, L. Overend, and F. Tucci for helpful discussions. Graphical abstract created with [BioRender.com](https://www.biorender.com)

## AUTHOR CONTRIBUTIONS

Conceptualization, C.H., J.R.B., P.K., P.A.L., and K.G.C.S.; software, K.H.; data acquisition, P.K., F.M., L.T., L.B., P.P.G., M.H., M.R.W., R.K.G., E.F.M., N.J.M., and I.G.G.; data analysis, P.K., F.M., L.T., L.B., P.P.G., M.H., W.M.R., Z.K.T., M.R.W., and M.R.C.; project administration, F.M., L.T., and L.B.; funding acquisition, C.H., J.R.B., P.A.L., and K.G.C.S.; writing – original draft, P.K., P.A.L., and K.G.C.S.; writing – review & editing, P.K., R.J.M.B.-R., C.H., J.R.B., M.R.C., J.N., P.A.L., and K.G.C.S.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 18, 2021

Revised: November 28, 2021

Accepted: January 24, 2022

Published: February 15, 2022

## REFERENCES

- Amanat, F., Thapa, M., Lei, T., Sayed Ahmed, S.M., Adelsberg, D.C., Carreno, J.M., Strohmeier, S., Schmitz, A.J., Zafar, S., Zhou, J.Q., et al. (2021). SARS-CoV-2 mRNA vaccination induces functionally diverse antibodies to NTD, RBD and S2. *Cell* **184**, 3936–3948.e10.
- Arunachalam, P.S., Wimmers, F., Mok, C.K.P., Perera, R.A.P.M., Scott, M., Hagan, T., Sigal, N., Feng, Y., Bristow, L., Tsang, O.T.Y., et al. (2020). Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* **369**, 1210–1220.
- Bashford-Rogers, R.J.M., Bergamaschi, L., McKinney, E.F., Pombal, D.C., Mescia, F., Lee, J.C., Thomas, D.C., Flint, S.M., Kellam, P., Jayne, D.R.W., et al. (2019). Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nature* **574**, 122–126.
- Bergamaschi, L., Mescia, F., Turner, L., Bradley, J.R., Lyons, P.A., and Smith, K.G.C. (2021). Longitudinal analysis reveals that delayed bystander CD8+ T cell activation and early immune pathology distinguish severe COVID-19 from mild disease. *Immunity* **54**, 1257–1275.e8.
- Bricker, T.L., Darling, T.L., Hassan, A.O., Harastani, H.H., Soung, A., Jiang, X., Dai, Y.N., Zhao, H., Adams, L.J., Holtzman, M.J., et al. (2021). A single intranasal or intramuscular immunization with chimpanzee adenovirus-vectored SARS-CoV-2 vaccine protects against pneumonia in hamsters. *bioRxiv*, 36.
- Brochet, X., Lefranc, M.P., and Giudicelli, V. (2008). IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* **36**, W503–W508.
- Buckland, M.S., Galloway, J.B., Fhogartaigh, C.N., Meredith, L., Provine, N.M., Bloor, S., Ogbé, A., Zelek, W.M., Smielewska, A., Yakovleva, A., et al. (2020). Treatment of COVID-19 with remdesivir in the absence of humoral immunity: a case report. *Nat. Commun.* **11**, 1–11.
- Collier, D.A., Ferreira, I.A.T.M., Kotagiri, P., Datir, R., Lim, E., Touizer, E., Meng, B., Abdullahi, A., Elmer, A., Kingston, N., et al. (2021). Age-related immune response heterogeneity to SARS-CoV-2 vaccine BNT162b2. *Nature* **596**, 417–422.
- Consortium, Co.-19 M.B.At. (COMBAT); Ahern, D.J., Ai, Z., Ainsworth, M., Allan, C., Allcock, A., Ansari, A., Arancibia-Carcamo, C. v, Aschenbrenner, D., Attar, M., et al. (2021). A blood atlas of COVID-19 defines hallmarks of disease severity and specificity COvid-19 Multi-omics Blood Atlas (COMBAT) Consortium 1\* 1. *MedRxiv*, 2021.05.11.21256877.
- Cox, R.J., and Brokstad, K.A. (2020). Not just antibodies: B cells and T cells mediate immunity to COVID-19. *Nat. Rev. Immunol.* **20**, 581–582.
- Daly, J.L., Simonetti, B., Klein, K., Chen, K.E., Williamson, M.K., Antón-Plágaro, C., Shoemark, D.K., Simón-Gracia, L., Bauer, M., Holland, R., et al. (2020). Neupilin-1 is a host factor for SARS-CoV-2 infection. *Science* **370**, 861–865.
- Davis, C.W., Jackson, K.J.L., McElroy, A.K., Halfmann, P., Huang, J., Chenreddy, C., Piper, A.E., Leung, Y., Albariño, C.G., Crozier, I., et al. (2019). Longitudinal Analysis of the Human B Cell Response to Ebola Virus Infection. *Cell* **177**, 1566–1582.e17.
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152.
- Gaebler, C., Wang, Z., Lorenzi, J.C.C., Muecksch, F., Finkin, S., Tokuyama, M., Cho, A., Jankovic, M., Schaefer-Babajew, D., Oliveira, T.Y., et al. (2021). Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639–644.
- Galsion, J.D., Schatzle, S., Bashford-Rogers, R.J.M., Raybould, M.I.J., Kovaltsuk, A., Kilpatrick, G.J., Minter, R., Finch, D.K., Dias, J., James, L.K., et al. (2020). Deep sequencing of B cell receptor repertoires from COVID-19 patients reveals strong convergent immune signatures. *Front. Immunol.* **11**, 3283.
- García-Beltrán, W.F., Lam, E.C., Astudillo, M.G., Yang, D., Miller, T.E., Feldman, J., Hauser, B.M., Caradonna, T.M., Clayton, K.L., Nitido, A.D., et al. (2021). COVID-19-neutralizing antibodies predict disease severity and survival. *Cell* **184**, 476–488.e11.
- Greaney, A.J., Loes, A.N., Gentles, L.E., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., and Bloom, J.D. (2021). Antibodies elicited by mRNA-1273 vaccination bind more broadly to the receptor binding domain than do those from SARS-CoV-2 infection.
- Godoy-Lozano, E.E., Téllez-Sosa, J., Sánchez-González, G., Sámano-Sánchez, H., Aguilar-Salgado, A., Salinas-Rodríguez, A., Cortina-Ceballos, B., Vivanco-Cid, H., Hernández-Flores, K., Pfaff, J.M., et al. (2016). Lower IgG somatic hypermutation rates during acute dengue virus infection is compatible with a germinal center-independent B cell response. *Genome Med.* **8**, 23.
- Gupta, N.T., vander Heiden, J.A., Uduman, M., Gadala-Maria, D., Yaari, G., and Kleinstein, S.H. (2015). Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358.
- Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Bousier, J., Smith, N., Péré, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., et al. (2020). Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718–724.
- Hernandez, A.M., and Holodick, N.E. (2017). Editorial: natural antibodies in health and disease. *Front. Immunol.* **8**, 1795.
- Horns, F., Dekker, C.L., and Quake, S.R. (2020). Memory B cell activation, broad anti-influenza antibodies, and bystander activation revealed by single-cell transcriptomics. *Cell Rep.* **30**, 905–913.e6.
- Jenks, S.A., Cashman, K.S., Zumaquero, E., Marigorta, U.M., Patel, A.V., Wang, X., Tomar, D., Woodruff, M.C., Simon, Z., Bugrovsky, R., et al. (2018). Distinct effector B cells induced by unregulated toll-like receptor 7 contribute to pathogenic responses in systemic lupus erythematosus. *Immunity* **49**, 725–739.e6.
- Jeyanathan, M., Afkhami, S., Smail, F., Miller, M.S., Lichty, B.D., and Xing, Z. (2020). Immunological considerations for COVID-19 vaccine strategies. *Nat. Rev. Immunol.* **20**, 615–632.
- Joyner, M.J., Carter, R.E., Senefeld, J.W., Klassen, S.A., Mills, J.R., Johnson, P.W., Theel, E.S., Wiggins, C.C., Bruno, K.A., Klompas, A.M., et al. (2021). Convalescent plasma antibody levels and the risk of death from Covid-19. *New Engl. J. Med.* **384**, 1015–1027.
- Kalinke, U., Bucher, E.M., Ernst, B., Oxenius, A., Roost, H.P., Geley, S., Kofler, R., Zinkernagel, R.M., and Hengartner, H. (1996). The role of somatic mutation in the generation of the protective humoral immune response against vesicular stomatitis virus. *Immunity* **5**, 639–652.
- Kaneko, N., Kuo, H.H., Boucau, J., Farmer, J.R., Allard-Chamard, H., Mahajan, V.S., Piechocka-Trocha, A., Lefteri, K., Osborn, M., Bals, J., et al. (2020). Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. *Cell* **183**, 143–157.e13.
- Kemp, S.A., Collier, D.A., Datir, R.P., Ferreira, I.A.T.M., Gayed, S., Jahun, A., Hosmillo, M., Rees-Spear, C., Mlcochova, P., Lumb, I.U., et al. (2021). SARS-CoV-2 evolution during treatment of chronic infection. *Nature* **592**, 277–282.
- Kim, S. Il, Noh, J., Kim, S., Choi, Y., Yoo, D.K., Lee, Y., Lee, H., Jung, J., Kang, C.K., Song, K.H., et al. (2021a). Stereotypic neutralizing VH antibodies against SARS-CoV-2 spike protein receptor binding domain in patients with COVID-19 and healthy individuals. *Sci. Translational Med.* **13**, eabd6990.

- Kim, Y. Il, Kim, S.M., Park, S.J., Kim, E.H., Yu, K.M., Chang, J.H., Kim, E.J., Casel, M.A.B., Rollon, R., Jang, S.G., et al. (2021b). Critical role of neutralizing antibody for SARS-CoV-2 reinfection and transmission. *Emerging Microbes and Infections* *10*, 152–160.
- Kreer, C., Zehner, M., Weber, T., Ercanoglu, M.S., Giesemann, L., Rohde, C., Halwe, S., Korenkov, M., Schommers, P., Vanshylla, K., et al. (2020). Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell* *182*, 843–854.e12.
- Kuri-Cervantes, L., Pampena, M.B., Meng, W., Rosenfeld, A.M., Ittner, C.A.G., Weisman, A.R., Agyekum, R.S., Mathew, D., Baxter, A.E., Vella, L.A., et al. (2020). Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci. Immunol.* *5*, eabd7114.
- Laing, A.G., Lorenc, A., del Molino del Barrio, I., Das, A., Fish, M., Monin, L., Muñoz-Ruiz, M., McKenzie, D.R., Hayday, T.S., Francos-Quijorna, I., et al. (2020). A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat. Med.* *26*, 1623–1635.
- Libster, R., Pérez Marc, G., Wappner, D., Coviello, S., Bianchi, A., Braem, V., Esteban, I., Caballero, M.T., Wood, C., Berrueta, M., et al. (2021). Early high-titer plasma therapy to prevent severe Covid-19 in older adults. *New Engl. J. Med.* *384*, 610–618.
- Mann, E.R., Menon, M., Knight, S.B., Konkol, J.E., Jagger, C., Shaw, T.N., Krishnan, S., Rattray, M., Ustianowski, A., Bakerly, N.D., et al. (2020). Longitudinal immune profiling reveals key myeloid signatures associated with COVID-19. *Sci. Immunol.* *5*, eabd6197.
- Mathew, D., Giles, J.R., Baxter, A.E., Oldridge, D.A., Greenplate, A.R., Wu, J.E., Alanio, C., Kuri-Cervantes, L., Pampena, M.B., D'Andrea, K., et al. (2020). Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* *369*, eabc8511.
- Moradi-kalbolandi, S., Majidzadeh-A, K., Abdolvahab, M.H., Jalili, N., and Farahmand, L. (2021). The role of mucosal immunity and recombinant probiotics in SARS-CoV2 vaccine development. *Probiotics Antimicrob. Proteins* *1*, 3.
- Nielsen, S.C.A., Yang, F., Jackson, K.J.L., Hoh, R.A., Röltgen, K., Jean, G.H., Stevens, B.A., Lee, J.Y., Rustagi, A., Rogers, A.J., et al. (2020). Human B cell clonal expansion and convergent antibody responses to SARS-CoV-2. *Cell Host and Microbe* *28*, 516–525.e5.
- Patterson, E.I., Prince, T., Anderson, E.R., Casas-Sanchez, A., Smith, S.L., Cansado-Utrilla, C., Solomon, T., Griffiths, M.J., Acosta-Serrano, Á., Turtle, L., et al. (2020). Methods of inactivation of SARS-CoV-2 for downstream biological assays. *J. Infect. Dis.* *222*, 1462–1467.
- Pereyra Gerber, P., Duncan, L.M., Greenwood, E.J., Marelli, S., Teixeira-Silva, A., Crozier, T.W., Gabaev, I., Zhan, J.R., Protasio, V., and Matheson, N.J. (2021). Protease-activatable biosensors of SARS-CoV-2 infection for cell. *BioRxiv*, 2021.03.22.435957.
- Pinheiro, J., Bates, D., DebRoy, S., and Sarkar, D.; R Core Team (2014). *Nlme: Linear and Nonlinear Mixed Effects Models*. R Package Version 3, 1–117.
- Quast, I., and Tarlinton, D. (2021). B cell memory: understanding COVID-19. *Immunity* *54*, 205–210.
- Raybould, M.I.J., Kovaltsuk, A., Marks, C., and Deane, C.M. (2020). CoV-Ab-Dab: the coronavirus antibody database. *Bioinformatics* *37*, 734–735.
- Rivett, L., Sridhar, S., Sparkes, D., Routledge, M., Jones, N.K., Forrest, S., Young, J., Pereira-Dias, J., Hamilton, W.L., Ferris, M., et al. (2020). Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission. *eLife* *9*, e58728.
- Robbiani, D.F., Gaebler, C., Muecksch, F., Lorenzi, J.C.C., Wang, Z., Cho, A., Agudelo, M., Barnes, C.O., Gazumyan, A., Finkin, S., et al. (2020). Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* *584*, 437–442.
- Röltgen, K., and Boyd, S.D. (2021). Antibody and B cell responses to SARS-CoV-2 infection and vaccination. *Cell Host Microbe* *29*, 1063–1075.
- Russell, M.W., Moldoveanu, Z., Ogra, P.L., and Mestecky, J. (2020). Mucosal immunity in COVID-19: a neglected but critical aspect of SARS-CoV-2 infection. *Front. Immunol.* *11*, 3221.
- Rydzynski Moderbacher, C., Ramirez, S.I., Dan, J.M., Grifoni, A., Hastie, K.M., Weiskopf, D., Belanger, S., Abbott, R.K., Kim, C., Choi, J., et al. (2020). Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* *183*, 996–1012.e19.
- Scharer, C.D., Blalock, E.L., Mi, T., Barwick, B.G., Jenks, S.A., Deguchi, T., Cashman, K.S., Neary, B.E., Patterson, D.G., Hicks, S.L., et al. (2019). Epigenetic programming underpins B cell dysfunction in human SLE. *Nat. Immunol.* *20*, 1071–1082.
- Schatz, D.G., and Swanson, P.C. (2011). V(D)J recombination: mechanisms of initiation. *Annu. Rev. Genet.* *45*, 167–202.
- Schulte-Schrepping, J., Reusch, N., Paclik, D., Saliba, A.-E., Sander, L.E., Baßler, K., Schlickeiser, S., Zhang, B., Krä, B., Krammer, T., et al. (2020). Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* *182*, 1419–1440.e23.
- Schultheiß, C., Paschold, L., Ciesek, S., Addo, M., Correspondence, M.B., Simnica, D., Mohme, M., Willscher, E., Von Wenserski, L., Scholz, R., et al. (2020). Next-generation sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease. *Immunity* *53*, 442–455.e4.
- Seydoux, E., Homad, L.J., MacCamy, A.J., Parks, K.R., Hurlburt, N.K., Jennewein, M.F., Akins, N.R., Stuart, A.B., Wan, Y.H., Feng, J., et al. (2020). Analysis of a SARS-CoV-2-infected individual reveals development of potent neutralizing antibodies with limited somatic mutation. *Immunity* *53*, 98–105.e5.
- Shugay, M., Bagaev, D.v., Turchaninova, M.A., Bolotin, D.A., Britanova, O.v., Putintseva, E.v., Pogorelyy, M.v., Nazarov, V.I., Zvyagin, I.v., Kirgizova, V.I., et al. (2015). VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput. Biol.* *11*, 1004503.
- Smith, K.G.C., Hewitson, T.D., Nossal, G.J.V., and Tarlinton, D.M. (1996). The phenotype and fate of the antibody-forming cells of the splenic foci. *Eur. J. Immunol.* *26*, 444–448.
- Smith, K.G.C., Light, A., Nossal, G.J.V., and Tarlinton, D.M. (1997). The extent of affinity maturation differs between the memory and antibody-forming cell compartments in the primary immune response. *EMBO J.* *16*, 2996–3006.
- Sokal, A., Chappert, P., Barba-Spaeth, G., Roeser, A., Fourati, S., Azzaoui, I., Vandenberghe, A., Fernandez, I., Meola, A., Bouvier-Alias, M., et al. (2021). Maturation and persistence of the anti-SARS-CoV-2 memory B cell response. *Cell* *184*, 1201–1213.e14.
- Song, G., He, W., Callaghan, S., Anzanello, F., Huang, D., Ricketts, J., Torres, J.L., Beutler, N., Peng, L., Vargas, S., et al. (2021). Cross-reactive serum and memory B-cell responses to spike protein in SARS-CoV-2 and endemic coronavirus infection. *Nat. Commun.* *12*, 2938.
- Stacey, H.D., Golubeva, D., Posca, A., Ang, J.C., Novakowski, K.E., Zahoor, M.A., Kaushic, C., Cairns, E., Bowdish, D.M.E., Mullarkey, C.E., et al. (2021). IgA potentiates NETosis in response to viral infection. *Proc Natl Acad Sci USA* *118*, e2101497118.
- Stamatatos, L., Czartoski, J., Wan, Y.-H., Homad, L.J., Rubin, V., Glantz, H., Neradilek, M., Seydoux, E., Jennewein, M.F., MacCamy, A.J., et al. (2021). mRNA vaccination boosts cross-variant neutralizing antibodies elicited by SARS-CoV-2 infection. *Science*, eabg9175.
- Stavnezer, J., Guikema, J.E.J., and Schrader, C.E. (2008). Mechanism and regulation of class switch recombination. *Annu. Rev. Immunol.* *26*, 261–292.
- Stephens, D.S., and McElrath, M.J. (2020). COVID-19 and the path to immunity. *JAMA* *324*, 1279–1281.
- Stephenson, E., Reynolds, G., Botting, R.A., Calero-Nieto, F.J., Morgan, M.D., Tuong, Z.K., Bach, K., Sungnak, W., Worlock, K.B., Yoshida, M., et al. (2021). Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med.* *27*, 904–916.
- Theofilopoulos, A.N., Kono, D.H., and Baccala, R. (2017). The multiple pathways to autoimmunity. *Nat. Immunol.* *18*, 716–724.
- Tipton, C.M., Fucile, C.F., Darce, J., Chida, A., Ichikawa, T., Gregoretti, I., Schieferl, S., Hom, J., Jenks, S., Feldman, R.J., et al. (2015). Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat. Immunol.* *16*, 755–765.

- Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* 302, 575–581.
- Turner, J.S., Zhou, J.Q., Han, J., Schmitz, A.J., Rizk, A.A., Alsoussi, W.B., Lei, T., Amor, M., McIntire, K.M., Meade, P., et al. (2020). Human germinal centres engage memory and naive B cells after influenza vaccination. *Nature* 586, 127–132.
- Turner, J.S., Kim, W., Kalaidina, E., Goss, C.W., Raueo, A.M., Schmitz, A.J., Hansen, L., Haile, A., Klebert, M.K., Pusic, I., et al. (2021). SARS-CoV-2 infection induces long-lived bone marrow plasma cells in humans. *Nature* 595, 421–425.
- Victora, G.D., and Nussenzweig, M.C. (2012). Germinal centers. *Annu. Rev. Immunol.* 30, 429–457.
- Voss, W.N., Hou, Y.J., Johnson, N.v., Delidakis, G., Kim, J.E., Javanmardi, K., Horton, A.P., Bartzoka, F., Paresi, C.J., Tanno, Y., et al. (2021). Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. *Science* 372, 1108–1112.
- Wang, Z., Muecksch, F., Schaefer-Babajew, D., Finkin, S., Viant, C., Gaebler, C., Hoffmann, H.-H., Barnes, C.O., Cipolla, M., Ramos, V., et al. (2021a). Naturally enhanced neutralizing breadth against SARS-CoV-2 one year after infection. *Nature*, 1–10.
- Wang, Z., Schmidt, F., Weisblum, Y., Muecksch, F., Barnes, C.O., Finkin, S., Schaefer-Babajew, D., Cipolla, M., Gaebler, C., Lieberman, J.A., et al. (2021b). mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* 592, 616–622.
- Wang, Z., Lorenzi, J.C.C., Muecksch, F., Finkin, S., Viant, C., Gaebler, C., Cipolla, M., Hoffmann, H.H., Oliveira, T.Y., Oren, D.A., et al. (2021c). Enhanced SARS-CoV-2 neutralization by dimeric IgA. *Sci. Translational Med.* 13, 1555.
- Wardemann, H., Yurasov, S., Schaefer, A., Young, J.W., Meffre, E., and Nussenzweig, M.C. (2003). Predominant autoantibody production by early human B cell precursors. *Science* 301, 1374–1377.
- Wen, W., Su, W., Tang, H., Le, W., Zhang, X., Zheng, Y., Liu, X., Xie, L., Li, J., Ye, J., et al. (2020). Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell Discov.* 6, 1–18.
- Woodruff, M.C., Ramonell, R.P., Nguyen, D.C., Cashman, K.S., Saini, A.S., Haddad, N.S., Ley, A.M., Kyu, S., Howell, J.C., Ozturk, T., et al. (2020). Extra-follicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19. *Nat. Immunol.* 21, 1506–1516.
- Xiong, X., Qu, K., Ciazynska, K.A., Hosmillo, M., Carter, A.P., Ebrahimi, S., Ke, Z., Scheres, S.H.W., Bergamaschi, L., Grice, G.L., et al. (2020). A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat Struct Mol Biol.* 27, 934–941.
- Xu, Z., Zan, H., Pone, E.J., Mai, T., and Casali, P. (2012). Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat. Rev. Immunol.* 12, 517–531.
- Yang, F., Nielsen, S.C.A., Hoh, R.A., Röltgen, K., Wirz, O.F., Haraguchi, E., Jean, G.H., Lee, J.Y., Pham, T.D., Jackson, K.J.L., et al. (2021). Shared B cell memory to coronaviruses and other pathogens varies in human age groups and tissues. *Science* 372, 738–741.



## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Chemicals, peptides, and recombinant proteins</b>		
Spike SARS-CoV-2 protein	(Xiong et al., 2020)	N/A
SARS-CoV-2 neutralisation assay	(Pereyra Gerber et al., 2021)	N/A
SARS-CoV-2 pseudo-neutralisation assay	(Collier et al., 2021)	N/A
<b>Critical commercial assays</b>		
SuperScript IV	Invitrogen	Cat# 18090200
KAPA Dual-Indexed Adapter Kit	KAPA Biosystems	Cat# KK8727, Cat# KK8721
KAPA Hyper Prep Kit	KAPA Biosystems	Cat# KK8502, Cat# KK8503
KAPA Real-time Library Amplification Kit	KAPA Biosystems	Cat# KK2702
<b>Deposited data</b>		
Flow cytometry data: B cell	(Bergamaschi et al., 2021)	Flow Repository:FR-FCM-Z3XQ
Absolute cell count, SARS-CoV-2 specific antibody titers	(Bergamaschi et al., 2021)	<a href="https://www.covid19cellatlas.org/patient/citiid/">https://www.covid19cellatlas.org/patient/citiid/</a>
Supplemental Items	This paper	<a href="https://data.mendeley.com/datasets/fczzz47td2/1">https://data.mendeley.com/datasets/fczzz47td2/1</a> .
BCR repertoire	This paper	<a href="https://www.covid19cellatlas.org/patient/citiid/">https://www.covid19cellatlas.org/patient/citiid/</a>
BCR repertoire	This paper	EGA: EGAD00001008368
<b>Oligonucleotides</b>		
BCR primers used in library generation, See Table S3	(Consortium, Co.-19 M.B.At. et al., 2021)	N/A
<b>Software and algorithms</b>		
R	(Pinheiro et al., 2014)	<a href="https://www.r-project.org">https://www.r-project.org</a>
Immccantation	(Gupta et al., 2015)	<a href="https://immccantation.readthedocs.io/en/stable/">https://immccantation.readthedocs.io/en/stable/</a>
BCR repertoire	(Bashford-Rogers et al., 2019)	<a href="https://github.com/Bashford-Rogers-lab">https://github.com/Bashford-Rogers-lab</a>
CD-HIT	(Fu et al., 2012)	<a href="http://weizhong-lab.ucsd.edu/cd-hit/">http://weizhong-lab.ucsd.edu/cd-hit/</a>
IMG_T	(Brochet et al., 2008)	<a href="http://www.imgt.org/">http://www.imgt.org/</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact Prof Kenneth Smith ([kgcs2@cam.ac.uk](mailto:kgcs2@cam.ac.uk)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- The datasets generated during this study are available at NIHR CITIID COVID-19 Cohort (<https://www.covid19cellatlas.org/patient/citiid/>). In addition, sequence data have been deposited at the European Genome -phenome Archive (EGA) which is hosted by the EBI and the CRG under accession number EGAD00001008368 and flow cytometry data are available at FLOW Repository: FR-FCM-Z3XQ, FR-FCM-Z3SR, FR-FCM-Z3ST, FR-FCM-Z3SS.
- This paper does not report original code, all code can be downloaded from <https://github.com/Bashford-Rogers-lab>.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

#### COVID-19

This cohort has been previously described by Bergamaschi et al. (Bergamaschi et al., 2021). Briefly, study participants were recruited between 31/3/2020 and 20/7/2020 from patients attending Addenbrooke's Hospital, Royal Papworth Hospital NHS Foundation Trust

or Cambridge and Peterborough Foundation Trust with a confirmed diagnosis of COVID-19, together with Health Care Workers identified through staff screening as PCR positive for SARS-CoV-2 (Rivett et al., 2020). Controls were recruited among hospital staff attending Addenbrooke's for SARS-CoV-2 serology screening program and having a negative serology result. Ethical approval was obtained from the East of England – Cambridge Central Research Ethics Committee (“NIHR BioResource” REC ref 17/EE/0025, and “Genetic variation AND Altered Leukocyte Function in health and disease - GANDALF” REC ref 08/H0308/176). All participants provided informed consent.

Inpatients were sampled at study entry, and then at regular intervals as long as they remained admitted to hospital (approximately weekly up to 4 weeks, and then every 2 weeks up to 12 weeks). Discharged patients were invited to provide a follow-up sample 4–8 weeks after study enrollment. Health care workers were sampled at study entry, and subsequently after approximately 2 and 4 weeks.

### SARS-CoV-2 vaccination

Community participants or health care workers receiving the first dose of the BNT162b2 vaccine between the 14th of December 2020 to the 29th of January 2021 were consecutively recruited at Addenbrookes Hospital into the COVID-19 cohort of the NIHR Bioresource. The study was approved by the East of England – Cambridge Central Research Ethics Committee (17/EE/0025).

### Influenza vaccination

Community participants receiving a dose of Adjuvanted Trivalent Influenza Vaccine (Surface Antigen, Inactivated) Adjuvanted with MF59C.1 (2020/2021 SEASON) were recruited. The study was approved by the East of England – Cambridge Central Research Ethics Committee (REC ref: 20/SW/0134, IRAS id: 287,814, CBR#: 213)

### Healthy controls

Healthy control samples were obtained under the Gandalf ethics (08/H0308/176) and the BioResource Study (150).

## METHOD DETAILS

### Clinical data collection

Clinical data were retrospectively collected by review of medical charts and extraction of data (laboratory test results, vital signs, medications) from Epic electronic health records (Addenbrooke's Hospital) and from MetaVision ICU (Royal Papworth Hospital).

### Study volunteers were classified in 5 groups

- Group A: health care workers who were asymptomatic at the time of positive SARS-CoV-2 testing. This group included 10 volunteers who had possible COVID-19 symptoms before PCR testing (median time from symptoms to COVID-19 PCR test 26 days, range 9–42 days).
- Group B: health care workers who had possible COVID-19 symptoms at the time of PCR testing.
- Group C: patients in hospital who did not receive any supplemental oxygen for COVID-19. Five patients were discharged soon after initial diagnosis and assessment but followed up as part of the study.
- Group D: patients in hospital who received supplemental oxygen using low flow nasal prongs, simple face mask, Venturi mask or non re-breather face mask.
- Group E: patients in hospital who received any of non-invasive ventilation (NIV), mechanical ventilation or ECMO. Patients who received supplemental oxygen (but no ventilation) and deceased in hospital were also assigned to group E.

Study results were analyzed according to time since onset of COVID-19 symptoms, or otherwise time since positive SARS-CoV-2 testing (in group A and in 4 asymptomatic patients in group C).

### Peripheral blood mononuclear cell preparation and flow cytometry immunophenotyping

For direct enumeration of T, B and NK cells, an aliquot of whole blood EDTA (50 $\mu$ L) was added to BD TruCount™ tubes with 20 $\mu$ L BD Multitest™ 6-colour TBNK reagent (BD Biosciences) and processed as per the manufacturer's instructions.

Peripheral venous blood (up to 27 mL per sample) for isolation of Peripheral Blood Mononuclear Cells (PBMCs) was collected into 10% sodium citrate tubes. PBMCs were isolated using Leucosep tubes (Greiner Bio-One) with Histopaque 1077 (Sigma) by centrifugation at 800x g for 15 min at room temperature. PBMCs at the interface were collected, rinsed twice with autoMACS running buffer (Miltenyi Biotech) and cryopreserved in FBS with 10% DMSO. All samples were processed within 4 h of collection.

Approximately 10<sup>6</sup> cells have been stained with: anti-human IgM (clone: G20-127, BD), CD19 (clone: SJ25C1, BD), CD38 (clone: HIT2, BD), IgD (clone: IA6-2, BD), CD20 (clone: 2H7, BD), CD3 (clone: UCHT1, BioLegend), CD14 (clone: 63D3, BioLegend), CD15 (clone: W6D3, BioLegend), CD193 (clone: 5 × 10<sup>8</sup>, BioLegend), CD27 (clone: O323, BioLegend), CD56 (clone: MEM188, Thermo), CD24 (clone: ML5, BD), IgA (polyclonal goat IgG, Jackson), IgG (clone: G18-145, BD), and Zombie Yellow (BioLegend) as described in detail by Bergamaschi et al. (Bergamaschi et al., 2021) Samples were stored at 4°C and acquired within 4 h using a 5-laser BD

Symphony X-50 flow cytometer. Single color compensation tubes (BD CompBeads) or cells were prepared for each of the fluorophores used and acquired at the start of each flow cytometer run.

Samples were gated in FlowJo v10.2 and number of cells falling within each gate was recorded. For analysis, these were expressed either as proportion of total B cells or an absolute concentration of cells per  $\mu\text{L}$ , calculated using the proportions of daughter populations present within the parent population determined using the BD TruCount™ system.

### Total immunoglobulin levels

Serum immunoglobulin levels were measured using the standard assay by the Immunology Department at Peterborough City hospital.

### SARS-CoV-2 serology COVID-19

Quantification of Spike SARS-CoV-2 specific antibodies was performed by ELISA as described by Xiong X et al. (Xiong et al., 2020). Briefly, serum samples collected at time of enrollment in the study and at the 4–8 weeks follow-up visit were first screened for positivity and then antibody titers were determined by an endpoint analysis. AUC values were calculated in R (3.6.3) using the flux (0.3–0) package. Kruskal–Wallis test was used to calculate p values among the different disease severities.

### SARS-CoV-2 vaccine recipients

SARS-CoV-2 serology by multiplex particle-based flow cytometry (Luminex). Recombinant SARS-CoV-2 N, S and RBD were covalently coupled to distinct carboxylated bead sets (Luminex; Netherlands) to form a 3-plex and analyzed as previously described (Collier et al., 2021). Specific binding was reported as mean fluorescence intensities (MFI).

### SARS-CoV-2 neutralisation assay

The virus used in this study was the clinical isolate SARS-CoV-2/human/Liverpool/REMRQ0001/2020, a kind gift from Ian Goodfellow (University of Cambridge), isolated by Lance Turtle (University of Liverpool) and David Matthews and Andrew Davidson (University of Bristol) (Daly et al., 2020; Patterson et al., 2020). Sera were heat-inactivated at 56°C for 30 min, then frozen in aliquots at 80°C. Neutralising antibody titers at 50% inhibition (NT50s) were measured as previously described (Pereyra Gerber et al., 2021). In brief, HEK293T reporter cells expressing Renilla luciferase (Rluc) and SARS-CoV-2 Papain-like protease-activatable circularly permuted firefly luciferase (FFluc) were seeded in flat-bottomed 96-well plates. The next day, SARS-CoV-2 viral stock (MOI = 0.01) was pre-incubated with a 3-fold dilution series of each serum for 2 h at 37°C, then added to the cells. After 24 h, cells were lysed in Dual-Glo Luciferase Buffer (Promega) diluted 1:1 with PBS and 1% NP-40. Lysates were transferred to white half-area 96-well plates, and infectious virus quantitated as the ratio of FFluc/Rluc activity measured using the Dual-Glo kit (Promega) according to the manufacturer's instructions.

Experiments were conducted in duplicate. To obtain NT50s, FFluc/Rluc ratios were analyzed using the Sigmoidal, 4PL, X is log(-concentration) function in GraphPad Prism. Neutralising activity was considered to be present if a sigmoidal neutralisation curve was observed, with an NT50 greater than the lowest dilution of serum tested (10:1).

### B cell receptor repertoire

#### Library preparation

PBMC were lysed and RNA extracted using Qiagen AllPrep DNA/RNA mini kits and Allprep DNA/RNA Micro kits according to the manufacturer's protocol. RNA was extracted from PAXgenes using Qiagen PAXgene Blood RNA kit. The RNA was quantified using a Qubit. B cell receptor repertoire libraries have been generated using the protocol described by Bashford-Rogers et al. (Bashford-Rogers et al., 2019). Briefly, 200ng of total RNA from PAXgenes/PBMCs (14ul volume) was combined with 1uL 10mM dNTP and 10uM reverse primer mix (Table S3) (Consortium et al., 2021) (2uL) and incubated for 5 min at 70°C. The mixture was immediately placed on ice for 1 min and then subsequently combined with 1uL DTT (0.1 M), 1uL SuperScriptIV (Thermo Fisher Scientific), 4uL SSIV Buffer (Thermo Fisher Scientific) and 1uL RNase inhibitor. The solution was incubated at 50°C for 60 min followed by 15 min inactivation at 70°C. cDNA was cleaned with AMPure XP beads and PCR-amplified with a 5' V-gene multiplex primer mix (Table S3) (Consortium, Co.-19 M.B.At. et al., 2021) and 3' universal reverse primer using the KAPA protocol and the following thermal cycling conditions: 1 cycle (95°C, 5min); 5 cycles (98°C, 20s; 72°C, 30s); 5 cycles (98°C, 15s; 65°C, 30s; 72°C, 30s); 19 cycles (98°C, 15s; 60°C, 30s; 72°C, 30s); 1 step (72°C, 5 min). Sequencing libraries were prepared using Illumina protocols and sequenced using 300-bp paired-end sequencing on a MiSeq.

#### Sequence analysis

Raw reads were filtered for base quality using a median Phred score of  $\geq 32$  (<http://sourceforge.net/projects/quasr/>). Forward and reverse reads were merged where a minimum 8bp identical overlapping region was present. Sequences were retained where over 80% base sequence similarity was present between all sequences with the same barcode. The constant-region allele with highest sequence similarity was identified by 10-mer matching to the reference constant-region genes from the IMGT database. Sequences without complete reading frames and non-immunoglobulin sequences were removed and only reads with significant similarity to

reference IGHV and J genes from the IMGT database using BLAST were retained (for read counts see [Table S1](#)). Immunoglobulin gene use and sequence annotation were performed in IMGT V-QUEST ([Brochet et al., 2008](#)).

Somatic hypermutation levels (including silent and non-silent mutations) per unique IGHV-D-J region per isotype were calculated over the CDR1 and CDR2 region for each individual sample using the `observedMutation` function within the SHazaM package ([Gupta et al., 2015](#)). BCR clones were assigned using the `changeo` package using the single-nucleotide Hamming distance model ([Gupta et al., 2015](#)). Lineage trees were generated using the `buildPhyloLineage` function within the Alakazam package after merging sequences from paired time points ([Gupta et al., 2015](#)). VDJtools was used to analyze the BCR sequencing data for diversity estimation of CDR3 sequences (Chao1); the diversity estimates were adjusted for sequencing depth via subsampling with 2,000 random iterations ([Shugay et al., 2015](#)). Convergent IGH clones among different individuals and the CoV-AbDab database ([Raybould et al., 2020](#)) were identified based on having the same CDR-H3 length with a minimum 85% CDR-H3 amino acid sequence identity. CDR-H3 amino acid sequence clustering was performed using CD-HIT ([Fu et al., 2012](#)) with options `-c 0.85 -L 4 -S 0 -g 1 -b 1`.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical tests were performed in R using Wilcoxon tests for significance (non-parametric test of differences between distributions). Longitudinal mixed modeling of BCR repertoire changes over time ( $y_{ij}$ ) was conducted using the `nlme` package in R (<https://cran.r-project.org/web/packages/nlme/>) including time with a linear trend and disease severity category as fixed effects, and sampled individuals as random effects.