




# Imprints of somatic hypermutation on B-cell receptor immunoglobulins post-infection versus post-vaccination against SARS-CoV-2

Elisavet Vlachonikola <sup>1</sup>, Nikolaos Pechlivanis<sup>1,2</sup>, Georgios Karakatsoulis<sup>1</sup>, Massimo Degano<sup>3,4</sup>, Fotis Psomopoulos <sup>1</sup>, Andrea Crisanti<sup>5</sup>, Giovanni Tonon <sup>4,6</sup>, Paolo Ghia<sup>4,7</sup>, Kostas Stamatopoulos<sup>1,8,\*</sup>, Enrico Lavezzo<sup>5,††</sup>, and Anastasia Chatzidimitriou<sup>1,8,††</sup>

<sup>1</sup>Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece

<sup>2</sup>Department of Genetics, Development and Molecular Biology, School of Biology, Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>3</sup>Biocrystallography Unit, Division of Immunology, Transplantation, and Infectious Diseases, IRCCS Scientific Institute San Raffaele, Milan, Italy

<sup>4</sup>Università Vita-Salute San Raffaele, Milan, Italy

<sup>5</sup>Department of Molecular Medicine, University of Padova, Padova, Italy

<sup>6</sup>Center for Omics Sciences, IRCCS Ospedale San Raffaele, Milan, Italy

<sup>7</sup>Division of Experimental Oncology, IRCCS Ospedale San Raffaele, Milan, Italy

<sup>8</sup>Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

\*Corresponding author: Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece, 6th km Charilaou-Thermi Road, Thermi, Thessaloniki 57001, Greece. Email: kostas.stamatopoulos@certh.gr

††Co-senior authors

## Abstract

Published evidence supports significant heterogeneity of immune responses among individuals infected with or vaccinated against SARS-CoV-2. This highlights the need for in-depth investigation of the implicated processes toward refined understanding and improved management of COVID-19. The main objective of the present study was to investigate the dynamics of B cell responses to SARS-CoV-2, focusing on how initial infection and subsequent vaccination influence the immunoglobulin gene repertoire, with special emphasis on the impact of somatic hypermutation (SHM) on antibody maturation. Samples were collected from 81 individuals infected by SARS-CoV-2 in the municipality of Vo' during the first pandemic wave in 2020. For 25 of them, sampling was repeated 7 d after completing the primary vaccination series. Deep immunogenetic analysis of the B-cell receptor immunoglobulin (BcR IG) gene repertoire was performed using targeted next-generation sequencing. Bioinformatics analysis focused on repertoire metrics, prediction of IG antigen specificity, and detailed profiling of the SHM patterns. Significant expansions of unmutated sequences early post-infection suggest extrafollicular B cell maturation. In contrast, vaccination promoted SHM acquisition, indicating a germinal center-dependent response, and pronounced repertoire renewal. Restricted SHMs in SARS-homologous clonotypes along with preferential targeting of specific codons within the VH domain post-vaccination support ongoing affinity maturation within germinal centers. Differences in the BcR IG profiles post-infection versus post-vaccination allude to distinct trajectories in B cell maturation. Distinct profiles of SHM targeting reflect ongoing affinity maturation post-vaccination, with implications for optimizing preventive and therapeutic interventions against COVID-19.

**Keywords:** adaptive immunity, B-cell receptor gene repertoire, COVID-19, somatic hypermutation

## Introduction

Immune responses in patients with COVID-19 show significant heterogeneity, likely reflected in diverse clinical manifestations ranging from no/mild symptoms to severe respiratory failure and multiorgan failure leading to death.<sup>1</sup> Adding to this complexity, varying levels of acquired protection from vaccination have also been observed.<sup>2</sup> This underscores the need to fully understand the immune mechanisms underlying these diverse outcomes in order to improve the management of COVID-19 and enhance immunization strategies against SARS-CoV-2.

Recent studies have outlined distinct trajectories in B-cell maturation and differentiation upon SARS-CoV-2 infection. Evidence suggests that a primary SARS-CoV-2 infection bypasses the normal germinal center (GC) reaction, which typically generates long-lived plasma cells, favoring instead

an extrafollicular reaction in which B cells rapidly proliferate and differentiate into short-lived plasma cells.<sup>3</sup> Contradictory results have been reported regarding the neutralizing capacity and affinity of these early-secreted antibodies.<sup>4,5</sup> Moreover, the molecular imprints of extrafollicular maturation on the B-cell receptor immunoglobulin (BcR IG) gene repertoire have not been conclusively determined.<sup>3,6</sup>

Early observations on the declining neutralizing antibody responses within the first year post-infection and the role of memory B-cell responses for stronger and longer-lasting protection during reinfection supported the design of vaccination strategies.<sup>7</sup> However, the relatively short period of time since the onset of the SARS-CoV-2 pandemic precludes definitive conclusions from being drawn regarding the optimal protection strategies. This explains why the debate continues on whether vaccinating previously infected individuals could

Received: April 7, 2025. Accepted: April 8, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of The American Association of Immunologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

further reduce the risk of reinfection and transmission. Therefore, understanding the temporal dynamics and the B-cell response molecular landscape both post-infection and post-vaccination is crucial for the future design of more effective immunization strategies, toward a better long-term control of COVID-19.

Here, utilizing next-generation sequencing (NGS), we investigated the BcR IG gene repertoire profiles in a unique cohort of cases infected by SARS-CoV-2 during the first pandemic wave in spring 2020 and subsequently vaccinated against the virus.<sup>8</sup> We focused on clonality patterns, immunoglobulin gene usage, and SARS-CoV-2–biased somatic hypermutation (SHM) signatures after the first encounter with SARS-CoV-2 and a subsequent round of vaccination. We report the existence of expanded clonotypes with minimal/no SHM post-infection, indicative of extrafollicular maturation. Moreover, we show that post-infection vaccination against SARS-CoV-2 following a prior infection caused a significant renewal of the BcR IG gene repertoire. Notably, post-vaccination SARS-homologous BcR IG clonotypes were predominantly mutated, showed preferential usage of particular IGHV genes, and carried restricted replacement SHMs, indicative of ongoing affinity maturation in germinal centers. These findings emphasize the dynamic interplay between SHM and immune responses, highlighting the impact of vaccination on SARS-specific B-cell repertoires.

## Materials and methods

### Patients and samples

Samples were collected as part of an epidemiological study investigating the prevalence of SARS-CoV-2 infection in the municipality of Vo', located in the province of Padua, Veneto region, Italy, during February to May 2020.<sup>8</sup> During this period, peripheral blood samples and nasopharyngeal specimens were obtained from the respiratory tract of 81 individuals who were either hospitalized or tested positive for COVID-19 within screening programs. The detection of SARS-CoV-2–positive individuals was performed using quantitative polymerase chain reaction (PCR). Anti-SARS-CoV-2 antibody response was assessed through 3 commercial serology testing kits: LIAISON SARS-CoV-2 S1/S2 IgG (DiaSorin Molecular), Elecsys Anti-SARS-CoV-2 assay (Roche Diagnostics), and ARCHITECT SARS-CoV-2 IgG (Abbott Laboratories).<sup>9</sup> Manufacturer cut-offs were used for determining if samples were positive by each test.

Sampling was repeated in 25 individuals of the study cohort in June 2021, at least 7d after vaccination with either the Pfizer-BioNTech or the Oxford-AstraZeneca COVID-19 vaccine.<sup>10</sup> Participants provided written informed consent before sampling, and all data used in this study were previously pseudonymized. The study was approved by the Ethics Committee for Clinical Research of the province of Padova (protocol numbers 0026971 and 0068830) and conducted in compliance with the ethical principles for medical research stated in the Declaration of Helsinki. A detailed description of the study cohort is given in [Table S1](#).

### NGS of IGHV-IGHD-IGHJ gene rearrangements

BcR IG gene rearrangements were amplified on genomic DNA isolated from peripheral blood mononuclear cells (PBMCs). PCR amplification of IGHV-IGHD-IGHJ gene rearrangements was performed by Platinum Taq DNA

Polymerase High Fidelity (Thermo Fisher Scientific) using IGHV subgroup specific primers annealing to the leader sequence and IGHJ gene primers.<sup>11</sup> Amplicons were purified by excising the PCR product from an agarose gel, followed by gel extraction (Qiagen). Library construction was performed according to the manufacturer's instructions (NEB Next Ultra II DNA Library Prep Kit for Illumina), and NGS was performed with the MiSeq Reagent Kit v3 (2 × 300 bp) on the MiSeq Benchtop Sequencer (Illumina).<sup>12</sup>

Paired-end read merging and strict length/quality filtering were performed by a purpose-built bioinformatics pipeline.<sup>13</sup> Annotation of the IGHV-IGHD-IGHJ gene rearrangements was carried out by the IMGT/HighV-QUEST tool<sup>14</sup> and final meta-data analysis was performed using the TRIP (T cell Receptor/Immunoglobulin Profiler) tool.<sup>15</sup>

For the interpretation of the NGS results, the term *clonotype* was used to describe IGHV-IGHD-IGHJ gene rearrangements with unique pairs of IGHV genes and identical heavy variable complementarity determining region 3 (VH CDR3) amino acid sequences within a sample. The 10 most frequent clonotypes within a sample are herein referred to as “major.”. The relative frequency of each clonotype in a given sample was calculated as the number of IGHV-IGHD-IGHJ gene rearrangements corresponding to this particular clonotype divided by the total number of productive, filtered-in IGHV-IGHD-IGHJ gene rearrangements of that sample. Repertoire clonality was estimated as the median cumulative frequency (MCF-10) of the major clonotypes. In addition, in order to evaluate the relative frequency of each IGHV gene that partakes in the clonotype formation, the number of clonotypes using particular IGHV genes was calculated over the total number of clonotypes.

### Prediction of SARS-homologous BcR IG clonotypes

To predict the SARS-homologous BcR IG clonotypes in the cohort, all published/patented antibodies (9,162 entries) able to bind to SARS-CoV-2 deposited in CoV-AbDab were used to calculate sequence similarities.<sup>16</sup> As a first step, we computed the length of the CDR3 region of all BcR IG sequences between the cohort and database. Subsequently, for CDR3 sequences with identical lengths, we calculated the Hamming distance to gauge their dissimilarity, and distances falling below 85% amino acid sequence similarity were filtered out. The similarity score was calculated following the standardized classes used in IMGT databases, allowing amino acids with common physicochemical properties to be clustered together.<sup>17</sup> BcR IG clonotypes sharing rearrangements of the same IGHV gene subgroup and displaying CDR3 amino acid sequence similarity >85% with a particular CoV-AbDab entry were designated as SARS-homologous.

### Mapping of SHM on BcR IG structures

The structures of representative fragment antigen-binding region (Fab) and single-chain fragment variable (scFv) fragments utilizing certain IGHV genes were downloaded from the RCSB Protein Data Bank (RCSB PDB) (<http://www.rcsb.org>)<sup>18</sup> and visualized using Pymol v1.8 (<http://www.pymol.org>). The associated PDB codes were 3FZU, 6XKP, and 7CDI (IGHV3-11); 7L09 and 5IFH (IGHV3-21); 6XYZ (IGHV3-30 and IGHV-30-3); 7LXW (IGHV3-33); 6ZTD (IGHV3-48); 3B9V (IGHV3-66); and 4UT7 (IGHV3-74). Intermolecular contacts within the deposited structures were analyzed using PISA.<sup>19</sup> Mutations were modeled using the

mutation module in Pymol without further energetic optimization.

### Statistical analysis

For the descriptive statistics, frequencies and relative frequencies were used for the categorical variables, with median (interquartile range) for the numeric variables. To compare the clonotype relative frequencies between the mutated and unmutated cases, and between the post-infection and post-vaccination cases, the binomial test was used. The Wilcoxon test was applied to investigating potential differences in the repertoire between the SARS-homologous and non-homologous clonotypes within rearrangements utilizing a specific IGHV gene or carrying CDR3s of identical length. Last, Fisher's exact test was applied for the comparison of relative IGHV gene frequencies in post-infection versus post-vaccination cases. The significance level was set to 5%, and all *P* values were exploratory. To correct for the multiplicity issues, the false discovery rate correction was used. All the analyses were conducted in R (v. 4.3.2; R Foundation for Statistical computing).

## Results

### Distinct BcR IG gene repertoire profiles post-infection versus post-vaccination

In total, 5,267,301 annotated reads were filtered-in and analyzed as productive gene rearrangements, resulting in the identification of 357,252 BcR IG clonotypes (i.e. IGHV-IGHD-IGHJ gene rearrangements with unique pairs of IGHV genes and identical heavy VH CDR3 amino acid sequences within a sample) in all individuals at the 2 sampling time points (median number of clonotypes per sample: 2,996 post-infection and 2,760 post-vaccination). Pairwise comparisons of the post-infection versus post-vaccination BcR IG gene repertoires in 25 individuals with available longitudinal samples revealed similar repertoire diversity at the two time points (median Shannon index: 5.86 post-infection versus 5.68 post-vaccination). Although oligoclonal expansions were evident in both subgroups (Fig. 1A), pairwise comparisons revealed significantly ( $P < 0.0007$ ) higher clonality in the post-infection (estimated as the median cumulative frequency of the 10 major clonotypes of each sample: MCF-10, 29.02% [range: 11.92%–70.87%]) versus the post-vaccination samples (MCF-10, 19.72% [range: 8.34%–49.03%]) (Fig. 1B). Moreover, differences in the frequency of individual IGHV genes were documented in the repertoires of the two time points, with the IGHV3-33, IGHV3-7, IGHV3-21, IGHV3-53, and IGH3-15 genes showing increased frequency post-vaccination ( $P < 0.05$ ) (Fig. 1C). Recombination events between the IGHV3-30 and IGHJ6 genes were among the most frequent at both time points. That said, other IGHV-IGHJ gene rearrangements were significantly more frequent at each time point ( $P < 0.05$ ) (Fig. 1D). Notably, the BcR IG gene repertoires appeared to be completely renewed post-vaccination, with only 302 (0.35%) of 86,067 clonotypes retained over time by 15 of 25 vaccinated individuals.

### Different SHM signatures post-infection versus post-vaccination

For the analysis of the SHM status of the BcR IG repertoires, we categorized BcR IG clonotypes into distinct subgroups: (i) those bearing IGHV genes with <100% identity compared

with the corresponding IGHV gene and allele germline sequence, which we refer to as mutated, and (ii) those displaying no SHM on the rearranged IGHV gene (i.e. displaying 100% IGHV gene and allele germline identity), which we refer to as unmutated.

In the post-infection samples, the majority of clonotypes (median: 78.27% [range: 64.21%–96.88%]) in all samples were classified as mutated. That notwithstanding, while unmutated clonotypes constituted a smaller fraction of the repertoire of unique clonotypes (median: 21.73% [range: 3.12%–35.79%]), they tended to be larger in size (MCF of unmutated clonotypes: 42% [range: 4.2%–72.3%] vs MCF of mutated clonotypes: 57.9% [range: 27.7%–95.8%]) (Fig. 2A).

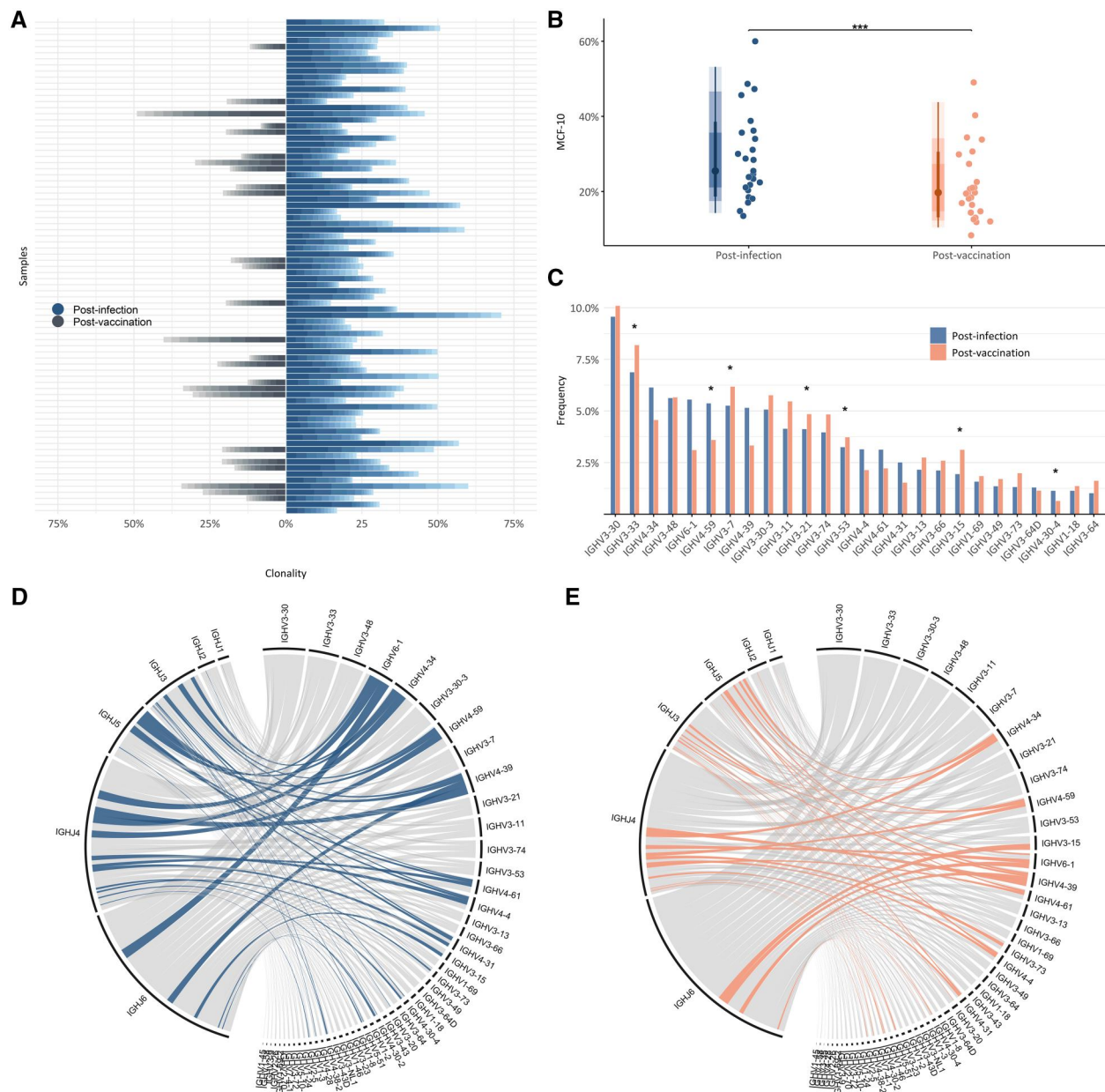
Predominance of mutated clonotypes was also noted in the post-vaccination samples (median: 74.84% [range: 64.17%–95.82%]). However, these samples included significantly less expanded unmutated clonotypes compared with the post-infection samples (MCF: 27.56% post-vaccination vs 42% post-infection;  $P < 0.0001$ ) (Fig. 2B).

### Immunogenetic characterization of SARS-homologous clonotypes

*In silico* prediction of the specificity of BcR IG clonotypes resulted in the identification of 764 unique SARS-homologous clonotypes within the BcR IG gene repertoires of 66 (81.5%) of 81 individuals who tested positive for COVID-19 during the first pandemic wave (Table S2). SARS-homologous clonotypes were identified with varying frequencies, and, among them, 26 (3.4%) ranked in the 100 most frequent clonotypes (Fig. 3A). Additionally, 276 SARS-homologous clonotypes were identified in the BcR IG gene repertoires of 21 (84%) of 25 vaccinated individuals, reflecting a comparable clonotype distribution (3.6% ranked in the 100 most frequent clonotypes). The majority of SARS-homologous clonotypes, specifically 70.5% of post-infection and 45.1% of post-vaccination clonotypes, exhibited a high similarity index to neutralizing antibodies targeting epitopes within the receptor-binding domain of the SARS-CoV-2 spike protein (Fig. 3B). Notably, 70 (9.16%) of 764 post-infection SARS-homologous clonotypes showed a high degree of similarity to therapeutic antibodies developed by Regeneron, derived either from humanized mice or convalescent patients. In sharp contrast, only 5 (1.8%) of 276 post-vaccination SARS-homologous clonotypes, found in different individuals than those post-infection cases, resembled known therapeutic antibodies (Table S3).

### IG gene repertoire

Analysis of the SARS-homologous clonotypes post-infection revealed a strong preference for the usage of particular IGHV genes compared with the non-SARS-homologous clonotypes. In particular, SARS-homologous clonotypes showed enrichment for rearrangements of the IGHV3 gene subgroup, while the remaining clonotypes exhibited a higher prevalence of IGHV4 gene rearrangements ( $P < 0.05$ ). Of note, SARS-homologous clonotypes post-vaccination showed biased usage of the IGHV3-30 gene (median frequency: 22% in SARS-homologous vs 10.19% in the remaining clonotypes;  $P = 0.0002$ ).



**Figure 1.** Comparison of Bcr IG gene repertoire features post-infection versus post-vaccination. (A) Bar plots depict the frequency of the 10 immunodominant (major) clonotypes in the post-infection and post-vaccination Bcr IG gene repertoires, emphasizing the oligoclonal profile of the post-infection repertoire. (B) Density box plots show the distribution of clonality across paired samples at the two time points: clonality was found to be increased post-infection versus post-vaccination ( $P = 0.0007$ ). Clonality was estimated as the MCF-10 and is represented as points, indicating clonality per sample. (C) Bar plots illustrate the frequency of the utilized IGHV genes, revealing differences between the two time points ( $P < 0.05$ ). (D) Circos diagrams depict IGHV-IGHJ gene rearrangements at the two time points, with colored connections indicating statistically significant signatures ( $***P < 0.05$ ).

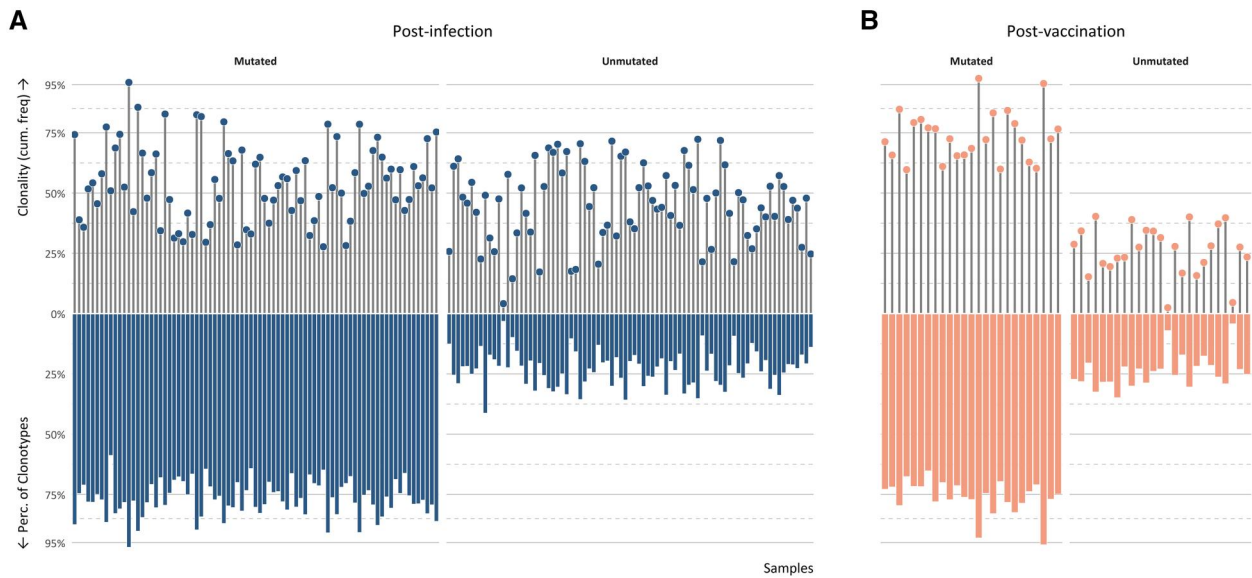
### SHM targeting

A total of 604 (79%) of 769 SARS-homologous clonotypes post-infection were mutated, whereas the remaining 165 (21%) were unmutated. In line with our observations in the entire repertoire, unmutated clonotypes, though fewer in number, were often considerably expanded, thus differing significantly ( $P = 0.0002$ ) (Fig. 4A) from mutated clonotypes, which were predominantly small. Similar to what we found in the entire repertoire, no statistically significant correlation was observed between SHM status and clonal size among post-vaccination SARS-homologous clonotypes (Fig. 4B).

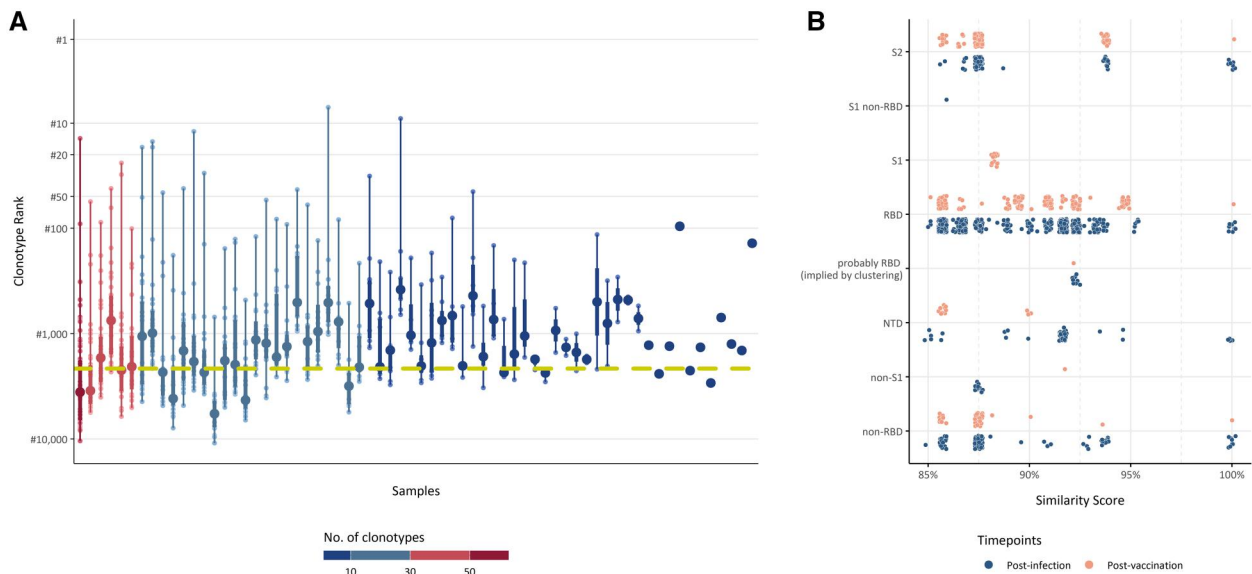
Next, we characterized in detail replacement SHMs toward obtaining hints regarding their potential relevance for

shaping the binding affinity of the Bcr IG clonotypes to SARS-related antigens. When comparing amino acid replacements per position in different SARS-homologous clonotypes that utilized the same IGHV gene, we found that, in many cases, distinct SHMs convergently encoded a particular conserved amino acid substitution. Furthermore, comparisons between post-infection versus post-vaccination SARS-homologous repertoires revealed biases in SHM topology within rearrangements utilizing particular IGHV genes. Specific codons within the VH domain were preferentially targeted at each time point, as shown in Table 1.

Detailed analysis of the topology of nucleotide mutations generated by SHM, and the resulting amino acid substitution



**Figure 2.** SHM status and clonality. Lollipop graphs depict clonality, showing cumulative frequencies of mutated and unmutated clonotypes per sample at the two time points. Bar plots illustrate the percentage of mutated and unmutated clonotypes per sample. The binomial test was employed to assess potential differences in the fraction occupied by the mutated and unmutated clonotypes in post-infection versus post-vaccination repertoires, revealing statistically significant results ( $P < 0.0001$ ) between the two time points.

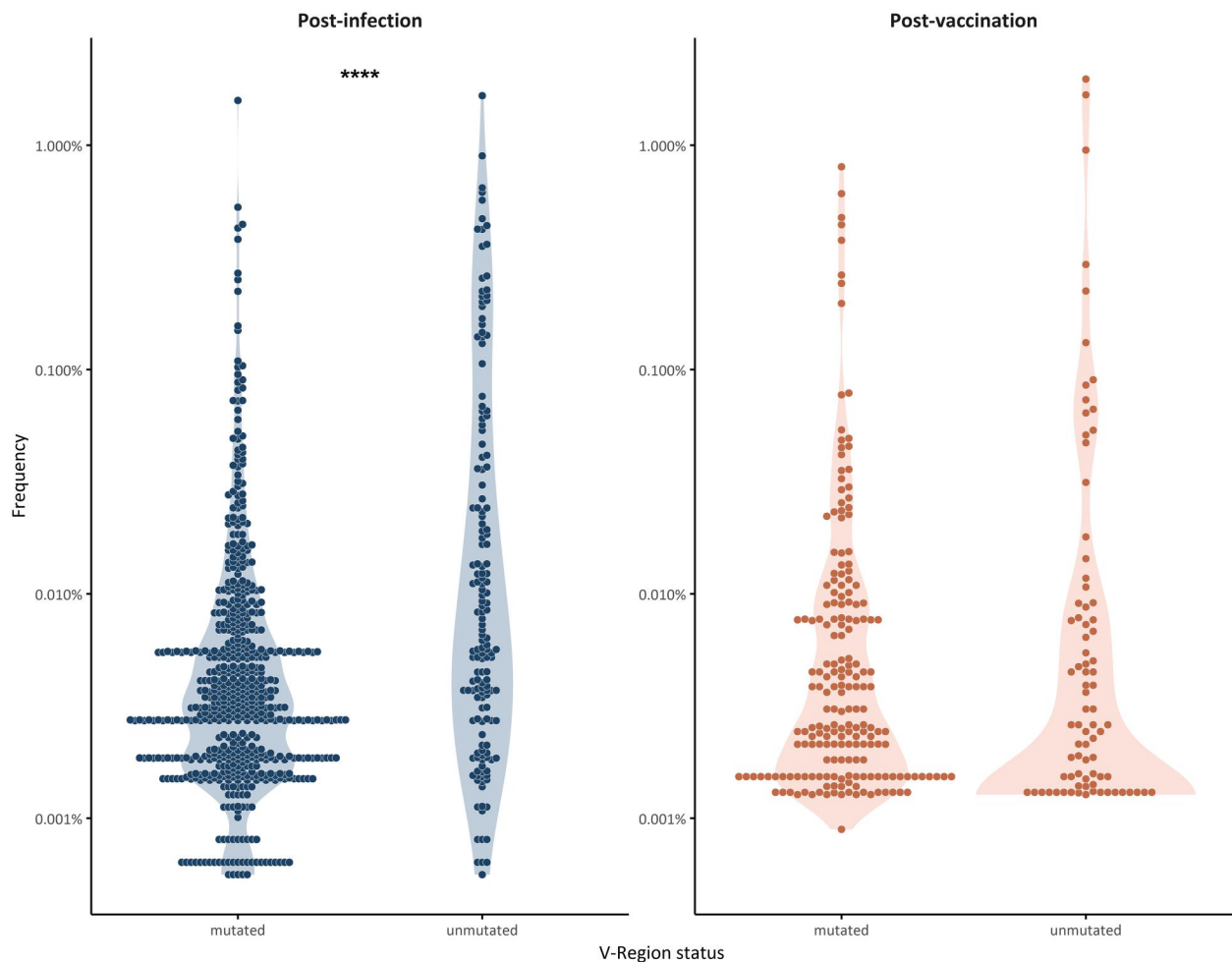


**Figure 3.** SARS-homologous clonotypes. (A) Density box plots illustrate the distributions of post-infection BcR IG gene clonotypes binding to SARS-derived epitopes. The colors represent the number of clonotypes, and the density indicates the percentage of values, with the darkest region encompassing 50% of the observations. (B) Points represent clonotypes that exhibit more than 85% similarity to known neutralizing antibodies in CDR3 amino acid sequences. The colors indicate different time points.

profiles, revealed that not all observed SHMs occurred in hot-spots. Instead, they were equally distributed throughout the variable domain with no significant differences between post-infection versus post-vaccination repertoires. However, in the post-vaccination repertoires we observed a statistically significant enrichment for replacement SHMs at particular positions within the variable domain (Table 1). In 60% of cases, these replacements led to the introduction of an amino acid residue with different physicochemical properties compared with the germline-encoded residue at this particular position.

Next, we sought to assess *in silico* the potential relevance of the observed amino acid substitutions in each

SARS-homologous clonotype. To this end, we took advantage of crystal structures of clonotypes available in the RCSB PDB.<sup>18</sup> In each case (clonotype) from our study, comparisons of the SHM patterns were made against public (RCSB PDB) clonotypes utilizing the same IGHV. These comparisons highlighted particular SHMs with significant predicted effects on IG conformation, shaping the antigen binding affinity and specificity. In general, non-conservative mutations at IMGT positions VH CDR1 26-38 and VH CDR2 56-65 appeared to critically impact their respective conformation, potentially affecting interactions with antigens. Indicatively, the introduction of G27>E by SHM in SARS-CoV-2 specific clonotypes utilizing the IGHV3-30-3



**Figure 4.** Frequency of SARS-homologous clonotypes with respect to SHM status. Points depict the frequency of all mutated and unmutated SARS-homologous clonotypes for the paired samples at (A) post-infection versus (B) post-vaccination, revealing a statistically significant expansion of the unmutated clonotypes post-infection ( $P = 0.0002$ ).

gene after vaccination alters the physicochemical properties of the wild-type residue, introducing a negatively charged amino acid at the beginning of VH CDR1 (Fig. 5). This alteration appears to influence antigen affinity and may even affect specificity compared with the unmutated clonotype with the germline-encoded sequence.

Another noteworthy finding concerned the existence of strings of distinct SHMs consistently co-occurring in rearrangements of a particular IGHV gene. These patterns highlighted the additive or synergistic effect of SHMs, while they differed between the post-infection and post-vaccination repertoires, as shown in Fig. 5. Interestingly, post-infection SARS-homologous clonotypes utilizing the IGHV3-30 gene acquired SHMs with a predicted major influence on critical BcR IG domains, anticipated to impact the final IG conformation and antigen recognition (Fig. 6).

Further comparisons of the amino acid substitution profiles in the BcR IG gene repertoires of individuals analyzed both post-infection as well as post-vaccination revealed two different patterns of SHM acquisition. On the one hand, a particular replacement mutation introduced post-infection increased in frequency post-vaccination (e.g. E51>D and V55>L in SARS-CoV-2 specific clonotypes utilizing the IGHV3-30 gene) (Fig. 7A). On the other hand, certain replacement SHMs introduced post-infection were not detected

post-vaccination; instead, the same position in the VH domain was found to carry a different replacement SHM (e.g. positions VH CDR1-27 and VH FR2-51 in SARS-CoV-2 specific clonotypes utilizing the IGHV3-30-3 gene) (Fig. 7B).

## Discussion

In-depth comprehension of the human antibody response to SARS-CoV-2 is a critical prerequisite for advancing effective immune strategies to restrain COVID-19. Previous immunogenetic studies on the B cell antibody response to infection by SARS-CoV-2 suggested a two-tier response, mainly through characterizing patterns of SHM within the IG genes. The identification of near-germline SARS-CoV-2 neutralizing antibodies, exhibiting no or limited degree SHM, was thought as consistent with a rapid maturation of naive B cells through an extrafollicular maturation pathway.<sup>2</sup> In contrast, studies on the B-cell memory response after the initial infection revealed additional SHMs and clonal expansion over time, consistent with GC maturation.<sup>20</sup> That said, previous studies on single time points, as well as variations in study designs and cohort composition, hinder solid conclusions on the molecular mechanisms of the B cell response to SARS-CoV-2.

In this study, we profiled the BcR IG gene repertoire after the initial infection by the SARS-CoV-2 virus as well as

**Table 1.** Significant amino acid substitutions observed in SARS-homologous clonotypes following infection or vaccination against SARS-CoV-2.

| Gene                               | Position | Reference | Alteration | Region | P value | Type of alteration | Hotspot |
|------------------------------------|----------|-----------|------------|--------|---------|--------------------|---------|
| <b>Post-infection repertoire</b>   |          |           |            |        |         |                    |         |
| IGHV3-30                           | 36       | S         | N          | CDR1   | 0.009   | Non-conservative   | Y       |
| IGHV3-30                           | 64       | N         | K          | CDR2   | 0.011   | Non-conservative   | Y       |
| IGHV3-30                           | 64       | N         | T          | CDR2   | 0.022   | Non-conservative   | Y       |
| IGHV3-30                           | 40       | H         | Y          | FR2    | 0.046   | Non-conservative   | N       |
| IGHV3-30                           | 97       | E         | D          | FR3    | 0.029   | Conservative       | N       |
| IGHV3-30                           | 86       | T         | S          | FR3    | 0.033   | Conservative       | N       |
| IGHV3-33                           | 86       | T         | S          | FR3    | 0.010   | Conservative       | N       |
| <b>Post-vaccination repertoire</b> |          |           |            |        |         |                    |         |
| IGHV3-11                           | 65       | I         | K          | CDR2   | 0.032   | Non-conservative   | Y       |
| IGHV3-11                           | 86       | S         | T          | FR3    | 0.042   | Conservative       | N       |
| IGHV3-21                           | 14       | K         | T          | FR1    | 0.034   | Non-conservative   | Y       |
| IGHV3-21                           | 38       | S         | T          | CDR1   | 0.031   | Conservative       | Y       |
| IGHV3-30                           | 51       | E         | D          | FR2    | 0.007   | Conservative       | N       |
| IGHV3-30                           | 55       | V         | A          | FR2    | 0.048   | Conservative       | Y       |
| IGHV3-30                           | 63       | S         | L          | CDR2   | 0.00004 | Non-conservative   | Y       |
| IGHV3-30                           | 68       | A         | T          | FR3    | 0.046   | Non-conservative   | N       |
| IGHV3-30                           | 69       | D         | T          | FR3    | 0.00008 | Non-conservative   | N       |
| IGHV3-30                           | 77       | T         | S          | FR3    | 0.00001 | Conservative       | N       |
| IGHV3-30                           | 84       | K         | N          | FR3    | 0.00002 | Non-conservative   | N       |
| IGHV3-30-3                         | 3        | Q         | R          | FR1    | 0.032   | Non-conservative   | N       |
| IGHV3-30-3                         | 27       | G         | E          | CDR1   | 0.029   | Non-conservative   | N       |
| IGHV3-30-3                         | 72       | K         | R          | FR3    | 0.003   | Conservative       | Y       |
| IGHV3-30-3                         | 103      | Y         | F          | FR3    | 0.019   | Conservative       | Y       |
| IGHV3-33                           | 24       | A         | S          | FR1    | 0.038   | Non-conservative   | N       |
| IGHV3-33                           | 36       | S         | N          | CDR1   | 0.020   | Non-conservative   | Y       |
| IGHV3-33                           | 55       | V         | A          | FR2    | 0.034   | Conservative       | Y       |
| IGHV3-33                           | 56       | I         | V          | CDR2   | 0.040   | Conservative       | Y       |
| IGHV3-33                           | 72       | K         | E          | FR3    | 0.037   | Non-conservative   | N       |
| IGHV3-48                           | 38       | S         | Y          | CDR1   | 0.047   | Non-conservative   | Y       |
| IGHV3-48                           | 65       | I         | K          | CDR2   | 0.012   | Non-conservative   | Y       |
| IGHV3-48                           | 91       | M         | L          | FR3    | 0.048   | Non-conservative   | Y       |
| IGHV3-66                           | 38       | Y         | D          | CDR1   | 0.017   | Non-conservative   | Y       |
| IGHV3-66                           | 64       | S         | I          | CDR2   | 0.021   | Non-conservative   | N       |
| IGHV3-66                           | 84       | K         | R          | FR3    | 0.023   | Conservative       | Y       |
| IGHV3-66                           | 103      | Y         | F          | FR3    | 0.022   | Conservative       | Y       |
| IGHV3-74                           | 56       | I         | M          | CDR2   | 0.048   | Non-conservative   | Y       |
| IGHV3-74                           | 58       | S         | T          | CDR2   | 0.003   | Conservative       | N       |
| IGHV3-74                           | 96       | A         | T          | FR3    | 0.045   | Non-conservative   | N       |

The term *conservative* denotes substitutions resulting in the replacement of an amino acid by another one possessing similar physicochemical properties, as defined in standardized classes used in IMGT databases.  
N, no; Y, yes.

following a complete primary series of immunization with 2 vaccine doses. A particular strength of the study concerns the fact that the investigated cohort derived from a restricted geographical area in the municipality of Vo', and the sampling was conducted during the first pandemic wave and the initial phase of vaccination strategies,<sup>8-10</sup> thus offering the opportunity to interrogate the naïve immune responses to a novel antigen.

By in-depth immunogenetic analysis of the BcR IG repertoire, we documented that vaccination against SARS-CoV-2 following the initial infection by the virus resulted in a significant turnover of the BcR IG gene repertoire: of note, clonal expansions detected after the initial infection decreased significantly after vaccination. Responses to vaccination exhibited a bias toward BcR IG clonotypes encoded preferentially by IGHV3 subgroup genes, consistent with previous reports on the spike-specific antibody response induced by vaccination.<sup>21,22</sup> Notably, extensive comparisons of post-infection BcR IG repertoire profiles between symptomatic versus asymptomatic cases, hospitalized versus non-hospitalized individuals, and post-vaccination BcR IG repertoires between individuals vaccinated with different vaccine types

(Pfizer-BioNTech vs. Oxford-AstraZeneca), did not reveal any statistically significant differences (data not shown). This underscores the distinct effects that initial infection and subsequent vaccination have on shaping the immune repertoire. Analysis of SHM profiles revealed distinct patterns between post-infection versus post-vaccination, which were found to be independent of age distribution in each group. While most clonotypes carried SHMs at both time points, significant expansions of unmutated clonotypes were seen almost exclusively post-infection. This difference held also when focusing specifically on the predicted SARS-homologous clonotypes. These findings reinforce the concept that the initial anti-SARS-CoV-2 immune response following infection predominantly activates extrafollicular B cell maturation, leading to the generation of near-germline anti-SARS-CoV-2 antibodies,<sup>3,23</sup> while in contrast, vaccine-induced immunization progressively drives the acquisition of SHM, a hallmark of a canonical GC-dependent immune response.<sup>24</sup>

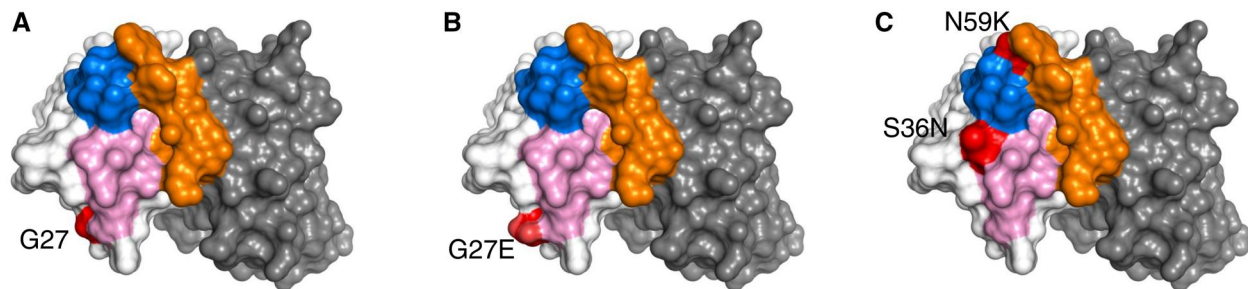
A notable finding of our study concerns the presence of restricted replacement SHMs in SARS-homologous BcR IG clonotypes utilizing the same IGHV gene, alluding to strong antigen selection that favors the introduction of particular

| Post infection mutational profile |  |                            |                     |                             |                      |  |
|-----------------------------------|--|----------------------------|---------------------|-----------------------------|----------------------|--|
| IGHV gene                         | Effect   | FR1                        | CDR1                | FR2                         | CDR2                 | FR3                                    |
| IGHV3-30                          | S36>N and N64>T collectively impact VH CDR1 and CDR2 with an additive effect   | QVQLVESGG.GVVQPGKSLRLSCAAS | GFTF... <b>SNYA</b> | MHWVRRQAPGKGLWEWVAV         | ISYD... <b>GS</b> TK | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC |
| IGHV3-30                          | The mutations influence antigen binding, potentially due to a rearrangement in FR4 and variations in VH CDR1. H40>Y may affect VH-VL chain pairing | QVQLVESGG.GVVQPGKSLRLSCAAS | GFTF... <b>SNYA</b> | <b>MY</b> WVRRQAPGKGLWEWVAV | ISYD...GSNK          | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC |
| IGHV3-30                          | The mutations affect VH CDR1, CDR2 and FR4   | QVQLVESGG.GVVQPGKSLRLSCAAS | GFTF... <b>SNYA</b> | MHWVRRQAPGKGLWEWVAV         | ISYD... <b>GS</b> TK | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC |
| IGHV3-30                          | Direct effect on antigen recognition, as H40>Y affects VH-VL chain pairing and N64>T VH CDR2   | QVQLVESGG.GVVQPGKSLRLSCAAS | GFTF... <b>SSYA</b> | <b>MY</b> WVRRQAPGKGLWEWVAV | ISYD... <b>GS</b> TK | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC |
| IGHV3-30                          | The mutations influence antigen binding, potentially due to changes in the conformation of FR4 and VH CDR2   | QVQLVESGG.GVVQPGKSLRLSCAAS | GFTF... <b>SSYA</b> | MHWVRRQAPGKGLWEWVAV         | ISYD... <b>GS</b> TK | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC |

| Post vaccination mutational profile |  |   |                                     |                             |   |  |
|-------------------------------------|--|---|-------------------------------------|-----------------------------|---|--|
| IGHV gene                           | Effect   | FR1   | CDR1                                | FR2                         | CDR2  | FR3  |
| IGHV3-30                            | S63>L and K84>N may affect direct antigen recognition  | QVQLVESGG.GVVQPGKSLRLSCAAS                  | GFTF... <b>SSYA</b>                 | MHWVRRQAPGKGL <b>W</b> WVAV | ISYD... <b>GL</b> NK                          | <b>YY</b> TT <b>SVK</b> .GRFTISRDNKNTLYLQMNLSRAEDTAVYYC        |
| IGHV3-30-3                          | The mutations may favor the formation of a salt bridge between them, stabilizing a specific VH CDR1 conformation   | <b>QV</b> R <b>L</b> VESGG.GVVQPGKSLRLSCAAS | <b>E</b> FTF... <b>SSYA</b>         | MHWVRRQAPGKGLWEWVAV         | ISYD...GSNK                                   | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC                         |
| IGHV3-33                            | V55>A and I56>V mutations, albeit conservative, may affect the combining site in fine details of the molecular surface for antigen interaction             | QVQLVESGG.GVVQPGKSLRLS <b>S</b> AS          | GFTF... <b>SNYG</b>                 | MHWVRRQAPGKGLWEW <b>A</b>   | <b>V</b> WYD...GSNK                           | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC                         |
| IGHV3-48                            | The mutations likely affect antigen binding directly and through alteration of the VH CDR1 and CDR2 conformation   | EVQLVESGG.GLVQPGKSLRLSCAAS                  | GFTF... <b>SSY</b> <b>Y</b>         | MHWVRRQAPGKGLWEWV <b>SY</b> | ISS <b>S</b> ... <b>S</b> ST <b>K</b>         | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC                         |
| IGHV3-66                            | S64>I, K84>R and Y103F may influence the geometry of the VH-VL chain pairing, while the additional mutation at Y38 may also change the epitope specificity | EVQLVESGG.GLVQPGKSLRLSCAAS                  | GFTV... <b>SS</b> <b>N</b> <b>D</b> | MSWVRRQAPGKGLWEWV <b>S</b>  | YIS <b>G</b> ... <b>G</b> IT                  | YYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAV <b>Y</b> <b>F</b> <b>C</b> |
| IGHV3-74                            | The mutations may affect VH CDR2, perhaps restructuring the combining site   | EVQLVESGG.GLVQPGKSLRLSCAAS                  | GFTF... <b>SSY</b> <b>W</b>         | MHWVRRQAPGKGLWV <b>W</b> SR | <b>M</b> IT <b>D</b> ... <b>G</b> ST <b>S</b> | SYADSVK.GRFTISRDNKNTLYLQMNLSRAEDTAVYYC                         |

**Figure 5.** Significant amino acid substitutions observed in SARS-homologous clonotypes as a string of co-existing SHMs in rearrangements of a particular IGHV gene.



**Figure 6.** Example of localization of frequent mutations post-vaccination or infection by SARS-CoV-2. The structure of a BcR IG encoded by the IGHV3-30 gene (PDB code 6KYZ) is shown through its molecular surface, with the heavy chain in white and the light chain in gray. The HCDR1, HCDR2, and HCDR3 regions are colored pink, blue, and orange, respectively. (A) The position of residue G27 is shown in red. (B) The effect of the mutation to a glutamic acid, a frequent mutation in antibodies from individuals post-vaccination, has a minor effect on the combining site but introduces a negative charge at the edge of the region devoted to antigen binding. (C) The simultaneous introduction of the two Ser36Asn and Asn59Lys mutations observed in antibodies from individuals post-infection affects directly the character of the surface involved in antigen recognition.

changes in IG structure, likely due to their relevance in anti-viral responses. These findings also suggest a dynamic interplay between SHM profiles and the immune response following vaccination, potentially reflecting ongoing affinity maturation processes and selective pressure exerted by the vaccine antigens.

Comparisons between predicted SARS-homologous versus non-homologous clonotypes highlighted a strong preference for particular IGHV genes in the former, with a notable enrichment for IGHV3 subgroup genes. Vaccination led to a significant further enrichment of SARS-homologous clonotypes utilizing IGHV3 subgroup genes, especially IGHV3-30, indicating a targeted response to vaccine antigens. Nevertheless, these results should be interpreted cautiously due to the inherent limitations of the *in silico* prediction software, which relied on the assessment of previously published or patented antibodies submitted to CoV-AbDab, potentially introducing biases into the analysis.<sup>16</sup> Future work should focus on generating

single-cell BcR IG repertoire data to gain insights into VH and VL pairing in SARS-homologous clonotypes, allowing the generation of monoclonal antibodies from IG sequences for experimental characterization of their antigenic specificity and other physicochemical attributes.

In conclusion, this study underscores the intricate dynamics of the B cell response to SARS-CoV-2, delineating how initial infection and subsequent vaccination uniquely shape the immune repertoire, particularly through the introduction of precisely targeted SHM. Future research should focus on further refining these immunogenetic insights to optimize vaccine design and therapeutic interventions against COVID-19.

## Acknowledgments

The authors thank the population of Vo' who participated to multiple serosurveys follow up and acknowledge Fondazione Umberto Veronesi, Misura Ricerca Covid 19, year 2020.



**Figure 7.** Mutational analysis of SARS-homologous clonotypes. Stacked bar plots illustrate amino acid substitutions resulting from SHM on the FR1–FR3 part of the VH domain of all SARS-homologous IG clonotypes sharing the same IGHV gene in the post-infection and post-vaccination BcR IG gene repertoires. The x-axis letters represent the amino acids encoded by the germline gene sequence and the letters in red depict mutational hotspots, while the named bars indicate amino acid substitutions with a frequency exceeding 5%. For the sake of clarity, SHMs called by a single sequencing read were excluded from the figure. The color code employed to describe the physicochemical properties of each amino acid per position is based on the IMGT database classification: dark blue for aliphatic (A, V, I, L, G, P), yellow for sulfur (C, M), light blue for hydroxyl (S, T), red for acidic (D, E), orange for amide (N, Q), purple for basic (H, K, R), and green for aromatic (F, W, Y) amino acids.

## Author contributions

E.V. designed and conducted the experimental work, analyzed the produced data, prepared the figures, and drafted the manuscript. N.P. developed the software for computational analysis and contributed to figure creation. G.K. performed the statistical analysis. M.D. modeled and analyzed SHM profiles on crystal structures and created the corresponding figure. F.P. provided guidance on computational analysis. G.T. offered guidance and secured funding. P.G. advised on the experimental design and analysis. A.C. provided guidance and funding support. E.L. coordinated sample collection at the different time points and secured funding. K.S.

guided the experimental design and analysis, supervised the project, and critically reviewed the manuscript. All authors reviewed and approved the final manuscript.

Elisavet Vlachonikola (Data curation [Equal], Formal analysis [Lead], Investigation [Lead], Methodology [Lead], Resources [Lead], Software [Supporting], Validation [Lead], Visualization [Equal], Writing—original draft [Lead], Writing—review & editing [Equal]), Nikolaos Pechlivanis (Data curation [Equal], Software [Lead], Visualization [Lead], Writing—review & editing [Supporting]), Georgios Karakatsoulis (Data curation [Equal], Validation [Equal], Writing—review & editing [Equal]), Massimo Degano (Formal

analysis [Supporting], Investigation [Supporting], Methodology [Supporting], Visualization [Equal], Writing—review & editing [Equal]), Fotis Psomopoulos (Supervision [Supporting], Writing—review & editing [Supporting]), Andrea Crisanti (Funding acquisition [Supporting], Supervision [Equal], Writing—review & editing [Equal]), Giovanni Tonon (Conceptualization [Supporting], Funding acquisition [Lead], Project administration [Equal], Resources [Supporting], Supervision [Equal], Writing—review & editing [Supporting]), Paolo Ghia (Conceptualization [Supporting], Supervision [Supporting], Writing—review & editing [Supporting]), Kostas Stamatopoulos (Conceptualization [Equal], Funding acquisition [Lead], Project administration [Lead], Resources [Lead], Supervision [Equal], Writing—review & editing [Lead]), Enrico Lavezzo (Conceptualization [Equal], Funding acquisition [Lead], Project administration [Lead], Resources [Supporting], Supervision [Supporting], Writing—review & editing [Supporting]), and Anastasia Chatzidimitriou (Funding acquisition [Lead], Resources [Equal], Supervision [Equal], Writing—review & editing [Supporting])

## Supplementary material

Supplementary material is available at *ImmunoHorizons* online.

## Funding

This work was supported by the Veneto region; the Fondazione Umberto Veronesi, Misura Ricerca COVID-19, year 2020; the EDCTP2 program supported by the European Union; the University of Padova and the Department of Molecular Medicine (STARS-CoG ISS-MYTH and PRID/SID 2020 [EL]); and the Greece versus Corona Action of the Hellenic Secretariat of Research and Innovation.

## Conflicts of interest

The authors declare no relevant conflict of interest.

## Data availability

Data is available at the Sequence Read Archive, National Center for Biotechnology Information, under accession number PRJNA1191651. The dataset can be accessed at <http://www.ncbi.nlm.nih.gov/bioproject/1191651>.

## References

- Mangge H et al. Immune responses against SARS-CoV-2—questions and experiences. *Biomedicines*. 2021;9:1342. <https://doi.org/10.3390/biomedicines9101342>
- Chen S et al. The role of B cells in COVID-19 infection and vaccination. *Front Immunol*. 2022;13:988536. <https://doi.org/10.3389/fimmu.2022.988536>
- Woodruff MC et al. Extrafollicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19. *Nat Immunol*. 2020;21:1506–1516. <https://doi.org/10.1038/s41590-020-00814-z>
- Gaebler C et al. Evolution of antibody immunity to SARS-CoV-2. *Nature*. 2021;591:639–644. <https://doi.org/10.1038/s41586-021-03207-w>
- Robbiani DF et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature*. 2020;584:437–442. <https://doi.org/10.1038/s41586-020-2456-9>
- van Schouwenburg P et al. Deciphering imprints of impaired memory B-cell maturation in germinal centers of three patients with common variable immunodeficiency. *Front Immunol*. 2022;13:959002. <https://doi.org/10.3389/fimmu.2022.959002>
- Stein C et al. Past SARS-CoV-2 infection protection against reinfection: a systematic review and meta-analysis. *Lancet*. 2023;401:833–842. [https://doi.org/10.1016/S0140-6736\(22\)02465-5](https://doi.org/10.1016/S0140-6736(22)02465-5)
- Lavezzo E et al.; Imperial College COVID-19 Response Team. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature*. 2020;584:425–429. <https://doi.org/10.1038/s41586-020-2488-1>
- Dorigatti I et al. SARS-CoV-2 antibody dynamics and transmission from community-wide serological testing in the Italian municipality of Vo'. *Nat Commun*. 2021;12:4383. <https://doi.org/10.1038/s41467-021-24622-7>
- Lavezzo E et al. Neutralising reactivity against SARS-CoV-2 delta and omicron variants by vaccination and infection history. *Genome Med*. 2022;14:61. <https://doi.org/10.1186/s13073-022-01066-2>
- Agathangelidis A et al. Immunoglobulin gene analysis in chronic lymphocytic leukemia. In: Malek SN, editor. *Chronic lymphocytic leukemia: methods and protocols*. Springer; 2019. p. 51–62. [https://doi.org/10.1007/978-1-4939-8876-1\\_5](https://doi.org/10.1007/978-1-4939-8876-1_5)
- Gemenetzi K et al. Higher-order immunoglobulin repertoire restrictions in CLL: the illustrative case of stereotyped subsets 2 and 169. *Blood*. 2021;137:1895–1904. <https://doi.org/10.1182/blood.2020005216>
- Vardi A et al. Restrictions in the T-cell repertoire of chronic lymphocytic leukemia: high-throughput immunoprofiling supports selection by shared antigenic elements. *Leukemia*. 2017;31:1555–1561. <https://doi.org/10.1038/leu.2016.362>
- Alamyar E, Duroux P, Lefranc M-P, Giudicelli V. IMGT® tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol Biol*. 2012;882:569–604. [https://doi.org/10.1007/978-1-61779-842-9\\_32](https://doi.org/10.1007/978-1-61779-842-9_32)
- Kotouza MT, et al. Hellenic precision medicine network in oncology. TRIP—T cell receptor/immunoglobulin profiler. *BMC Bioinform*. 2020;21:422. <https://doi.org/10.1186/s12859-020-03669-1>
- Raybould MIJ, Kovaltuk A, Marks C, Deane CM. CoV-AbDab: the coronavirus antibody database. *Bioinformatics*. 2021;37:734–735. <https://doi.org/10.1093/bioinformatics/btaa739>
- Pommié C, Levadoux S, Sabatier R, Lefranc G, Lefranc M. IMGT standardized criteria for statistical analysis of immunoglobulin V-REGION amino acid properties. *J Mol Recognit*. 2004;17:17–32. <https://doi.org/10.1002/jmr.647>
- Berman HM et al. The protein data bank. *Nucleic Acids Res*. 2000;28:235–242. <https://doi.org/10.1093/nar/28.1.235>
- Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007;372:774–797. <https://doi.org/10.1016/j.jmb.2007.05.022>
- Kreer C et al. Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell*. 2020;182:843–854.e12. <https://doi.org/10.1016/j.cell.2020.06.044>
- Zhang Y et al. Analysis of B cell receptor repertoires reveals key signatures of the systemic B cell response after SARS-CoV-2 infection. *J Virol*. 2022;96:e0160021. <https://doi.org/10.1128/jvi.01600-21>
- Wang Y et al. A large-scale systematic survey reveals recurring molecular features of public antibody responses to SARS-CoV-2. *Immunity*. 2022;55:1105–1117.e4. <https://doi.org/10.1016/j.immuni.2022.03.019>
- Sokal A et al. Maturation and persistence of the anti-SARS-CoV-2 memory B cell response. *Cell*. 2021;184:1201–1213.e14. <https://doi.org/10.1016/j.cell.2021.01.050>
- Nielsen SCA et al. Human B cell clonal expansion and convergent antibody responses to SARS-CoV-2. *Cell Host Microbe*. 2020;28:516–525.e5. <https://doi.org/10.1016/j.chom.2020.09.002>

© The Author(s) 2025. Published by Oxford University Press on behalf of The American Association of Immunologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

ImmunoHorizons, 2025, 9, 1–10

<https://doi.org/10.1093/immhor/vlaf021>

Research Article